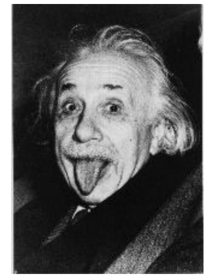


Lecture Notes on General Relativity

Matthias Blau

Albert Einstein Center for Fundamental Physics
Institut für Theoretische Physik
Universität Bern
CH-3012 Bern, Switzerland

The latest version of these notes is available from

`http://www.blau.itp.unibe.ch/GRLecturenotes.html`

`Last update April 29, 2024`

Contents

0 Introduction

1905 was Einstein's magical year. In that year, he published three articles, on light quanta, on Brownian motion, and on the foundations of the theory of Special Relativity (and, almost as an afterthought, a short note containing a first derivation of the iconic $E = mc^2$), each one separately worthy of a Nobel prize.

Immediately after his work on Special Relativity, Einstein started thinking about gravity and how to give it a relativistically invariant formulation. He kept on working on this problem during the next ten years, doing little else. This work, after many trials and errors, culminated in his masterpiece, the *General Theory of Relativity*, presented in 1915/1916. It is widely considered to be one of the greatest scientific and intellectual achievements of all time, a beautiful theory derived from pure thought and physical intuition, capable of explaining, or at least describing, still today, more than 100 years later, every aspect of gravitational physics ever observed.

Einstein's key insight was what is now known as the *Einstein Equivalence Principle*, the (local) equivalence of gravitation and inertia. This ultimately led him to the realisation that gravity is best described and understood not as a physical external force like the other forces of nature but rather as a manifestation of the *geometry and curvature of space-time* itself. This realisation, in its simplicity and beauty, has had a profound impact on theoretical physics as a whole, and Einstein's vision of a geometrisation of all of physics is still with us today.

The aim of these lecture notes is to provide a reasonably self-contained introduction to General Relativity, including a variety of applications of the theory, ranging from the solar system to gravitational waves, black holes and cosmology.

These lecture notes for an introductory course on General Relativity are based on a course that I originally gave in the years 1998-2003 in the framework of the Diploma Course of the ICTP (Trieste, Italy). Currently these notes form the basis of a course that I teach as part of the Master in Theoretical Physics curriculum at the University of Bern.

In the intervening years, I have made (and keep making) various additions to the lecture notes, and they now include much more material than is needed for (or can realistically be covered in) an introductory 1- or even 2-semester course, say, but I hope to have nevertheless preserved (at least in parts) the introductory character and accessible style of the original notes.

Invariably, any set of (introductory) lecture notes has its shortcomings, due to lack of space and time, the requirements of the audience and the expertise (or lack thereof) and interests of the lecturer. These lecture notes are, of course, no exception. In particular, the emphasis in these notes is on developing the theory (I am a theoretical physicist), not on experiments or connecting the theory with observation, but stops short of doing real mathematical general relativity (i.e. proving theorems), as this would require significantly more mathematical sophistication and machinery than I want to assume (or can develop) in these notes. I hope that these lecture notes nevertheless provide the necessary background for studying these or other more advanced topics not covered in these notes.

I should also stress that I have written these notes primarily for myself, and for my students. I am making them publicly available just in case somebody else happens to find them useful, and because I know that previous versions of these notes have enjoyed some popularity. However, if you do not like these notes or my way of explaining things, or do not find what you are looking for, please do not complain to me (yes, this has happened in the past). There will occasionally be further additions and updates to these notes, reflecting however my personal preferences and taste rather than any (futile) aim for completeness.

Lecture notes of this length unavoidably contain some minor mistakes somewhere. However, I hope that these notes are free of major conceptual errors and blunders. I am of course grateful for any constructive criticism and corrections. If you have such comments, or also if you just happen to find these notes useful, please let me know (`blau at itp.unibe.ch`).

In these notes, the pronoun "we" is used to refer to the author along with you, the reader (whereas, as you may have already noticed, I unashamedly use "I" to refer to myself, the author - no *pluralis auctoris* or *pluralis modestiae*, let alone a *pluralis maiestatis* . . .).

0.1 Prerequisites

General Relativity may appear to be a difficult subject at first, since it requires a certain amount of new mathematics and takes place in an unfamiliar arena. However, this course is meant to be essentially self-contained, requiring only a familiarity with

- Special Relativity,
- Lagrangian mechanics,
- vector calculus and calculus in $\mathbb{R}^n$.

To be precise, by special relativity I mean the covariant formulation in terms of the Minkowski metric and Lorentz tensors etc.; special relativity is (regardless of what you may have been taught) not fundamentally a theory about people changing trains erratically, running into barns with poles, or doing strange things to their twins; rather, it is a theory of a fundamental symmetry principle of physics, namely that the laws of

physics are invariant under Lorentz transformations and that they should therefore also be formulated in a way which makes this symmetry manifest.
[Litmus Test: does the content of section 1.2 look familiar to you?]

I will thus attempt to explain every single other thing that is required to understand the basics of Einstein's theory of gravity. However, this also means that I will not be able to discuss some mathematically more advanced and yet equally important aspects of General Relativity.

0.2 Overview

Currently, these notes are organised into 7 parts, namely

A: Physics in a Gravitational Field and Tensor Calculus

B: General Relativity and Geometry

C: Dynamics of the Gravitational Field

D: General Relativity and the Solar System

E: Black Holes

F: Cosmology

G: Varia

I refer to the Table of Contents for rather detailed information about the contents of the individual parts and sections of these notes and want to just provide some remarks here for a first orientation.

Part A of the lecture notes is dedicated to explaining and exploring the consequences of Einstein's insights into the relation between gravity and space-time geometry, and to developing the machinery (of tensor calculus and Riemannian geometry) required to describe physics in a curved space-time, i.e. in a gravitational field.

From about section 4 onwards, Part A can be read in parallel with other parts of these notes which deal with various applications of General Relativity. In particular, at this point in the course I find it useful to develop in parallel (and suggest to read in parallel) the more formal material on tensor analysis in Part A, and Part D (dealing with solar system tests of general relativity) – cf. the more detailed suggestions at the end of section 3. Not only does this provide an interesting and physically relevant application and illustration of the machinery developed so far, it also serves to provide an appropriate balance between physics and formalism in the lectures.

The topics covered in Parts A and D, together with the first section 19 of Part C dealing with the Einstein field equations, probably form the core of most introductory courses on general relativity. This provides (or is meant to provide) the basis for other applications or investigations of general relativity, and other sections of Part C and Parts E-G provide a reasonably large variety of topics to choose from.

In Part B of the lecture notes I have collected a number of different more mathematical topics that develop the formalism of tensor calculus and differential geometry in one way or another. Stricly speaking, none of these topics are essential for understanding some of the more elementary aspects of general relativity to be treated later on (so Part B can also be regarded as a mathematical appendix to the notes). However, some of them are required at a later stage to understand, or even formulate, certain somewhat more advanced aspects of general relativity (and it is perhaps best to then go back to this section if and when needed), and others are included simply because they are fun or beautiful (or, usually, both).

0.3 Literature

Most of the material covered in these notes, in particular in the introductory parts, is completely standard and can be found in many places. While my way of explaining things is my own, and numerous gratuitous "Remarks" throughout the notes as well as the selection of more advanced topics reflect my own interests, I make no claim to major originality in these notes and have not attempted to reinvent the wheel.

In particular, in earlier versions of these notes the presentation of much of the introductory material followed quite closely the treatment in Weinberg's classic

- S. Weinberg, *Gravitation and Cosmology*

and readers familiar with this book may still recognise the similarities in some places. Even though my own way of thinking about general relativity is much more geometric (and this has definitely influenced later versions of and additions to these notes), I have found that the pragmatic approach adopted by Weinberg is ideally suited to introduce general relativity to students with little mathematical background.

As far as more recent and modern books are concerned, here is a short personal selection of my favourites:

1. At an introductory level, a book that I like and highly recommend is
   - J. Hartle, *Gravity. An Introduction to Einstein's General Relativity*
2. At an intermediate level (i.e. more or less at the level of these notes), my favourite modern book is

   - S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*

3. At a more advanced level, my favourites are

   - S. Hawking, G. Ellis, *The large scale structure of space-time*
   - E. Poisson, *A Relativist's Toolkit: the Mathematics of Black Hole Mechanics*
   - R. Wald, *General Relativity*

   and I will frequently refer to these books in the body of the notes for discussions of more advanced and/or more mathematical topics.

4. The history of the development of general relativity is an important and complex subject, crucial for a thorough appreciation of general relativity. My remarks on this subject are scarce and possibly even misleading at times and should not be taken as gospel. For an authoritative and informative account, I strongly recommend the scientific biography of Einstein

   - A. Pais, *Subtle is the Lord: the science and life of Albert Einstein*

0.4 References and Footnotes

As mentioned before, much of the material covered in these notes is quite standard, and can be found in many places, and I have not attempted to provide references or attributions for this.

Nevertheless, these lecture notes contain a large number of footnotes, with significantly higher density in the sections of the notes dealing with more advanced and, specifically, more recent developments. For the most part, these are meant as pointers to the literature for further reading and with more information.

However, I have also attempted to indicate explicitly in footnotes whenever I have knowingly used or adopted something specific from a specific source that should perhaps not be considered common knowledge. If you feel that somewhere in these notes I have written or used something that should not be considered common knowledge and that has not been properly attributed, please let me know.

For referencing I have adopted the following procedure:

- When referring to textbooks, I usually just refer to them in the form "Author, Title" (as above), without indicating publisher, year, ... If you actually need this information, it will be easy for you to find it.

- When referring to articles, if they are available from the preprint server at

  `https://arXiv.org/`

I usually just refer to the `arXiv` number, regardless of whether or not that article has been published elsewhere (this just reflects the by now standard practice that people are more likely to first go there rather than to the library to look for or at an article).

- References to pre-`arXiv` articles are given in the traditional complete "Author(s), Title, Journal, ..." form.

0.5 Exercises

Exercises are, of course, an indispensable part of any course, in particular a course on general relativity, since it is impossible to familiarise oneself with the formalism (of tensor calculus) without actually doing calculations. Nevertheless, these lecture notes contain no exercises, or at least none that are explicitly labelled as such.

This simply reflects my own style of teaching, where exercises are very much integrated into the course and mainly serve the purpose of getting students to look at what was done in the course and to perhaps fill in some details that I skipped in class. In particular, I am no fan of exercises that go significantly beyond what is covered in class or in the notes (if it is relevant, I should explain or include it, if it is not then we may as well not bother).

However, most (sub-)sections contain numerous "Remarks", and many of them contain supplementary and/or more advanced information and material, and these may be regarded as (annotated) exercises or used as a basis for exercises.

If that does not provide enough or satisfactory material, see also

- A. Lightman, W. Press, R. Price, S. Teukolsky, *Problem Book in Relativity and Gravitation*

for almost 500 fully solved problems in relativity.

A: Physics in a Gravitational Field and Tensor Calculus

1 Einstein Equivalence Principle: from Gravity to Geometry

1.1 Motivation: The Einstein Equivalence Principle

The highly successful Newtonian theory of gravity can be succinctly summarised by two sets of differential equations, one describing the dynamics (motion) of particles in a given gravitational field (described by a potential ϕ), and the other describing the dynamics of the gravitational field itself, namely how ϕ is to be determined from a given mass configuration. The former takes the standard Newtonian form

$$m\ddot{\vec{x}} = \vec{F}_g = -m\vec{\nabla}\phi \tag{1.1}$$

(but we will come back in some detail below to the question if/why the same mass parameter m appears on both sides of this equation, so as to incorporate the observation, going back to Galileo, that "all bodies fall at the same rate in a a gravitational field"). The latter is the Poisson equation

$$\Delta\phi = 4\pi G_N \mu = (4\pi G_N/c^2)\rho \ \ , \tag{1.2}$$

with G_N denoting, here and throughout, Newton's constant, i.e. the gravitational coupling constant, and where μ is the mass density, and $\rho = \mu c^2$ the associated rest mass energy density - I will set $c = 1$ in the following and use ρ.

Let us start with the field equation. It is immediately evident that this cannot be the final story. Not only is this equation not Lorentz invariant. Because of the absence of time-derivatives in (1.2), it actually describes an "action at a distance" and an instantaneous propagation of the gravitational field to every point in space (if you wiggle your mass distribution here now, this will immediately effect the gravitational potential arbitrarily far away). This is something that Einstein had just successfully exorcised from other aspects of physics, and clearly Newtonian gravity had to be revised as well.

It is then also immediately clear that what would have to replace Newton's theory is something rather more complicated. The reason for this is that, according to Special Relativity, mass is just another form of energy. Then, since gravity couples to masses, in a relativistically invariant theory gravity will also have to couple to energy. In particular, therefore, gravity would have to couple to gravitational energy, i.e. to itself. As a consequence, the new gravitational field equations will, unlike Newton's, have to be non-linear: the field of the sum of two masses cannot equal the sum of the gravitational fields of the two masses because it should also take into account the gravitational energy of the two-body system.

Now, having realised that Newton's theory cannot be the final word on the issue, how does one go about finding a better theory?

I will first very briefly discuss (and then dismiss) what at first sight may appear to be the most natural and naive approach to formulating a relativistic theory of gravity, namely the simple replacement of Newton's field equation (1.2) by its relativistically covariant version

$$\Delta\phi = 4\pi G_N \rho \quad \longrightarrow \quad \Box\phi = 4\pi G_N\rho \ \ , \tag{1.3}$$

where $\Box$ is the Lorentz invariant d'Alembert or wave operator. While this looks promising, something can't be quite right about this equation. We already know (from Special Relativity) that ρ is not a scalar but rather the 00-component of a tensor, the energy-momentum tensor T_{ab}, so if actually ρ appears on the right-hand side, ϕ cannot be a scalar, while if ϕ is a scalar something needs to be done to fix the right-hand side.

Turning first to the latter possibility, one option that suggests itself is to replace ρ by the trace $T = T^a_{\ a}$ of the energy-momentum tensor. This is by definition / construction a scalar, and it will agree with ρ in the non-relativistic limit (where rest mass dominates over other contributions). With the sign conventions that we will use, one has $T_{00} = \rho \Rightarrow T = -\rho + \ldots$, and thus a first attempt at fixing the above equation might look like

$$\Box\phi = -4\pi G_N T \ \ . \tag{1.4}$$

This is certainly an attractive equation, but it definitely has the drawback that it is too linear. Recall from the discussion above that the universality of gravity (coupling to all forms of matter) and the equivalence of mass and energy lead to the conclusion that gravity should couple to gravitational energy, invariably predicting non-linear (self-interacting) equations for the gravitational field. However, the left hand side could be such that it only reduces to $\Box$ or Δ of the Newtonian potential in the Newtonian limit of weak time-independent fields. Thus a second attempt at fixing the above equation might look like

$$\Box\Phi(\phi) = -4\pi G_N T \ \ , \tag{1.5}$$

where $\Phi(\phi) \approx \phi$ for weak fields.

Such scalar relativistic theories of gravity (or rather some minor variants thereof) were proposed and studied among others by Abraham, Mie, and Nordstrøm. As it stands, the field equation appears to be perfectly consistent (and it may be interesting to discuss if/how the Einstein equivalence principle, which will put us on our route towards metrics and space-time curvature is realised in such a theory).

However, regardless of this, this theory is incorrect simply because it is ruled out experimentally. The easiest way to see this (with hindsight) is to note that the energy-momentum tensor of Maxwell theory (7.47) is traceless (7.121), and thus the above equation would predict no coupling of gravity to the electro-magnetic field, in particular to light. Hence in such a theory there would be no deflection of light by the sun etc.[1]

[1] For more on the history and properties of scalar theories of gravity see the review by D. Giulini, *What is (not) wrong with scalar gravity?*, `arXiv:gr-qc/0611100`.

The other possibility to render (1.3) consistent is the, a priori perhaps much less compelling, option to think of ϕ and $\Delta\phi$ or $\Box\phi$ not as scalars but as (00)-components of some tensor, in which case one could try to salvage (1.3) by promoting it to a tensorial equation

$$\{\text{Some tensor generalising } \Delta\phi\}_{ab} \sim 4\pi G_N T_{ab} \quad . \tag{1.6}$$

This appears to require not just one potential, but actually 10 of them,

$$\phi \to \phi_{ab} = \phi_{ba} \tag{1.7}$$

and this seems to be a rather crazy thing to do at this stage (in particular, without any insight into the nature of these potentials). However, this is indeed the form of the field equations for gravity (the Einstein equations) we will ultimately be led to (see section 19.4), but Einstein arrived at this in a completely different, and much more insightful, way.

Let us now, very briefly and in a streamlined way, try to retrace (one aspect of) Einstein's thoughts, namely on the relation between inertial and gravitational mass, which, as we will see, will lead us rather quickly to the geometric picture of gravity sketched in the Introduction.

To that end we return to the Newtonian equation of motion (1.1). Recall that in this Newtonian theory, there are two *a priori* completely independent concepts of mass:

- inertial mass m_i (or acceleration mass), which accounts for the resistance of a body or particle against acceleration and appears universally on the left-hand-side of the Newtonian equation of motion

$$m_i \vec{a} = \vec{F} \tag{1.8}$$

  in conjunction with the acceleration $\vec{a}$;

- gravitational mass m_g which is the mass the gravitational field couples to, i.e. it is the gravitational charge of a particle,

$$\vec{F}_g = -m_g \vec{\nabla}\phi \quad . \tag{1.9}$$

Now it is an important empirical fact that the inertial mass of a body is equal to its gravitational mass. This realisation, at least with this clarity, is usually attributed to Newton, although it goes back to experiments and observations by Galileo usually paraphrased as "all bodies fall at the same rate in a gravitational field". (It is not true, though, that Galileo dropped objects from the leaning tower of Pisa to test this - he used an inclined plane, a water clock and a pendulum).

These experiments were later on improved, in various forms, by Huygens, Newton, Bessel and others and reached unprecedented accuracy with the work of Baron von

Eötvös (starting in 1889), who was able to show that inertial and gravitational mass of different materials (like wood and platinum) agree to one part in 10^9. In the 1950/60's, this was still further improved by R. Dicke to something like one part in 10^{11}. More recently, rumours of a 'fifth force', based on a reanalysis of Eötvös' data (but buried in the meantime) motivated experiments with even higher accuracy and no difference between m_i and m_g was found.

Newton's theory would in principle be perfectly consistent with $m_i \neq m_g$, just as the formally analogous equation for an electrically charged particle with charge q_e in an electrostatic field $\vec{E} = -\vec{\nabla}\phi$,

$$m_i \ddot{\vec{x}} = -q_e \vec{\nabla}\phi \quad , \tag{1.10}$$

is perfectly acceptable for any ratio q_e/m_i, and Einstein was very impressed with the observed equality of m_i and m_g. This should, he reasoned, not be a mere coincidence but is probably trying to tell us something rather deep about the nature of gravity.

To see what this could be, let us recall that there is a very common class of "forces" for which the equality between the inertial mass and the coupling constant is actually built in and automatic. These are the "pseudo-forces" or "fictitious forces" $\vec{P}$ (like centrifugal forces) which arise when one transforms the Newtonian equations of motion via a non-linear coordinate transformation to accelerated (or other non-Cartesian) coordinates,

$$x^i \to z^m = z^m(x^i) \tag{1.11}$$

(like spherical coordinates). These "forces" arise from the non-trivial transformation behaviour of the acceleration $\ddot{\vec{x}}$ under such non-linear coordinate transformations, and are therefore inevitably and automatically proportional to m_i,

$$m_i \ddot{\vec{x}} = \vec{F} \quad \Rightarrow \quad m_i \ddot{\vec{z}} = \vec{F} + \vec{P} \quad \text{with} \quad \vec{P} \sim m_i \quad . \tag{1.12}$$

To be explicit, note that if we perform such a time-independent coordinate transformation, the velocity and acceleration of a particle with trajectory $x^i(t)$ transform as

$$\begin{aligned} \dot{z}^m &= \frac{\partial z^m}{\partial x^i} \dot{x}^i \\ \ddot{z}^m &= \frac{\partial z^m}{\partial x^i} \ddot{x}^i + \frac{\partial^2 z^m}{\partial x^i \partial x^j} \dot{x}^i \dot{x}^j \end{aligned} \tag{1.13}$$

(cf. section 1.5 for an analogous calculation in relativistic mechanics in Minkowski space). The first line expresses the fact that velocities transform linearly (with the Jacobi matrix) under arbitrary coordinate transformations (and are thus the protyotype of what we will call vectors or tensors later on). In the coordinates z^m, the equations of motion thus take the form

$$m_i \ddot{z}^m = \frac{\partial z^m}{\partial x^i} F^i + P^m \equiv F^m + P^m \quad , \tag{1.14}$$

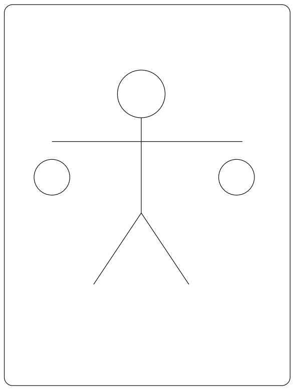


Figure 1: Experimenter and his two stones freely floating somewhere in outer space, i.e. in the absence of forces.

where it is the second term in the second line of (1.13) that gives rise to the (centrifugal etc.) pseudo-force

$$P^m = m_i \frac{\partial^2 z^m}{\partial x^i \partial x^j} \dot{x}^i \dot{x}^j \quad , \tag{1.15}$$

manifestly proportional to m_i. If one considers 3-dimensional coordinate transformations that depend explicitly on time, $z^m = z^m(x^i, t)$ (such as a transformation to a rotating reference system), then one finds additional contributions to the pseudo-force, like Coriolis forces etc. Conversely, any such pseudo-forces can be eliminated by transforming the equations of motion to a suitable (inertial) coordinate system or reference system.

With his unequalled talent for discovering profound truths in such simple observations, he concluded (calling this “der glücklichste Gedanke meines Lebens” (the happiest thought of my life)) that the equality of inertial and gravitational mass suggests a close relation between inertia and gravity itself, suggests, in fact, that locally effects of gravity and acceleration (or non-linear transformations of the reference system) are indistinguishable,

gravitational mass = inertial mass <u>because</u> (locally) GRAVITY = ACCELERATION

He substantiated this with some classical thought experiments, *Gedankenexperimente*, as he called them, which in the meantime have morphed into and have come to be known as the *elevator thought experiments*, which we will now discuss.

1. Consider somebody in a small sealed box (elevator) somewhere in outer space. In the absence of any forces, this person will float. Likewise, two stones he has just dropped (see Figure 1) will float with him.

2. Now assume (Figure 2) that somebody on the outside suddenly pulls the box up with a constant acceleration. Then of course, our friend will be pressed to the

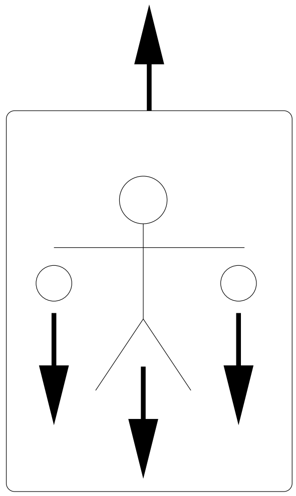


Figure 2: Constant acceleration upwards mimics the effect of a gravitational field: experimenter and stones drop to the bottom of the box.

bottom of the elevator with a constant force and he will also see his stones drop to the floor.

3. Now consider (Figure 3) this same box brought into a constant gravitational field. Then again, the experimenter will be pressed to the bottom of the elevator with a constant force and will see the stones drop to the floor. There is no experiment inside the elevator that permits him to decide if this is actually due to a gravitational field or due to the fact that somebody is pulling the elevator upwards.

   Thus our first lesson is that, indeed, locally the effects of acceleration and gravity are indistinguishable.

4. Now consider somebody cutting the cable of the elevator (Figure 4). Then the elevator will fall freely downwards but, as in Figure 1, our experimenter and his stones will float as in the absence of gravity.

   Thus lesson number two is that, locally the effect of gravity can be eliminated by going to a freely falling reference frame (or coordinate system). This should not come as a surprise, since this is of course built into the Newtonian theory: free fall in a constant gravitational field (in the vertial z-direction, say) is described by the equation

$$\ddot{z} = -g \ \ (+ \text{ other forces}) \ . \tag{1.16}$$

   In the accelerated (freely falling) coordinate system

$$Z(z,t) = z + gt^2/2 \tag{1.17}$$

   the same physics is described by the equation

$$\ddot{Z} = 0 \ \ (+ \text{ other forces}) \ , \tag{1.18}$$

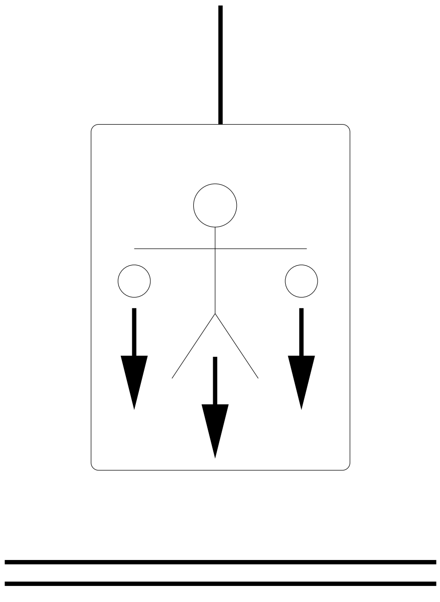


Figure 3: Effect of a constant gravitational field: indistinguishable for our experimenter from that of a constant acceleration in Figure 2.

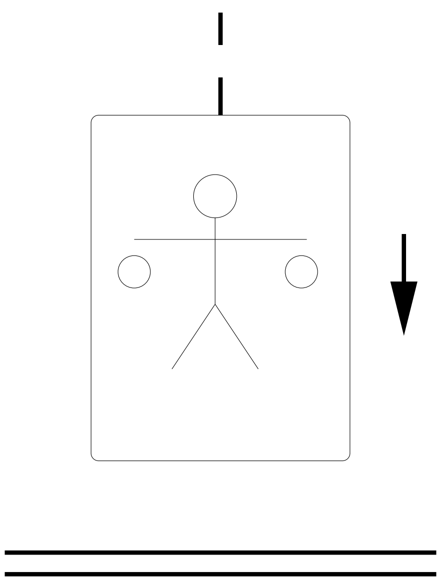


Figure 4: Free fall in a gravitational field has the same effect as no gravitational field (Figure 1): experimenter and stones float.

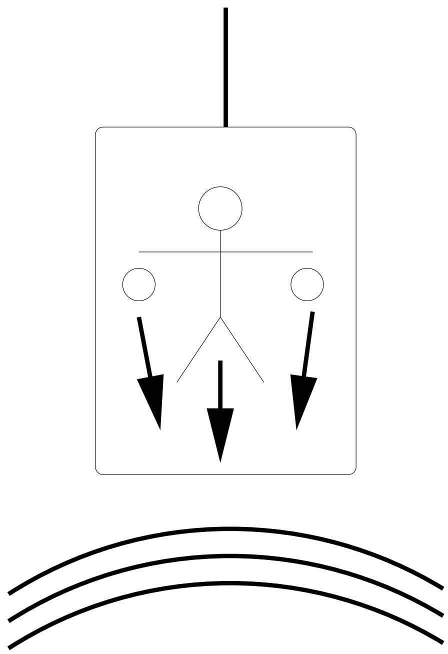


Figure 5: Experimenter and his stones in a non-uniform gravitational field: the stones will approach each other slightly as they fall to the bottom of the elevator.

and the effect of gravity has been eliminated completely by going to the freely falling reference system described by Z. The crucial point here is that in such a reference frame not only our observer will float freely, but because of the equality of inertial and gravitational mass he will also observe all other objects obeying the usual laws of motion in the absence of gravity.

5. In the above discussion, I have put the emphasis on constant accelerations and on 'locally'. To see the significance of this, consider our experimenter with his elevator in the gravitational field of the earth (Figure 5). This gravitational field is not constant but spherically symmetric, pointing towards the center of the earth. Therefore the stones will slightly approach each other as they fall towards the bottom of the elevator, in the direction of the center of the gravitational field.

6. Thus, if somebody cuts the cable now and the elevator is again in free fall (Figure 6), our experimenter will float again, so will the stones, but our experimenter will also notice that the stones move closer together for some reason. He will have to conclude that there is some force responsible for this.

   This is lesson number three: in a non-uniform gravitational field the effects of gravity cannot be eliminated by going to a freely falling coordinate system. This is only possible locally, on such scales on which the gravitational field is essentially constant.

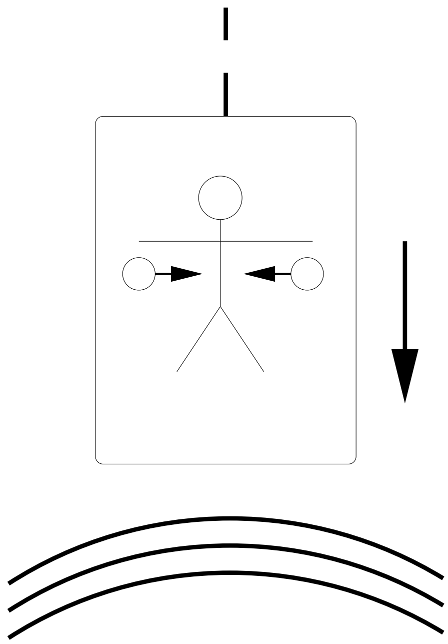


Figure 6: Experimentator and stones freely falling in a non-uniform gravitational field. The experimenter floats, so do the stones, but they move closer together, indicating the presence of some external force.

Einstein formalised the outcome of these thought experiments in what is now known as the *Einstein Equivalence Principle* which roughly states that physics in a freely falling frame in a gravitational field is the same as physics in an inertial frame in Minkowski space in the absence of gravitation. Two formulations are

> At every space-time point in an arbitrary gravitational field it is possible to choose a *locally inertial* (or *freely falling*) *coordinate system* such that, within a sufficiently small region of this point, the laws of nature take the same form as in unaccelerated Cartesian coordinate systems in the absence of gravitation.[2]

and

> Experiments in a sufficiently small freely falling laboratory, over a sufficiently short time, give results that are indistinguishable from those of the same experiments in an inertial frame in empty space.[3]

There are different versions of this principle depending on what precisely one means by 'the laws of nature'. If one just means the laws of Newtonian (or relativistic) mechanics,

[2] S. Weinberg, *Gravitation and Cosmology.*
[3] J. Hartle, *Gravity. An Introduction to Einstein's General Relativity.*

then this principle essentially reduces to the statement that inertial and gravitational mass are equal. Usually, however, this statement is taken to imply also Maxwell's theory, quantum mechanics etc.[4] What it pragmatically asserts in one of its stronger forms is that

> [...] there is no experiment that can distinguish a uniform acceleration from a uniform gravitational field. (J. Hartle, loc. cit.)

The power of the above principle, which we will regard as a heuristic guideline, rather than trying to (prematurely) give it a mathematically precise formulation, lies in the fact that we can combine it with our understanding of physics in accelerated reference systems to gain insight into the physics in a gravitational field. Two immediate consequences of this (which cannot be derived on the basis of Newtonian physics or Special Relativity alone) are

- light is deflected by a gravitational field just like material objects;
- clocks run slower in a gravitational field than in the absence of gravity.

To see the inevitability of the first assertion, imagine a lightray entering the rocket / elevator in Figure 1 horizontally through a window on the left hand side and exiting again at the same height through a window on the right. Now imagine, as in Figure 2, accelerating the elevator upwards. Then clearly the lightray that enters on the left will exit at a lower point of the elevator on the right because the elevator is accelerating upwards. By the equivalence principle one should observe exactly the same thing in a constant gravitational field (Figure 3). It follows that in a gravitational field the lightray is bent downwards, i.e. it experiences a downward acceleration with the (locally constant) gravitational acceleration g.

To understand the second assertion, one can e.g. simply appeal to the so-called "twin-paradox" of Special Relativity: the accelerated twin is younger than his unaccelerated inertial sibling. Hence accelerated clocks run slower than inertial clocks. Hence, by the equivalence principle, clocks in a gravitational field run slower than clocks in the absence of gravity.

Alternatively, one can imagine two observers at the top and bottom of the elevator, having identical clocks and sending light signals to each other at regular intervals as determined by their clocks. Once the elevator accelerates upwards, the observer at the bottom will receive the signals at a higher rate than he emits them (because he is accelerating towards the signals he receives), and he will interpret this as his clock

[4] For a discussion of different formulations of the equivalence principle and the logical relations among them, see E. di Casola, S. Liberati, S. Sonego, *Nonequivalence of equivalence principles*, `arXiv:1310.7426 [gr-qc]`.

running more slowly than that of the observer at the top. By the equivalence principle, the same conclusion now applies to two observers at different heights in a gravitational field. This can also be interpreted in terms of a gravitational redshift or blueshift (photons losing or gaining energy by climbing or falling in a gravitational field), and we will return to a more quantitative discussion of this effect in section 3.5.

1.2 Lorentz-Covariant Formulation of Special Relativity (Review)

What the equivalence principle tells us is that we can expect to learn something about the effects of gravitation by transforming the laws of nature (equations of motion) from an inertial Cartesian coordinate system to other (accelerated, curvilinear) coordinates. As a first step, we will, in section 1.3 below, discuss the above example of an observer undergoing constant acceleration in the context of special relativity.

As a preparation for this, and the remainder of the course, this section will provide a lightning review of the Lorentz-covariant formulation of special relativity, mainly to set the notation and conventions that will be used throughout, and only to the extent that it will be used in the following.

1. Minkowski space(-time)

   (a) The arena of special relativity is Minkowski space-time [henceforth *Minkowski space* for short, the union of space and time is implied by uttering the word "Minkowski"]. It is the space of events, labelled by 3 Cartesian spatial coordinates x^k and a time-coordinate t or, more usefully, by the coordinates

   $$(\xi^a) = (\xi^0 = ct, \xi^k = x^k) \quad , \tag{1.19}$$

   where c is the speed of light. Typically in these notes ξ^a will indicate such a (locally) inertial coordinate system, whereas generic coordinates will be called x^μ etc. We will almost always work in units in which $c = 1$.

   (b) Minkowski space is equipped with a prescription for measuring distances, encoded in a line-element which, in these coordinates, takes the form

   $$ds^2 = -(d\xi^0)^2 + \sum_k (d\xi^k)^2 \quad . \tag{1.20}$$

   (c) This can be written as

   $$ds^2 = -(d\xi^0)^2 + \sum_k (d\xi^k)^2 \equiv \eta_{ab}\, d\xi^a d\xi^b \quad . \tag{1.21}$$

   with metric $(\eta_{ab}) = \text{diag}(-1, +1, +1, +1)$ or, more explicitly,

   $$\eta_{ab} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} \tag{1.22}$$

(thus we are using the "mostly plus" convention).

2. Lorentz Transformations

   (a) Lorentz transformations are by definition those linear transformations

$$\xi^a \mapsto \bar{\xi}^a = L^a_b \xi^b \tag{1.23}$$

   that leave the Minkowski line-element invariant,

$$d\bar{s}^2 \equiv \eta_{ab} d\bar{\xi}^a d\bar{\xi}^b = \eta_{ab} d\xi^a d\xi^b = ds^2 \quad \Leftrightarrow \quad \eta_{ab} L^a_{\ c} L^b_{\ d} = \eta_{cd} \ . \tag{1.24}$$

   In matrix notation this can also be written as

$$\bar{\xi} = L\xi \quad , \quad L^t \eta L = \eta \tag{1.25}$$

   where L^t is the transpose of L. This is the defining condition for Lorentz transformations, and the Lorentz signature analogue of the condition $R^t \mathbb{1} R = \mathbb{1}$ for an orthogonal transformation (rotation or reflection) in Euclidean space, with metric $\eta_{ab} \to \mathbb{1}_{ik} = \delta_{ik}$.
   Alternative notation:

$$\xi^{\bar{a}} = L^{\bar{a}}_b \xi^b \quad \text{or} \quad \xi^{\bar{a}} = L^{\bar{a}}_b \xi^b \tag{1.26}$$

   Strictly speaking $\bar{\xi}^a$ and $\xi^{\bar{a}}$ may be considered to refer to two different quantities, to the coordinates of the new point $\bar{\xi}$ in the old coordinate system versus the coordinates of the old point ξ in the new coordinate system. However, for many elementary purposes this difference between what is known as the active (moving points) versus the passive (relabelling points) point of view is not crucial, and one should not be hung-up on notation: coordinates (and in particular indices referring to them) are fundamentally just bookkeeping devices so use whatever is convenient for current bookkeeping or other purposes.

   (b) Infinitesimal Lorentz rotations, i.e. Lorentz transformations with L of the form $L = \mathbb{1} + \omega$, ω infinitesimal, are characterised by

$$(\mathbb{1} + \omega)^t \eta (\mathbb{1} + \omega) = \eta \quad \Rightarrow \quad (\eta\omega) + (\eta\omega)^t = 0 \ . \tag{1.27}$$

   Thus the matrix $\eta\omega$ is anti-symmetric. In components, an infinitesimal Lorentz transformation therefore has the form

$$\delta\xi^a = \omega^a_{\ b} \xi^b \quad \text{with} \quad \omega_{ab} \equiv \eta_{ac} \omega^c_{\ b} = -\omega_{ba} \ . \tag{1.28}$$

   (c) Poincaré transformations are those affine transformations that leave the Minkowski line-element invariant. They are composed of Lorentz transformations and arbitrary constant translations and thus have the form

$$\xi^a \mapsto \bar{\xi}^a = L^a_b \xi^b + \zeta^a \ , \tag{1.29}$$

infinitesimally

$$\delta\xi^a = \omega^a_{\ b}\xi^b + \epsilon^a \quad . \tag{1.30}$$

Any two inertial systems in the sense of the equivalence principle of special relativity are related by a Poincaré transformation.

3. Distance & Causal Structure

   (a) The Minkowski metric defines the Lorentz (and Poincaré) invariant distance

   $$(\Delta\xi)^2 = \eta_{ab}(\xi^a_P - \xi^a_Q)(\xi^b_P - \xi^b_Q) \tag{1.31}$$

   betwen two events P and Q with coordinates ξ^a_P and ξ^a_Q respectively.

   (b) Depending on the sign of $(\Delta\xi)^2$, the two events P, Q are called, spacelike, lightlike (null) or timelike separated,

   $$(\Delta\xi)^2 = \begin{cases} > 0 & \text{spacelike} \quad \text{separated} \\ = 0 & \text{lightlike} \quad \text{separated} \\ < 0 & \text{timelike} \quad \text{separated} \end{cases} \tag{1.32}$$

   (c) The set of events that are lightlike separated from P define the lightcone at P. It consists of two components (joined at P), the future and the past lightcone, distinguished by the sign of $\xi^0_Q - \xi^0_P$ (positive for Q on the future lightcone, $\xi^0_Q > \xi^0_P$, negative for Q on the past lightcone).

4. Curves and Tangent Vectors

   (a) A parametrised curve is given by a map $\lambda \mapsto \xi^a(\lambda)$. The tangent vector to the curve at the point $\xi(\lambda_0)$ has components

   $$\xi'^a(\lambda_0) = \frac{d}{d\lambda}\xi^a(\lambda)|_{\lambda=\lambda_0} \quad . \tag{1.33}$$

   It is called spacelike, lightlike (null) or timelike, depending on the sign of $\eta_{ab}\xi'^a\xi'^b$,

   $$\eta_{ab}\xi'^a\xi'^b \quad \begin{cases} > 0 & \text{spacelike} \\ = 0 & \text{lightlike} \\ < 0 & \text{timelike} \end{cases} \tag{1.34}$$

   This sign (and hence this classification) depends only on the image of the curve, not its parametrisation.

   (b) A curve whose tangent vector is everywhere timelike is called a timelike curve (and likewise for lightlike and spacelike curves). A curve whose tangent vector is everywhere timelike or null (i.e. non-spacelike) is called a causal curve. Worldlines of massive particles are timelike curves, those of massless particles (light) are null curves.

(c) A natural Lorentz-invariant parametrisation of timelike curves is provided by the Lorentz-invariant proper time τ along the curves,

$$\xi^a = \xi^a(\tau) \quad , \tag{1.35}$$

with

$$\begin{aligned} c d\tau &= \sqrt{-ds^2} = \sqrt{-\eta_{ab} d\xi^a d\xi^b} = \sqrt{-\eta_{ab}\xi'^a\xi'^b} d\lambda \\ \Rightarrow \quad \eta_{ab} &\frac{d\xi^a(\tau)}{d\tau}\frac{d\xi^b(\tau)}{d\tau} = -c^2 \quad . \end{aligned} \tag{1.36}$$

Likewise spacelike curves are naturally parametrised by proper distance ds. The derivative with respect to proper time will be denoted by an overdot,

$$\dot{\xi}^a(\tau) = \frac{d}{d\tau}\xi^a(\tau) \quad . \tag{1.37}$$

Because τ is Lorentz-invariant, $\bar{\tau} = \tau$, tangent vectors $\dot{\xi}^a$ of τ-parametrised curves transform linearly under Lorentz transformations,

$$\dot{\bar{\xi}}^a(\tau) = \frac{d}{d\tau}\bar{\xi}^a(\tau) = \frac{\partial \bar{\xi}^a}{\partial \xi^b}\frac{d}{d\tau}\xi^b(\tau) = L^a_{\ b}\dot{\xi}^b(\tau) \quad . \tag{1.38}$$

These are the prototypes of what are called Lorentz vectors or, more generally, Lorentz tensors.

5. Lorentz Vectors

(a) Lorentz vectors (or 4-vectors) are objects with components v^a which transform under Lorentz transformations with the matrix $L^a_{\ b}$ (to be thought of as the Jacobian of the transformation relating $\bar{\xi}^a$ and ξ^a),

$$\bar{v}^a = L^a_{\ b} v^b \quad . \tag{1.39}$$

(b) η_{ab} can be regarded as defining an indefinite scalar product on the space of Lorentz vectors, and the Minkowski norm $\eta_{ab}v^a v^b$ and the Minkowski scalar product $\eta_{ab}v^a w^b$ of Lorentz vectors are invariant under Lorentz transformations,

$$\eta_{ab}\bar{v}^a\bar{v}^b = \eta_{ab}v^a v^b \quad , \quad \eta_{ab}\bar{v}^a\bar{w}^b = \eta_{ab}v^a w^b \quad . \tag{1.40}$$

A vector is called, spacelike, lightlike (null) or timelike depending on the sign of its Minkowski norm.

(c) One can identify Minkowski space with its "tangent space", i.e. with the vector space $\mathbb{V} = \mathbb{R}^{1,3}$ of 4-vectors equipped with the quadratic form or scalar product η_{ab} with signature (1,3).

6. Other Lorentz Tensors

(a) Lorentz scalars are objects that are invariant under Lorentz transformations. Examples are scalar products and norms of Lorentz vectors.

(b) Lorentz covectors are objects u_a that transform under Lorentz transformations with the dual (or contragredient = transpose inverse) representation

$$\Lambda = (L^t)^{-1} = \eta L \eta^{-1} \quad , \tag{1.41}$$

i.e.

$$\bar{u}_a = \Lambda_a^{\ b} u_b \quad , \quad \Lambda_a^{\ b} = \eta_{ac} L^c_{\ d} \eta^{db} \quad , \tag{1.42}$$

where η^{ab} denotes the components of the inverse metric η^{-1}. In terms of Λ, the condition that L is a Lorentz transformation (i.e. preserves η_{ab}) can evidently be written as

$$L^t \eta L = \eta \quad \Leftrightarrow \quad \Lambda \eta \Lambda^t = \eta \quad \Leftrightarrow \quad \Lambda_a^{\ c} \Lambda_b^{\ d} \eta_{cd} = \eta_{ab} \quad . \tag{1.43}$$

Covectors can be regarded as elements of the dual $\mathbb{V}^*$ of the space $\mathbb{V}$ of 4-vectors, with u_a defining the Lorentz-invariant linear mapping

$$u: \ v \in \mathbb{V} \mapsto u(v) = u_a v^a \in \mathbb{R} \quad . \tag{1.44}$$

Examples are $u_a = \eta_{ab} v^b \equiv v_a$ with v^a a Lorentz vector, the scalar product η_{ab} defining the isomorphism $\mathbb{V}^* \cong \mathbb{V}$.

(c) Lorentz (p,q)-tensors are objects that transform under Lorentz transformations like a product of p vectors and q covectors,

$$T^{a_1 \ldots a_p}_{c_1 \ldots c_q} \rightarrow \bar{T}^{a_1 \ldots a_p}_{c_1 \ldots c_q} = L^{a_1}_{\ b_1} \ldots L^{a_p}_{\ b_p} \Lambda_{c_1}^{\ d_1} \ldots \Lambda_{c_q}^{\ d_q} \, T^{b_1 \ldots b_p}_{d_1 \ldots d_q} \quad . \tag{1.45}$$

In particular, direct products of vectors and covectors like $V^a W^b U_c$ are tensors. A special case is η_{ab}, which is a Lorentz-invariant $(0,2)$-tensor by definition,

$$\bar{\eta}_{ab} = \Lambda_a^{\ c} \Lambda_b^{\ d} \eta_{cd} = \eta_{ab} \quad . \tag{1.46}$$

Linear combinations of (p,q)-tensors are again (p,q)-tensors. Arbitrary products and contractions of Lorentz tensors are again Lorentz tensors (and the tensor type can be read off from the number and position of the "free" indices).

7. Tensor Fields

(a) Lorentz tensor fields are assignments of Lorentz tensors to each point of Minkowski space,

$$T: \xi \mapsto T^{a_1 \ldots a_p}_{c_1 \ldots c_q}(\xi) \quad . \tag{1.47}$$

(b) Given a vector field $V^a(\xi)$, $\eta_{ab}V^a(\xi)V^b(\xi)$ is an example of a scalar field, and given a scalar field $f(\xi)$, its partial derivatives give a covector field

$$U_a(\xi) = \partial_{\xi^a} f(\xi) \equiv \partial_a f(\xi) \tag{1.48}$$

(providing the justification for abbreviating $\partial_{\xi^a} = \partial_a$). More generally, the partial derivatives of the components of a (p, q)-tensor,

$$T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) \quad \rightarrow \quad \partial_a T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) \tag{1.49}$$

are the components of a $(p, q+1)$-tensor, and the wave operator

$$\Box = \eta^{ab}\partial_a\partial_b \tag{1.50}$$

is a Lorentz-invariant differential operator mapping (p, q) tensor fields to (p, q) tensor fields.

(c) Tensorial equations of the form

$$T^{a_1 \dots a_p}_{c_1 \dots c_q}(\xi) = 0 \tag{1.51}$$

are Lorentz invariant in the sense that they are satisfied in one inertial system iff they are satisfied in all inertial systems. (Here and throughout these notes "iff" is the usual mathematicians' shorthand for "if and only if".)

8. Worldlines of Massive Particles

(a) In the covariant formulation, the timelike worldline of a massive particle is parametrised by proper time, $\xi^a = \xi^a(\tau)$. The velocity (tangent) vector

$$u^a \equiv \dot{\xi}^a(\tau) \tag{1.52}$$

is a Lorentz vector, normalised as

$$u^a u_a \equiv \eta_{ab} u^a u^b = -c^2 \quad . \tag{1.53}$$

(b) The Lorentz-covariant acceleration is the 4-vector

$$a^c = \frac{d}{d\tau} u^c = \frac{d^2}{d\tau^2} \xi^c \quad , \tag{1.54}$$

and the equation of motion of a massive free particle is

$$a^c = \frac{d^2}{d\tau^2} \xi^c(\tau) = 0 \quad . \tag{1.55}$$

We will study this equation further (in any arbitrary coordinate system) in section 1.5. For observers with non-zero acceleration it follows from (1.53) by differentiation that a^c is orthogonal to u^b,

$$a^c u_c \equiv \eta_{cb} a^c u^b = 0 \quad , \tag{1.56}$$

and therefore spacelike,

$$\eta_{cb} a^c a^b \equiv \mathsf{a}^2 > 0 \quad . \tag{1.57}$$

Observers with constant acceleration will be the subject of section 1.3.

(c) The action for a free massive particle with worldline $\xi^a(\tau)$ is essentially the total proper time along the path,

$$S[\xi] = -mc^2 \int d\tau = -mc \int \sqrt{-\eta_{ab} d\xi^a d\xi^b} \quad , \tag{1.58}$$

worldlines of free massive particles extremising (maximising) the proper time. In terms of an arbitrary parametrisation $\xi^a = \xi^a(\lambda)$ of the path, this action can be written as

$$S[\xi] = \int d\lambda \, L_\lambda \quad , \quad L_\lambda = -mc \left(-\eta_{ab} \frac{d\xi^a}{d\lambda} \frac{d\xi^b}{d\lambda} \right)^{1/2} \quad . \tag{1.59}$$

A special choice is $\lambda = t$, for which

$$L_t = -mc^2 \sqrt{1 - \vec{v}^2/c^2} \qquad \vec{v} = d\vec{\xi}/dt = d\vec{x}/dt \quad . \tag{1.60}$$

9. Energy-Momentum 4-Vector

(a) The covariant momenta p_a are defined by

$$p_a = \frac{\partial L_\lambda}{\partial (d\xi^a/d\lambda)} = m\eta_{ab} u^b \quad \Rightarrow \quad p^a = mu^a = m(d\xi^a/d\tau) \tag{1.61}$$

(independently of the choice of λ).

(b) Its components are

$$p^0 = E/c \quad , \quad p^k = p^{(c)k} \tag{1.62}$$

where $p^{(c)k}$ are the canonical momenta associated to the Lagrangian L_t,

$$p^{(c)}_k = \frac{\partial L_t}{\partial v^k} = m\gamma(v) v_k \quad , \tag{1.63}$$

with $\gamma(v) = (1 - \vec{v}^2/c^2)^{-1/2}$, and $E = H$ is the corresponding energy or Hamiltonian

$$H = p^{(c)}_k v^k - L_t = m\gamma(v) c^2 \quad . \tag{1.64}$$

(c) The p^a are the components of a Lorentz vector, the energy-momentum 4-vector. It satisfies the mass-shell relation

$$\eta_{ab} p^a p^b = -m^2 c^2 \quad \Leftrightarrow \quad E^2 = m^2 c^4 + \vec{p}^2 c^2 \quad . \tag{1.65}$$

1.3 Accelerated Observers in Minkowski Space and the Rindler Metric

We return to the issue discussed in the context of the Einstein equivalence principle in section 1.1, namely physics as experienced by an observer undergoing constant acceleration (as a precursor to studying this observer in a genuine gravitational field), now specifically within the framework of special relativity.

Specialising (1.56) to an observer accelerating in the ξ^1-direction (so that in the momentary restframe of this observer one has $u^a = (1,0,0,0), a^a = (0,\mathsf{a},0,0)$), we will say that the observer undergoes constant acceleration if a is time-independent. To determine the worldline of such an observer, we note that the general solution to (1.53) with $u^2 = u^3 = 0$,

$$\eta_{ab}u^a u^b = -(u^0)^2 + (u^1)^2 = -1 \quad , \tag{1.66}$$

is

$$u^0 = \cosh F(\tau) \quad , \quad u^1 = \sinh F(\tau) \tag{1.67}$$

for some function $F(\tau)$. Thus the acceleration is

$$a^a = \dot{F}(\tau)(\sinh F(\tau), \cosh F(\tau), 0, 0) \quad , \tag{1.68}$$

with norm

$$\eta_{ab}a^a a^b = \mathsf{a}^2 = \dot{F}^2 \quad , \tag{1.69}$$

and an observer with constant acceleration is characterised by $F(\tau) = \mathsf{a}\tau$,

$$u^a(\tau) = (\cosh \mathsf{a}\tau, \sinh \mathsf{a}\tau, 0, 0) \quad . \tag{1.70}$$

This can now be integrated, and in particular

$$\xi^a(\tau) = (\mathsf{a}^{-1}\sinh \mathsf{a}\tau, \mathsf{a}^{-1}\cosh \mathsf{a}\tau, 0, 0) \tag{1.71}$$

is the worldline of an observer with constant acceleration a and initial condition $\xi^a(\tau = 0) = (0, \mathsf{a}^{-1}, 0, 0)$. The worldlines of this observer is the hyperbola

$$\eta_{ab}\xi^a\xi^b = -(\xi^0)^2 + (\xi^1)^2 = \mathsf{a}^{-2} \tag{1.72}$$

in the quadrant $\xi^1 > |\xi^0|$ of Minkowski space-time.

We can now ask the question what the Minkowski metric or line-element looks like in the restframe of such an observer. Note that one cannot expect this to be again the constant Minkowski metric η_{ab}: the transformation to an accelerated reference system, while certainly allowed in special relativity, is not a Lorentz transformation, while η_{ab} is, by definition, invariant under Lorentz-transformations.

We are thus looking for coordinates that are adapted to these accelerated observers in the same way that the inertial coordinates are adapted to static observers (ξ^0 is proper time, and the spatial components ξ^i remain constant). In other words, we seek a coordinate transformation $(\xi^0, \xi^1) \to (\eta, \rho)$ such that the worldlines of these accelerated observers are characterised by $\rho =$ constant (this is what we mean by restframe, the observer stays at a fixed value of ρ) and ideally such that then η is proportional to the proper time of the observer.

Comparison with (1.71) suggests the coordinate transformation

$$\xi^0(\eta, \rho) = \rho \sinh \eta \qquad \xi^1(\eta, \rho) = \rho \cosh \eta \quad . \tag{1.73}$$

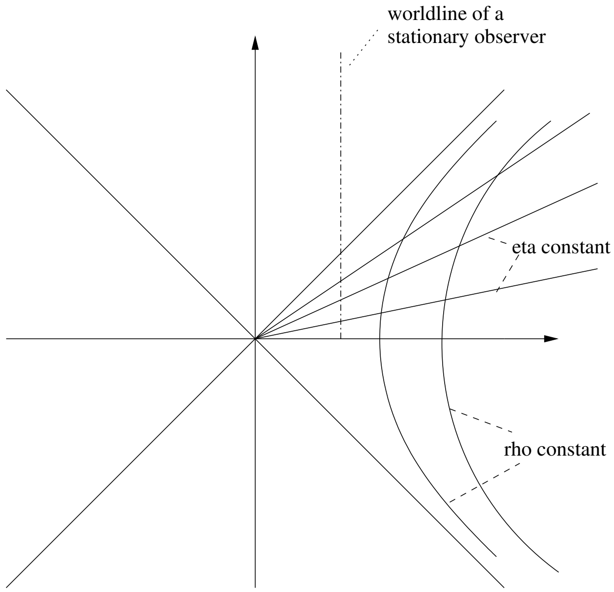


Figure 7: Rindler metric: Rindler coordinates (η, ρ) cover the first quadrant $\xi^1 > |\xi^0|$. Indicated are lines of constant ρ (hyperbolas, worldlines of constantly accelerating observers) and lines of constant η (straight lines through the origin). The quadrant is bounded by the lightlike lines $\xi^0 = \pm\xi^1 \Leftrightarrow \eta = \pm\infty$. An inertial observer reaches and crosses the line $\eta = \infty$ in finite proper time $\tau = \xi^0$.

It is now easy to see that in terms of these new coordinates the 2-dimensional Minkowski metric $ds^2 = -(d\xi^0)^2 + (d\xi^1)^2$ (we are now suppressing, here and in the remainder of this subsection, the transverse spectator dimensions 2 and 3) takes the form

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 \quad . \tag{1.74}$$

This is the so-called *Rindler metric*.

Let us try to gain a better understanding of these Rindler coordinates (illustrated in Figure 7 - see also Figure 25 in section 28.4 for a *Penrose Diagram* illustration).

- The Rindler coordinates ρ and η are obvisouly in some sense hyperbolic (Lorentzian) analogues of polar coordinates ($x = r\cos\phi, y = r\sin\phi, ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2$). In particular, since

$$(\xi^1)^2 - (\xi^0)^2 = \rho^2 \qquad , \qquad \frac{\xi^0}{\xi^1} = \tanh\eta \quad , \tag{1.75}$$

  by construction the lines of constant ρ, $\rho = \rho_0$, are hyperbolas, $(\xi^1)^2 - (\xi^0)^2 = \rho_0^2$, while the lines of constant $\eta = \eta_0$ are straight lines through the origin, $\xi^0 = (\tanh\eta_0)\xi^1$.

- The null lines $\xi^0 = \pm\xi^1$ correspond to $\eta = \pm\infty$. Thus the Rindler coordinates cover the first quadrant $\xi^1 > |\xi^0|$ of Minkowski space and can be used as coordinates there.

- The metric in these new coordinates is time-independent, where time means η, and time-independent means that the coefficients of the metric or line-element in (1.74) do not depend on η. This is due to the fact that the generator ∂_η of η-"time evolution" is actually the generator of a Lorentz boost in the (ξ^0, ξ^1)-plane in Minkowski space,

$$\partial_\eta = (\partial_\eta \xi^0)\partial_{\xi^0} + (\partial_\eta \xi^1)\partial_{\xi^1} = \xi^1 \partial_{\xi^0} + \xi^0 \partial_{\xi^1} \quad . \tag{1.76}$$

  Since a Lorentz boost leaves the Minkowski metric invariant, the latter has to be invariant under translations in η, i.e. it has to be η-independent, as is indeed the case.

- Along the worldline of an observer with constant ρ one has $d\tau = \rho_0 d\eta$, so that his proper time parametrised path is

$$\xi^0(\tau) = \rho_0 \sinh \tau/\rho_0 \qquad \xi^1(\tau) = \rho_0 \cosh \tau/\rho_0 \quad , \tag{1.77}$$

  and his 4-velocity is given by

$$u^0 = \frac{d}{d\tau}\xi^0(\tau) = \cosh \tau/\rho_0 \qquad u^1 = \frac{d}{d\tau}\xi^1(\tau) = \sinh \tau/\rho_0 \quad . \tag{1.78}$$

  These satisfy $-(u^0)^2 + (u^1)^2 = -1$ (as they should), and comparison with (1.70,1.71) shows that the observer's (constant) acceleration is $\mathsf{a} = 1/\rho_0$.

Even though (1.74) is just the metric of Minkowski space-time, written in accelerated coordinates, this metric exhibits a number of interesting features that are prototypical of more general metrics that one encounters in general relativity:

1. First of all, we notice that the coefficients of the line element (metric) in (1.74) are no longer constant (space-time independent). Since in the case of constant acceleration we are just describing a "fake" gravitational field, this dependence on the coordinates is such that it can be completely and globally eliminated by passing to appropriate new coordinates (namely inertial Minkowski coordinates). Since, by the equivalence principle, locally an observer cannot distinguish between a fake and a "true" gravitational field, this now suggests that a "true" gravitational field can be described in terms of a space-time coordinate dependent line-element

$$ds^2 = g_{\alpha\beta}(x)dx^\alpha dx^\beta \quad , \tag{1.79}$$

   where the coordinate dependence on the x^α is now such that it cannot be eliminated globally by a suitable choice of coordinates.

2. We observe that (1.74) appears to be ill-defined at $\rho = 0$. However, in this case we already know that this is a mere *coordinate singularity* at $\rho = 0$ (akin to the

coordinate singularity at the origin of standard polar coordinates in the Cartesian plane). More generally, whenever a metric written in some coordinate system appears to exhibit some singular behaviour, one needs to investigate whether this is just a coordinate singularity or a true singularity of the gravitational field itself.

3. The above coordinates do not just fail at $\rho = 0$, they actually fail to cover large parts of Minkowski space. Thus the next lesson is that, given a metric in some coordinate system, one has to investigate if the space-time described in this way needs to be extended beyond the range of the original coordinates. One way to analyse this question (which we will make extensive use of in sections 26 and 27 when trying to understand and come to terms with black holes) is to study lightrays or the worldlines of freely falling (inertial) observers.

   In the present case, an example of an inertial observer is a static observer in Minkowski space, i.e. an observer at a fixed value of ξ^1, say, with $\xi^0 = \tau$ his proper time. In Rindler coordinates this is described by the condition that $\xi^1 = \rho\cosh\eta$ is a constant, so this is most certainly not a straight line in an (η, ρ)-diagram.

   Such an observer will of course "discover" that $\eta = +\infty$ is not the end of the world (indeed, he crosses this line at finite proper time $\tau = \xi^1$) and that Minkowski space continues (at the very least) into the quadrant $\xi^0 > |\xi^1|$ (see Figure 7 for an illustration of this).

4. Related to this is the behaviour of lightcones when expressed in terms of the coordinates (η, ρ) or when drawn in the (η, ρ)-plane (do this!). These lightcones satisfy $ds^2 = 0$, i.e.

$$\rho^2 d\eta^2 = d\rho^2 \quad \Rightarrow \quad d\eta = \pm\rho^{-1} d\rho \ . \tag{1.80}$$

   describing outgoing (ρ grows with η) respectively ingoing (ρ decreases with increasing η) lightrays. These lightcones have the familiar Minkowskian shape at $\rho = 1$, but the lightcones open up for $\rho > 1$ and become more and more narrow for $\rho \to 0$, once again exactly as we will find for the Schwarzschild black hole metric (see Figure 16 in section 26).

5. It follows from (1.76) that the Minkowski norm of ∂_η is

$$|\partial_\eta|^2 = (\xi^0)^2 - (\xi^1)^2 \ . \tag{1.81}$$

   Thus this generator of Rindler time-translations really is timelike in the region $\xi^1 > |\xi^0|$ covered by the Rindler coordinates, but it actually becomes lightlike on the lightlike boundary $\xi^1 = |\xi^0|$ of that region. As we will discuss in section 27.10, such a *Killing horizon* also happens to be one of the characteristic properties of a black hole.

6. Finally we note that there is a large region of Minkowski space that is "invisible" to the constantly accelerated observers. While a static observer will eventually receive information from any event anywhere in space-time (his past lightcone will eventually cover all of Minkowski space ...), the past lightcone of one of the Rindler accelerated observers (whose worldlines asymptote to the lightcone direction $\xi^0 = \xi^1$) will asymptotically only cover one half of Minkowski space, namely the region $\xi^0 < \xi^1$. Thus any event above the line $\xi^0 = \xi^1$ will forever be invisible to this class of observers. Such an observer-dependent horizon has some similarities with the *event horizon* characterising a black hole (see section 27.5 for a first encounter with such an object, and section 32 for a detailed discussion).

For more on Rindler coordinates, see sections 3.4 and 7.8.

1.4 General Coordinate Transformations in Minkowski Space I: Metric

In order to move away from constant accelerations (as models of observers in constant gravitational fields only), we now consider the effect of arbitrary (general) coordinate transformations on the laws of special relativity and the geometry of Minkowski space. This may look like a somewhat exaggerated move at this point (should we perhaps not just look at coordinate transformations to coordinates that somehow correspond to adapted coordinates for some arbitrary accelerated observer?), but

- it is actually easier to just do this than to understand what is meant precisely by this parenthetical remark and how to implement it;
- there are many useful things that one can learn from doing this;
- and we will see later (when discussing the relation between the *Einstein Equivalence Principle* and the *Principle of General Covariance* in section 4.1), that the relation between the description of physics in an arbitrary gravitational field and the behaviour of this description under arbitrary coordinate transformations is much closer and more far-reaching than we perhaps have the right to expect at the moment.

Let us see what things look like when written in some other (non-inertial, accelerating) coordinate system. It is extremely useful for bookkeeping purposes and for avoiding algebraic errors to use different kinds of indices for different coordinate systems. Thus we will usually call the new coordinates $x^\mu(\xi^b)$ or $x^\alpha(\xi^a)$, and not, say, $x^a(\xi^b)$ (although there would be nothing wrong with that).

We start with the definition of proper time, as described in inertial coordinates by the Minkowski line element,

$$d\tau^2 = -\eta_{ab}d\xi^a d\xi^b \tag{1.82}$$

First of all, this proper time should not depend on which coordinates we use to describe the motion of the particle, but only on the world line of the particle itself. After all, the particle could not care less what coordinates we experimenters or observers use to describe the particle's proper time.

[By the way: this is the best way to resolve the so-called "twin-paradox", which should really be referred to (and presented) as the "twin-non-paradox" or simply the "twin-fact" - everything else is deliberate and unhelpful obfuscation! It does not matter which reference system you use - the accelerating twin in the rocket will always be younger than her brother when they meet again.]

Thus all we need to know is how the same proper time τ is expressed in terms of the new coordinates, which simply follows from

$$d\tau^2 = -\eta_{ab} d\xi^a d\xi^b = -\eta_{ab} \frac{\partial \xi^a}{\partial x^\mu} \frac{\partial \xi^b}{\partial x^\nu} dx^\mu dx^\nu \quad . \tag{1.83}$$

Here

$$J^a_\mu(x) = \frac{\partial \xi^a}{\partial x^\mu} \tag{1.84}$$

is the Jacobi matrix associated to the coordinate transformation $\xi^a = \xi^a(x^\mu)$, and we will make the assumption that (locally) this matrix is non-degenerate, thus has an inverse $J^\mu_a(x)$ or $J^\mu_a(\xi)$ which is the Jacobi matrix associated to the inverse coordinate transformation $x^\mu = x^\mu(\xi^a)$,

$$J^a_\mu J^\mu_b = \delta^a_b \qquad J^\mu_a J^a_\nu = \delta^\mu_\nu \quad . \tag{1.85}$$

We see that in the new coordinates, proper time and distance are no longer measured by the Minkowski metric in its standard form (the constant matrix η_{ab}), but by

$$d\tau^2 = -\eta_{\mu\nu}(x) dx^\mu dx^\nu \quad , \tag{1.86}$$

where the *metric tensor* (or *metric* for short) $\eta_{\mu\nu}(x)$, the Minkowski metric in the coordinates x^μ, is

$$\eta_{\mu\nu}(x) = \eta_{ab} \frac{\partial \xi^a}{\partial x^\mu} \frac{\partial \xi^b}{\partial x^\nu} \quad . \tag{1.87}$$

<u>REMARKS:</u>

1. The fact that the Minkowski metric written in the coordinates x^μ in general depends on x should not come as a surprise - after all, this also happens when one writes the Euclidean metric in spherical coordinates, $ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2$ etc.

2. Even though the components of the metric are not those of the Minkowski metric in inertial coordinates, this metric or line element (with its associated presciption for time- and space-measurements) still describes exactly the same Minkowskian

geometry as the standard Minkowski metric η_{ab}: just as passing from Cartesian to spherical coordinates in $\mathbb{R}^3$ does not change the Euclidean geometry of the space, using some other than inertial coordinates does not change the Minkowskian geometry of Minkowski space.

3. It is easy to check, using (1.85), that the inverse metric, which we will denote by $\eta^{\mu\nu}$,

$$\eta^{\mu\nu}(x)\eta_{\nu\lambda}(x) = \delta^\mu_{\ \lambda} \ , \tag{1.88}$$

is given by

$$\eta^{\mu\nu}(x) = \eta^{ab}\frac{\partial x^\mu}{\partial \xi^a}\frac{\partial x^\nu}{\partial \xi^b} \ . \tag{1.89}$$

In terms of Jacobi matrices we have

$$\eta_{\mu\nu} = J^a_\mu J^b_\nu \eta_{ab} \quad , \quad \eta^{\mu\nu} = J^\mu_a J^\nu_b \eta^{ab} \ . \tag{1.90}$$

The rationale for using a different positioning of the indices for the metric and its inverse is that, as is apparent form the above equations, they transform in a *different* way under coordinate transformations: one with the Jacobi matrices, the other with their inverses. Thus geometrically they are different kinds of objects, and (as in the tensor algebra for Lorentz tensors of special relativity) the positioning of the indices is used to indicate and keep track of this fact in a simple and intuitive manner.

We will have much more to say about the metric below and, indeed, throughout this course.

1.5 General Coordinate Transformations in Minkowski Space II: Free Particle

We now turn to the equation of motion of a free particle, given in inertial coordinates by

$$\frac{d^2}{d\tau^2}\xi^a(\tau) = 0 \ . \tag{1.91}$$

The usual rules for a change of variables give

$$\frac{d}{d\tau}\xi^a = \frac{\partial \xi^a}{\partial x^\mu}\frac{dx^\mu}{d\tau} \ , \tag{1.92}$$

where $J^a_\mu = \frac{\partial \xi^a}{\partial x^\mu}$ is the invertible Jacobi matrix. This shows that, as usual, velocities transform in a particularly simple (linear, vectorial) way under arbitrary coordinate transformations, namely just with the Jacobi matrix,

$$\dot{\xi}^a = J^a_\mu \dot{x}^\mu \ . \tag{1.93}$$

Differentiating once more, one finds

$$\begin{aligned}
\frac{d^2}{d\tau^2}\xi^a &= \frac{\partial \xi^a}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \frac{\partial^2\xi^a}{\partial x^\nu \partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} \\
&= \frac{\partial \xi^a}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \delta^a_{\ b}\frac{\partial^2\xi^b}{\partial x^\nu \partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} \\
&= \frac{\partial \xi^a}{\partial x^\mu}\left[\frac{d^2x^\mu}{d\tau^2} + \frac{\partial x^\mu}{\partial \xi^b}\frac{\partial^2\xi^b}{\partial x^\nu \partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \quad . \qquad (1.94)
\end{aligned}$$

Thus, since the matrix appearing outside the square bracket is invertible, in terms of the coordinates x^μ the equation of motion, or the equation for a straight (and, in the case at hand, timelike) line in Minkowski space, becomes

$$\frac{d^2x^\mu}{d\tau^2} + \frac{\partial x^\mu}{\partial \xi^a}\frac{\partial^2\xi^a}{\partial x^\nu \partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \quad . \qquad (1.95)$$

We will write this as

$$\frac{d^2x^\mu}{d\tau^2} + \gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \quad , \qquad (1.96)$$

or just

$$\ddot{x}^\mu + \gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \quad , \qquad (1.97)$$

where

$$\gamma^\mu_{\ \nu\lambda} = \frac{\partial x^\mu}{\partial \xi^a}\frac{\partial^2\xi^a}{\partial x^\nu \partial x^\lambda} \quad . \qquad (1.98)$$

Remarks:

1. Of course the statement (1.93) regarding the linear transformation rule of velocities under coordinate transformations remains true for transformations between arbitary coordinate systems in Minkowski space, $\{x^\mu\}$ and $\{y^\alpha\}$, say, i.e. one has

$$\dot{y}^\alpha = \frac{\partial y^\alpha}{\partial x^\mu}\dot{x}^\mu \equiv J^\alpha_\mu \dot{x}^\mu \quad . \qquad (1.99)$$

In general this simply follows from the chain rule. It can also be deduced (in a somewhat unnecessarily long way included here only for later reference purposes) from what we have already done, namely by simply repeating the calculation leading to (1.93), but now for the coordinates y^α,

$$\dot{\xi}^a = J^a_\alpha \dot{y}^\alpha \quad . \qquad (1.100)$$

This now implies

$$\dot{y}^\alpha = J^\alpha_a \dot{\xi}^a = J^\alpha_a J^a_\mu \dot{x}^\mu = J^\alpha_\mu \dot{x}^\mu \qquad (1.101)$$

where

$$J^\alpha_\mu = \frac{\partial y^\alpha}{\partial \xi^a}\frac{\partial \xi^a}{\partial x^\mu} = \frac{\partial y^\alpha}{\partial x^\mu} \qquad (1.102)$$

and in the last step the chain rule (or multiplicativity of the Jacobi matrix under consecutive coordinate transformations) was used.

2. While (1.97) looks a bit complicated and unattractive, it is simply the general variant of a calculation that you have probably done numerous times before in various specific contexts. In particular, the second term in this equation, is just the general expression for a pseudo-force or fictitious gravitational force (like a centrifugal force or the Coriolis force) that arises whenever one describes inertial motion in non-inertial coordinates.

3. More compactly, this pseudo-force term can be written as

$$\gamma^{\mu}_{\nu\lambda} = J^{\mu}_{a}\partial_{\nu}J^{a}_{\lambda} = J^{\mu}_{a}\partial_{\lambda}J^{a}_{\nu} \equiv J^{\mu}_{a}J^{a}_{\nu\lambda} \quad . \tag{1.103}$$

It is absent precisely for linear coordinate transformations $\xi^a(x^\mu) = M^a_\mu x^\mu$,

$$\gamma^{\mu}_{\nu\lambda} = 0 \quad (\forall\, ^{\mu},_{\nu},_{\lambda}) \quad \Leftrightarrow \quad \xi^a = M^a_\mu x^\mu \tag{1.104}$$

for some constant matrix M^a_μ. In particular, this means that the equation of motion for a free particle is invariant under Lorentz transformations, as it should be.

4. By the same reasoning, the quantity $\gamma^{\mu}_{\nu\lambda}$ is independent of the choice of reference inertial coordinate system ξ^a. I.e. if $\zeta^a = L^a_b\xi^b$ for some Lorentz (or more general linear) transformation matrix L^a_b, then

$$\frac{\partial x^\mu}{\partial \xi^a}\frac{\partial^2 \xi^a}{\partial x^\nu \partial x^\lambda} = \frac{\partial x^\mu}{\partial \zeta^a}\frac{\partial^2 \zeta^a}{\partial x^\nu \partial x^\lambda} \tag{1.105}$$

5. In the same way that the equation of motion for a free particle in inertial coordinates follows from the extremisation of the proper time (written in inertial coordinates),

$$\delta \int d\tau = \delta \int \sqrt{-\eta_{ab} d\xi^a d\xi^b} = 0 \quad \Rightarrow \quad \ddot{\xi}^a = 0 \ , \tag{1.106}$$

the equation of motion for a free particle in noninertial coordinates follows from the extremisation of the proper time (written in these noninertial coordinates),

$$\delta \int d\tau = \delta \int \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu} = 0 \quad \Rightarrow \quad \ddot{x}^\mu + \gamma^{\mu}_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \ . \tag{1.107}$$

It is a straightforward exercise to establish this, and simply reflects the well-known covariance of the Euler-Lagrange equations under coordinate transformations. The proof will not be given here also because we will straightaway establish a more general statement (for a space equipped with an arbitrary metric) in section 2.3 below.

6. It is immediately apparent from this action principle, that the (pseudo-)force terms in the equation of motion will arise from the fact that in general the components of the metric $\eta_{\mu\nu}(x)$ are not constant in these cooordinates. In particular, therefore, it must be possible to write the pseudo-force terms $\gamma^{\mu}_{\nu\lambda}$ in terms of the first partial derivatives of the metric $\eta_{\mu\nu}$. We will come back to this in section 1.6 below.

Even though the resulting equations look a bit uninviting at the moment, that is just what you get when you do write things in arbitrary coordinates. Moreover, there are at least two very useful things that we can extract or anticipate from this, namely

1. candidates for the appropriate generalisation of the Newtonian gravitational potential
2. the prototypical *general covariance* of physical equations

in *any* theory of gravity satisfying the Einstein equivalence principle. Let us now discuss these features in turn (relegating some uninspiring calculational details to the end of this subsection):

1. the Metric as a Candidate for the Gravitational Potential

   Recall that in non-inertial coordinates the metric takes the form

   $$\eta_{\mu\nu} = J^a_\mu J^b_\nu \eta_{ab} \tag{1.108}$$

   and the equation of motion of a free particle is (1.97) with the pseudo-force term (1.103)

   $$\gamma^\mu_{\ \nu\lambda} = J^\mu_a \partial_\nu J^a_\lambda \ \ . \tag{1.109}$$

   It turns out that this term can be expressed in terms of the partial derivatives of the metric (as anticipated from the action principle above). Indeed, given any symmetric and invertible object $g_{\mu\nu}(x)$ (a "metric"), with inverse $g^{\mu\nu}$, define

   $$\begin{aligned} \Gamma^\mu_{\ \nu\lambda} &= g^{\mu\rho}\Gamma_{\rho\nu\lambda} \\ \Gamma_{\rho\nu\lambda} &= \tfrac{1}{2}(g_{\rho\nu},_\lambda + g_{\rho\lambda},_\nu - g_{\nu\lambda},_\rho) \end{aligned} \tag{1.110}$$

   (this definition is such that it remains valid for an arbitary metric). Then one finds

   $$g_{\mu\nu} = \eta_{\mu\nu} \equiv J^a_\mu J^b_\nu \eta_{ab} \quad \Rightarrow \quad \Gamma^\mu_{\ \nu\lambda} = \gamma^\mu_{\ \nu\lambda} \ \ . \tag{1.111}$$

   It is an elementary but nevertheless useful exercise to check this (see below - but do try this yourself as well).

   This shows that the components $\eta_{\mu\nu}$ of the metric appear to play the role of "potentials" for the gravitational pseudo-force. In particular, since in principle all components of the metric can contribute to $\gamma_{\rho\nu\lambda}$, we learn the interesting fact that in order to achieve this a single scalar potential is completely insufficient (and one could have discovered the possibility or necessity of a multitude of potentials simply by the study of pseudo-forces in non-gravitational Newtonian mechanics).

It is enormously pleasing to note that the "number" of potentials that we seem to have discovered, namely 10 (for the symmetric (4×4)-matrix $g_{\mu\nu}(x)$), agrees with the number of potentials anticipated in our discussion of section 1.1 when we contemplated a tensorial generalisations (1.6) of the Poisson equation (with source T_{ab}).

If the metric indeed plays the role of the gravitational potential, as suggested by these considerations, then it will play the role of the fundamental dynamical variable of gravity. Since the metric encodes what one usually refers to as the geometry of a space(-time), as we will discuss in much more detail below, namely the information required to determine distances, areas, volumes etc., this means that we are being led to the conclusion that any theory of gravity based on the equivalence principle is a theory of dynamical geometry. Wow ...

2. the General Covariance of the Equation of Motion

   The equation of motion (1.97) has one other fundamental redeeming and attractive feature which will also make it the prototype of the kind of equations that we will be looking for in general. This feature is its *covariance* under general coordinate transformations, i.e. its *general covariance*, which means that the equation takes the same form in any coordinate system. Equivalently, but somewhat more to the point, as discussed in more detail in section 4.1 below, this is the statement that the equation is satisified in one coordinate system if and only if it is satisfied in all coordinate systems.

   This covariance is in some sense tautologically true since the coordinate system $\{x^\mu\}$ that we have chosen is indeed arbitrary. However, it is instructive to see how this comes about by explicitly transforming (1.97) from one coordinate system x^μ to another, say y^α.

   If one does this (cf. below for a proof), one finds that the equations of motion (1.97) in the coordinates x^μ and y^α are related by

$$\frac{d^2 y^\alpha}{d\tau^2} + \gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = \frac{\partial y^\alpha}{\partial x^\mu}\left[\frac{d^2 x^\mu}{d\tau^2} + \gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \tag{1.112}$$

   Thus the geodesic equation transforms in the simplest possible non-trivial way under coordinate transformations $x \to y$, namely with the Jacobi matrix

$$J^\alpha_\mu = \frac{\partial y^\alpha}{\partial x^\mu} \quad . \tag{1.113}$$

   We will see later that this transformation behaviour characterises/defines tensors, in this particular case a vector (or contravariant tensor of rank 1).

   In particular, since this matrix is assumed to be invertible, we reach the conclusion that the left hand side of (1.112) is zero if and only if the term in square brackets on the right hand side is zero,

$$\frac{d^2y^\alpha}{d\tau^2} + \gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = 0 \quad \Leftrightarrow \quad \frac{d^2x^\mu}{d\tau^2} + \gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \tag{1.114}$$

This is what is meant by the statement that the equation takes the same form in any coordinate system, and is therefore satisfied in one coordinate system if and only if it is satisfied in all coordinate systems. We see that in this case this is achieved by having the equation transform in a particularly simple way under coordinate transformations, namely as a tensor. We will discuss this in more generality in section 4.1.

---

1. Proof of (1.111):

   - From
     $$\eta_{\mu\nu} = \eta_{ab}J^a_\mu J^b_\nu \tag{1.115}$$
     one deduces
     $$\eta_{\mu\nu,\lambda} = \eta_{ab}(J^a_{\mu\lambda}J^b_\nu + J^a_\mu J^b_{\nu\lambda}) \tag{1.116}$$
     where
     $$J^a_{\mu\lambda} = \partial_\lambda J^a_\mu = \frac{\partial^2\xi^a}{\partial x^\mu\partial x^\lambda} = J^a_{\lambda\mu} \quad . \tag{1.117}$$
   - Therefore, form the definition (1.110) of the Γ-symbols, one has (for $g_{\mu\nu} = \eta_{\mu\nu}$)
     $$\begin{aligned}\Gamma_{\mu\nu\lambda} &= \tfrac{1}{2}(\eta_{\mu\nu,\lambda} + \eta_{\mu\lambda,\nu} - \eta_{\nu\lambda,\mu}) \\ &= \tfrac{1}{2}\eta_{ab}(J^a_{\mu\lambda}J^b_\nu + J^a_\mu J^b_{\nu\lambda} + J^a_{\mu\nu}J^b_\lambda + J^a_\mu J^b_{\lambda\nu} - J^a_{\nu\mu}J^b_\lambda - J^a_\nu J^b_{\lambda\mu}) \\ &= \eta_{ab}J^a_\mu J^b_{\nu\lambda} \quad ,\end{aligned} \tag{1.118}$$
     where the cancellations in passing to the last line arise from the symmetries $\eta_{ab} = \eta_{ba}$, $J^b_{\lambda\mu} = J^b_{\mu\lambda}$ etc.
   - Thus, finally (and writing out everything in detail for once),
     $$\begin{aligned}\Gamma^\mu{}_{\nu\lambda} &= \eta^{\mu\rho}\Gamma_{\rho\nu\lambda} = \eta^{cd}J^\mu_c J^\rho_d \eta_{ab}J^a_\rho J^b_{\nu\lambda} = \eta^{cd}J^\mu_c \delta^a_d \eta_{ab}J^b_{\nu\lambda} \\ &= \eta^{ca}J^\mu_c \eta_{ab}J^b_{\nu\lambda} = \delta^c_b J^\mu_c J^b_{\nu\lambda} = J^\mu_b J^b_{\nu\lambda} \quad ,\end{aligned} \tag{1.119}$$
     as was to be shown.

2. Proof of (1.112):

   - We proceed as in the proof of (1.99). Thus consider transforming the free particle equation of motion in inertial coordinates (1.55) not to the coordinate systsem x^μ, as we did before, but to another coordinate system $\{y^\alpha\}$.

Following the same steps as above, one arrives at the y-version of (1.94), namely

$$\frac{d^2}{d\tau^2}\xi^a = \frac{\partial \xi^a}{\partial y^\alpha}\left[\frac{d^2 y^\alpha}{d\tau^2} + \frac{\partial y^\alpha}{\partial \xi^b}\frac{\partial^2 \xi^b}{\partial y^\beta \partial y^\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau}\right] \quad . \tag{1.120}$$

- Equating this result to (1.94) and using the chain rule for partial derivatives

$$\frac{\partial y^\alpha}{\partial x^\mu} = \frac{\partial y^\alpha}{\partial \xi^a}\frac{\partial \xi^a}{\partial x^\mu} \quad , \tag{1.121}$$

one finds

$$\frac{d^2 y^\alpha}{d\tau^2} + \gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = \frac{\partial y^\alpha}{\partial x^\mu}\left[\frac{d^2 x^\mu}{d\tau^2} + \gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \tag{1.122}$$

as claimed.

2 Metrics, Geometry and Geodesics

Above we saw that the motion of free particles in Minkowski space in curvilinear coordinates is described in terms of a modified metric, $g_{\mu\nu}$, and a force term $\gamma^{\mu}_{\nu\lambda}$ representing the "pseudo-force" on the particle. Thus the Einstein Equivalence Principle suggests that an appropriate description of true gravitational fields is in terms of a metric tensor $g_{\mu\nu}(x)$ (and its associated Γ-symbols) which can only locally be related to the Minkowski metric via a suitable coordinate transformation (to locally inertial coordinates). We adopt this as our working hypothesis.

2.1 Metrics and Geometry I: Definition and Examples

Thus our starting point will now be a space-time equipped with some metric $g_{\mu\nu}(x)$, which (by analogy with the Euclidean and Minkowski metrics) we will assume to be symmetric and non-degenerate, i.e.

$$g_{\mu\nu}(x) = g_{\nu\mu}(x) \qquad \mathsf{det}(g_{\mu\nu}(x)) \neq 0 \ . \tag{2.1}$$

The metric encodes the information how to measure (spatial and temporal) distances, as well as areas, volumes etc., via the associated line element

$$g_{\mu\nu}(x) \quad \Rightarrow \quad ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \ . \tag{2.2}$$

As an example, the most general 2-dimensional line element (on a space with local coordinates $(x^i) = (x^1, x^2)$) has the form

$$ds^2 = g_{ij} dx^i dx^j = g_{11}(dx^1)^2 + (g_{22})(dx^2)^2 + 2g_{12} dx^1 dx^2 \tag{2.3}$$

(which is non-degenerate if $g_{11}g_{22} - (g_{12})^2 \neq 0$).

A metric determines a geometry (in the literal sense of a prescription for measuring distances etc.), but different metrics may well determine the same geometry, namely those metrics which are just related by coordinate transformations. In particular, distances should not depend on which coordinate system is used. Hence, changing coordinates from the $\{x^\mu\}$ to new coordinates $\{y^\alpha(x^\mu)\}$ and demanding that

$$g_{\mu\nu}(x) dx^\mu dx^\nu = g_{\alpha\beta}(y) dy^\alpha dy^\beta \ , \tag{2.4}$$

one finds that under a coordinate transformation a metric transforms as

$$g_{\alpha\beta}(y) = g_{\mu\nu}(x) \frac{\partial x^\mu}{\partial y^\alpha} \frac{\partial x^\nu}{\partial y^\beta} \equiv J^\mu_\alpha J^\nu_\beta \, g_{\mu\nu} \ . \tag{2.5}$$

Objects which transform in such a nice and simple way under coordinate transformations are known as *tensors* - the metric is an example of what is known as (and we will get to know as) a covariant symmetric rank two tensor. We will study tensors in much more detail and generality later, starting in section 4.

Remarks:

1. Here I have denoted the components of the metric in the new coordinates y^α simply by $g_{\alpha\beta}$. Occasionally it is more convenient to use a more elaborate notation, such as
$$x^\mu \to x'^\alpha = y^\alpha \quad \Rightarrow \quad g_{\mu\nu} \to g'_{\alpha\beta} = J^\mu_\alpha J^\nu_\beta g_{\mu\nu} \quad , \tag{2.6}$$
which allows one to distinguish notationally specific components of the metric in 2 different coordinate systems, such as g'_{11} (the (11)-component of the metric in the y-coordinates) from g_{11} (the (11)-component of the metric in the x-coordinates). As mentioned before, indices and other decorations are primarily bookkeeping devices; therefore I will usually not be overly-pedantic about these things in the following and will use whatever notation is more convenient in the case at hand.

2. As a consequence of the non-degeneracy condition, pointwise $g_{\mu\nu}(x)$ possesses an inverse, whose components we will denote by $g^{\mu\nu}(x)$, i.e.
$$g^{\mu\nu}(x) g_{\nu\lambda}(x) = \delta^\mu_\lambda \quad , \quad g_{\mu\nu}(x) g^{\nu\lambda}(x) = \delta^\lambda_\mu \quad . \tag{2.7}$$
Clearly, the inverse metric then transforms inversely, i.e. with the inverse Jacobi matrices J^α_μ, and this is now nicely compatible with the convention to denote the inverse metric by upper indices,
$$g^{\alpha\beta} = J^\alpha_\mu J^\beta_\nu g^{\mu\nu} \quad . \tag{2.8}$$
This is also the rationale for writing the inverse metric with "upper" indices: the positioning of indices is used to indicate how an object transforms under coordinate transformations (and we will formalise this in the discussion of section 4 on tensor algebra).

3. A space-time equipped with a metric tensor $g_{\mu\nu}(x)$ is called a metric space-time or (pseudo-)Riemannian space-time. Here "Riemannian" usually refers to a space equipped with a positive-definite metric (all eigenvalues positive), while pseudo-Riemannian (or Lorentzian) refers to a space-time with a metric with one negative and 3 (or 27, or whatever) positive eigenvalues.

4. One point to note about the tensorial transformation behaviour is that pointwise it is a similarity transformation in the sense of linear algebra, in matrix notation
$$g \mapsto J^t g J \quad . \tag{2.9}$$
In particular, therefore, if in one coordinate system the space-time metric tensor has one negative and three positive eigenvalues (as in a locally inertial coordinate

system), then the same will be true in any other coordinate system (even though the eigenvalues themselves will in general be different) - this statement should be familiar from linear algebra (e.g. as Sylvester's law, but it also goes under various other names).

Here are some examples of Riemannian metrics that you may already be familiar with.

<u>EXAMPLES:</u>

1. The Euclidean metrics or line-elements on $\mathbb{R}^2$ or $\mathbb{R}^3$, but written in polar or spherical coordinates,

$$\begin{aligned} ds^2(\mathbb{R}^2) &= dx^2 + dy^2 = dr^2 + r^2 d\phi^2 \\ ds^2(\mathbb{R}^3) &= dx^2 + dy^2 + dz^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \end{aligned} \tag{2.10}$$

   E.g. for the latter case one has

$$(x, y, z) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta) \ , \tag{2.11}$$

   and plugging this into the Euclidean line-element $dx^2 + dy^2 + dz^2$, one finds the above result.

   Denoting the Cartesian coordinates by x^α and the spherical coordinates by y^α, with $(y^1 = r, y^2 = \theta, y^3 = \phi)$, the non-vanishing components of the metric in the two coordinate systems are thus (using the prime notation (2.6))

$$g_{11} = g_{22} = g_{33} = 1 \quad , \quad g'_{11} = 1 \ , \ g'_{22} = r^2 \ , \ g'_{33} = r^2 \sin^2\theta \ . \tag{2.12}$$

   Alternatively, it is often more informative (and very common) to use the coordinates themselves, rather than indices, as the labels of the components of the metric tensor. In this case one can dispense with the prime notation and simply write the components of the metric in spherical coordinates as

$$g_{rr} = 1 \ , \ g_{\theta\theta} = r^2 \ , \ g_{\phi\phi} = r^2 \sin^2\theta \ . \tag{2.13}$$

2. Restricting the first example above to constant radius $r = R$, this gives us the line-element on the circle S^1_R of radius R,

$$ds^2(S^1_R) = R^2 d\phi^2 \ . \tag{2.14}$$

   Restricting the second to the 2-sphere S^2_R of radius R,

$$x^2 + y^2 + z^2 = r^2 = R^2 \quad \text{or} \quad r = R \ , \tag{2.15}$$

   one finds the line-element

$$ds^2(S^2_R) = R^2(d\theta^2 + \sin^2\theta d\phi^2) \equiv R^2 d\Omega^2 \ . \tag{2.16}$$

Here

$$d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2 \tag{2.17}$$

is usually called the solid angle, and we can now interpret it as the line element on the unit 2-sphere. We will use the notation / abbrevation $d\Omega^2$ for this line element throughout the notes.

This example provides a nice illustration of the fact that by drawing the coordinate grid / infinitesimal parallelograms determined by the metric tensor, one can get a feeling for the geometry and can in particular convince oneself that in general a metric space or space-time need not or cannot be flat, i.e. is not the flat Euclidean space of Euclidean geometry.

Indeed, the coordinate grid of the metric $d\theta^2 + \sin^2\theta d\phi^2$ cannot be drawn in flat space because the infinitesimal parallelograms described by ds^2 degenerate to triangles not just at $\theta = 0$ (as would also be the case for the flat metric $ds^2 = dr^2 + r^2 d\phi^2$ in polar coordinates at $r = 0$), but also at $\theta = \pi$. This coordinate grid can, on the other hand, of course be drawn on the 2-sphere.

3. This line-element on the unit 2-sphere generalises to the line-element on a unit 3-sphere,

$$ds^2(S^3) = d\psi^2 + \sin^2\psi(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{2.18}$$

This can be obtained by simply generalising the construction of spherical coordinates from $\mathbb{R}^3$ to $\mathbb{R}^4$, and (if required) this can be continued iteratively to yet higher-dimensional spheres.

Alternatively, by thinking of the 3-sphere as the locus

$$x^2 + y^2 + z^2 + w^2 = 1 \tag{2.19}$$

in $\mathbb{R}^4$, and "solving" this equation by first setting

$$x^2 + y^2 = \sin^2\alpha \quad , \quad z^2 + w^2 = \cos^2\alpha \quad , \tag{2.20}$$

and then refining this to

$$x = \sin\alpha\cos\beta \quad , \quad y = \sin\alpha\sin\beta \quad , \quad z = \cos\alpha\cos\gamma \quad , \quad w = \cos\alpha\sin\gamma \quad , \tag{2.21}$$

one finds that the standard Euclidean line-element

$$ds^2 = dx^2 + dy^2 + dz^2 + dw^2 \tag{2.22}$$

induces the line-element

$$ds^2(S^3) = d\alpha^2 + \sin^2\alpha d\beta^2 + \cos^2\alpha d\gamma^2 \tag{2.23}$$

on the sphere. This is the same metric on S^3 as above (2.18), namely the one induced from the Euclidean metric on $\mathbb{R}^4$, but written in different coordinates.

In particular, both are invariant under 4-dimensional rotations, i.e. under $SO(4)$-transformations.

However, we can obtain genuinely different metrics on the 3-sphere e.g. by starting with different metrics on $\mathbb{R}^4$. One of the simplest possibilities is to replace (2.22) by

$$d\tilde{s}^2 = a^2(dx^2 + dy^2) + b^2(dz^2 + dw^2) \quad , \tag{2.24}$$

with a, b real non-zero parameters. Then the induced metric on the 3-sphere $x^2 + y^2 + z^2 + w^2 = 1$ is

$$d\tilde{s}^2(S^3) = (a^2 \cos^2 \alpha + b^2 \sin^2 \alpha) d\alpha^2 + a^2 \sin^2 \alpha d\beta^2 + b^2 \cos^2 \alpha d\gamma^2 \quad . \tag{2.25}$$

For $a^2 \neq b^2$, this metric is not invariant under full 4-dimensional rotations, but only under rotations in the (x, y) and (z, w) planes, i.e. under $SO(2) \times SO(2)$ transformations. Thus this equips the 3-sphere with a genuinely different geometry (and is an example of what is sometimes referred to as a "squashed 3-sphere geometry").

4. If instead of the unit 2-sphere one considers the "unit" hyperboloid H^2, defined by

$$x^2 + y^2 + z^2 = +1 \quad \longrightarrow \quad x^2 + y^2 - z^2 = -1 \quad , \tag{2.26}$$

then this is naturally thought of as being embedded not in $\mathbb{R}^3$ but in $\mathbb{R}^{1,2}$, i.e. into the 3-dimensional vector space with line-element

$$ds^2 = dx^2 + dy^2 - dz^2 \quad . \tag{2.27}$$

The hyperbolic analogues (r, σ, ϕ) of the spherical coordinates, defined by

$$(x, y, z) = (r \sinh \sigma \cos \phi, r \sinh \sigma \sin \phi, r \cosh \sigma) \quad , \tag{2.28}$$

are naturally adapted to this situation, because

$$x^2 + y^2 - z^2 = -r^2 \tag{2.29}$$

so that the unit hyperboloid is evidently just the surface $r = 1$. In these coordinates, the metric (2.27) takes the form

$$ds^2 = -dr^2 + r^2(d\sigma^2 + \sinh^2 \sigma \, d\phi^2) \quad , \tag{2.30}$$

and therefore the induced metric on the unit hyperboloid $r = 1$ is

$$ds^2(H^2) = d\sigma^2 + \sinh^2 \sigma \, d\phi^2 \quad . \tag{2.31}$$

2.2 Metrics and Geometry II: Lorentzian (Pseudo-Riemannian) Metrics

We now turn to Lorentzian (pseudo-Riemannian) metrics and geometries. These will of course occupy and accompany us throughout these notes, so this section is meant to just provide a first brief encounter with these objects.

For a metric with Lorentzian signature, and with coordinates $x^\alpha = (x^0 = t, x^k)$, say, the metric has components $g_{00}, g_{0k} = g_{k0}$ and $g_{ik} = g_{ki}$, and the corresponding line element has the form

$$ds^2 = g_{00}dt^2 + 2g_{0k}dt\,dx^k + g_{ik}dx^i dx^k \tag{2.32}$$

Without any further conditions on the coefficients (except those ensuring non-degeneracy), this could a priori be a metric of any signature, and the signature of the metric may not always be readily apparent even when the coefficients of the metric are given explicitly.

Before looking at this in somewhat more detail, here are some simple examples, where the Lorentzian signature of the metrics is reasonably manifest:

<u>Examples:</u>

1. Of course any of the Riemannian metrics of the previous section can be promoted to space-time metrics by simply adding a $(-dt^2)$ (i.e. by taking the direct product with the time-axis). Thus the Minkowski metric in spatial spherical coordinates has the form

$$ds^2(\mathbb{R}^{1,3}) = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{2.33}$$

2. A generalisation of this is provided by the so-called *ultrastatic* metrics, i.e. metrics that are just a product of the standard metric $-dt^2$ along the time-direction and a spatial metric $\tilde{g}_{ij}(x)$

$$ds^2 = -dt^2 + \tilde{g}_{ij}(x)dx^i dx^j \tag{2.34}$$

(i.e. the components depend only on the spatial coordinates x^i, not on t).

3. Somewhat more generally, the spatial components of the metric can depend non-trivially on time. For example, a space-time metric describing a spatially spherical universe with a time-dependent radius (expansion of the universe!) might be described by the line element

$$ds^2 = -dt^2 + a(t)^2\left(d\psi^2 + \sin^2\psi(d\theta^2 + \sin^2\theta d\phi^2)\right) \quad , \tag{2.35}$$

and more generally one can consider the corresponding generalisation of (2.34), namely metrics of the form

$$ds^2 = -dt^2 + a(t)^2\tilde{g}_{ij}(x)dx^i dx^j \quad . \tag{2.36}$$

This describes a space-time with spatial metric $\tilde{g}_{ij}(x)dx^i dx^j$ and a time-dependent overall scale factor $a(t)$; in particular, such a space-time metric can describe an expanding universe in cosmology. We will discuss such metrics in detail later on in the context of cosmology, sections 33-38.

4. The (time-time)-component of the metric can of course in general depend non-trivially on the spatial coordinates. We already encountered this in the example of the Rindler metric (1.74), which has the form

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 \ . \tag{2.37}$$

5. A particularly prominent example is the *Schwarzschild Metric*

$$ds^2 = -(1 - 2m/r)dt^2 + (1 - 2m/r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{2.38}$$

   It is of fundamental importance for General Relativity, and perhaps the most important exact solution of the Einstein field equations for the gravitational field, as it describes the gravitational field outside a spherical star (as well as black holes, as it turns out ...). We will discuss this metric in great detail in sections 24-27.

The characteristic feature of metrics with Lorentzian signature is of course the presence of timelike and null (lightlike) directions, and thus in a pseudo-Riemannian space-time one has the same distinction between spacelike, timelike and lightlike separations as in Minkowski space(-time). Infinitesimal

- spacelike distances correspond to $ds^2 > 0$,
- timelike distances to $d\tau^2 = -ds^2 > 0$,
- and null or lightlike distances to $ds^2 = d\tau^2 = 0$.

Likewise, a vector $V^\mu(x)$ at a point x is called

- spacelike if $g_{\mu\nu}(x)V^\mu(x)V^\nu(x) > 0$,
- timelike if $g_{\mu\nu}(x)V^\mu(x)V^\nu(x) < 0$,
- and null or lightlike if $g_{\mu\nu}(x)V^\mu(x)V^\nu(x) = 0$,

and a curve $x^\mu(\lambda)$ is called spacelike if its tangent vector is everywhere spacelike etc.

Using the definition of a vector in general relativity (to be introduced in section 4), namely an object that transforms in the obvious way, with the Jacobi matrix, under coordinate transformations, one sees that $g_{\mu\nu}(x)V^\mu(x)V^\nu(x)$ is a scalar, i.e. invariant under coordinate transformations, and hence the statement that a vector is, say, spacelike is a coordinate-independent statement, as it should be.

Remarks:

1. When the metric (2.32) is (time-space) block-diagonal, i.e. when the mixed components $g_{0k} = 0$ (as in all of the above examples), then the timelike and spacelike directions are easy to distinguish by inspection. Typically then the "spatial" metric g_{ik} is positive definite, and thus necessarily $g_{00} < 0$.

2. When some of the g_{0k} are non-zero, on the other hand, one has a more intricate mixing of time- and space-directions. For example, consider the simple 2-dimensional metric
$$ds^2 = dx^2 + dy^2 + 2a\, dx\, dy \tag{2.39}$$
for some real parameter a. For any a, the coordinate lines of x and y are spacelike curves. However, this does not imply all by itself that the metric is Euclidean (just like the sum of two spacelike vectors in Minkowski space is not necessarily spacelike: it may be spacelike, timelike or null: e.g. if $v = (0,1)$, choose $w = (0,1), (1/2,-1), (1,-2)$ respectively).

   Indeed, by calculating the determinant of the metric,
$$\det \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} = 1 - a^2 \quad , \tag{2.40}$$
one sees that

   - when $a^2 < 1$, the metric has Euclidean signature
     (it is actually related to the standard 2-dimensional Euclidean metric by a linear transformation)
   - when $a^2 = 1$, the metric is degenerate
     (in this case, the line element can be written as $ds^2 = d(x \pm y)^2$, which is a 1-dimensional metric for the single coordinate $x \pm y$)
   - when $a^2 > 1$, the metric has Lorentzian signature
     (it is actually related to the standard 2-dimensional Minkowski metric by a linear transformation)

3. This mixing of time- and space-directions for a metric which is not block-diagonal can also be seen from the components of the inverse metric. Indeed, from (2.7), one finds
$$g_{0\nu} g^{\nu k} = g_{00} g^{0k} + g_{0i} g^{ik} = \delta_0^k = 0 \quad , \tag{2.41}$$
and thus (for $g_{00} \neq 0$)
$$g^{0k} = -\frac{1}{g_{00}} g_{0i} g^{ik} \quad . \tag{2.42}$$
Likewise from (2.7) one deduces
$$g_{i\nu} g^{\nu k} = g_{i0} g^{0k} + g_{ij} g^{jk} = \delta_i^k \quad . \tag{2.43}$$

In particular, this shows that in general (i.e. unless the off-diagonal components g_{0k} are all zero), the spatial components g^{ik} of the inverse metric are not the inverse of the spatial components g_{ij} of the metric. Rather, using (2.42) one has

$$\left(g_{ij} - \frac{1}{g_{00}} g_{i0} g_{j0}\right) g^{jk} = \delta_i^k \quad . \tag{2.44}$$

Outlook:

This ends our first brief encounter with metrics and geometries. At this point the question naturally arises how one can tell whether a given (perhaps complicated looking) metric is just the "flat" (Euclidean or Minkowski) metric written in other coordinates or whether it describes a genuinely new geometry.

We will see later that there is an object, the *Riemann curvature tensor*, constructed from the metric and its 1st and 2nd derivatives, which has the property that all of its components vanish if and only if the metric is a coordinate transform of the flat space Minkowski metric. Thus, given a metric, by calculating its curvature tensor one can decide if the metric is just the flat metric in disguise or not. The curvature tensor will be introduced in section 8, and the above statement will be established in section 11.2.

2.3 Geodesic Equation from the Extremisation of Proper Time

We have seen that the equation for a straight line in Minkowski space, written in arbitrary coordinates, is

$$\frac{d^2 x^\mu}{d\tau^2} + \gamma^\mu_{\nu\lambda} \frac{dx^\nu}{d\tau} \frac{dx^\lambda}{d\tau} = 0 \quad , \tag{2.45}$$

where the pseudo-force term $\gamma^\mu_{\nu\lambda}$ is given by (1.98). We have also seen in (1.111) (provided you checked this) that $\gamma^\mu_{\nu\lambda}$ can be expressed in terms of the metric (1.87) as

$$\gamma^\mu_{\nu\lambda} = \tfrac{1}{2} \eta^{\mu\rho} (\eta_{\rho\nu},_\lambda + \eta_{\rho\lambda},_\nu - \eta_{\nu\lambda},_\rho) \quad . \tag{2.46}$$

This gravitational force term is fictitious since it can globally be transformed away by going to the global inertial coordinates ξ^a. The equivalence principle suggests, however, that in general the equation for the worldline of a massive particle, i.e. a path that extremises proper time, in a true gravitational field described by a non-trivial metric $g_{\mu\nu}(x)$ (not related to the Minkowski metric by a coordinate transformation) is also of the above form.

We will now confirm this by deriving the equations for a timelike path that extremises proper time from a variational principle. These paths will be referred to as (timelike) *geodesics*. We will briefly return below to the (delicate) issue to which extent these can be regarded as world lines of actual massive particles.

Recall first of all from special relativity that the Lorentz-covariant description of the dynamics of a massive particle is based on describing the timelike worldline of the particle in the parametric form

$$\xi^a = \xi^a(\tau) \tag{2.47}$$

where τ is the proper time along the worldline,

$$d\tau^2 = -\eta_{ab} d\xi^a d\xi^b \ . \tag{2.48}$$

In particular, the 4-velocity

$$u^a = \frac{d\xi^a(\tau)}{d\tau} \tag{2.49}$$

is normalised as

$$\eta_{ab} u^a u^b = -1 \ . \tag{2.50}$$

The Lorentz-invariant action for a free massive particle with mass m is

$$S_0 = -m \int d\tau \ . \tag{2.51}$$

We can adopt the same set-up and action in the present setting. Thus we parametrise the worldlines by

$$x^\mu = x^\mu(\tau) \ , \tag{2.52}$$

with τ the proper time

$$d\tau^2 = -g_{\mu\nu}(x) dx^\mu dx^\nu \ , \tag{2.53}$$

invariant under general coordinate transformations (provided that one transforms the metric appropriately). The corresponding 4-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} \tag{2.54}$$

is again normalised as

$$g_{\mu\nu} u^\mu u^\nu = -1 \ , \tag{2.55}$$

and we are led to consider the coordinate-invariant Lagrangian

$$S_0[x] = -m \int d\tau = -m \int \sqrt{-g_{\mu\nu}(x) dx^\mu dx^\nu} \ . \tag{2.56}$$

Of course m drops out of the variational equations (as it should by the equivalence principle) and we will therefore ignore m in the following.

In order to perform the variation, it is useful to introduce an arbitrary auxiliary parameter λ in the initial stages of the calculation via

$$d\tau = (-g_{\mu\nu} \tfrac{dx^\mu}{d\lambda} \tfrac{dx^\nu}{d\lambda})^{1/2} d\lambda \ , \tag{2.57}$$

and to write

$$\int d\tau = \int (d\tau/d\lambda) d\lambda = \int (-g_{\mu\nu} \tfrac{dx^\mu}{d\lambda} \tfrac{dx^\nu}{d\lambda})^{1/2} d\lambda \ . \tag{2.58}$$

We are varying the paths

$$x^\mu(\tau) \to x^\mu(\tau) + \delta x^\mu(\tau) \tag{2.59}$$

keeping the end-points fixed, and will denote the τ-derivatives by $\dot{x}^\mu(\tau)$. Under this variation, the metric $g_{\mu\nu}(x)$ varies as

$$\delta g_{\mu\nu} = g_{\mu\nu},_\lambda \delta x^\lambda \quad . \tag{2.60}$$

By the standard variational procedure one then finds, first of all,

$$\delta \int d\tau = \tfrac{1}{2}\int (-g_{\mu\nu}\tfrac{dx^\mu}{d\lambda}\tfrac{dx^\nu}{d\lambda})^{-1/2} d\lambda \left[-\delta g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - 2g_{\mu\nu}\frac{d\delta x^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right] \quad . \tag{2.61}$$

Already at this stage we can revert from λ to τ, and the expression simplifies to

$$\delta \int d\tau = \tfrac{1}{2}\int d\tau \left[-(\delta g_{\mu\nu})\dot{x}^\mu \dot{x}^\nu - 2g_{\mu\nu}\frac{d\delta x^\mu}{d\tau}\dot{x}^\nu\right] \quad . \tag{2.62}$$

Integration by parts of the 2nd term (in order to eliminate the derivative of the variation) and use of (2.60) then leads to

$$\begin{aligned}\delta \int d\tau &= \frac{1}{2}\int d\tau \left[-g_{\mu\nu},_\lambda \dot{x}^\mu \dot{x}^\nu \delta x^\lambda + 2g_{\mu\nu}\ddot{x}^\nu \delta x^\mu + 2g_{\mu\nu},_\lambda \dot{x}^\lambda \dot{x}^\nu \delta x^\mu\right] \\ &= \int d\tau \left[g_{\mu\nu}\ddot{x}^\nu + \tfrac{1}{2}(g_{\mu\nu},_\lambda + g_{\mu\lambda},_\nu - g_{\nu\lambda},_\mu)\dot{x}^\nu \dot{x}^\lambda\right]\delta x^\mu\end{aligned} \tag{2.63}$$

after a suitable relabelling of the indices.

If we now adopt the definition (2.46) for an arbitrary metric,

$$\Gamma^\mu_{\ \nu\lambda} = g^{\mu\rho}\Gamma_{\rho\nu\lambda} = \tfrac{1}{2}g^{\mu\rho}(g_{\rho\nu},_\lambda + g_{\rho\lambda},_\nu - g_{\nu\lambda},_\rho) \quad , \tag{2.64}$$

we can write the result as

$$\delta \int d\tau = \int d\tau g_{\mu\nu}(\ddot{x}^\nu + \Gamma^\nu_{\ \rho\lambda}\dot{x}^\rho \dot{x}^\lambda)\delta x^\mu \quad . \tag{2.65}$$

Thus we see that indeed the equations for a timelike geodesic in an arbitrary gravitational field are

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \quad . \tag{2.66}$$

<u>Remarks:</u>

1. Given a metric $g_{\mu\nu}$, the quantities $\Gamma_{\rho\nu\lambda}$ and $\Gamma^\mu_{\ \nu\lambda}$ defined in (2.64) are known as the *Christoffel Symbols* of the 1st kind and 2nd kind.

2. The Christoffel symbols (2.64) play the role of the gravitational force term, and thus in this sense the components of the metric play the role of the gravitational potential. These Christoffel symbols play an important role not just in the geodesic equation but, as we will see later on, more generally in the definition of a covariant derivative operator and the construction of the curvature tensor, and thus ultimately also in the generally covariant description of the dynamics of the gravitational field itself.

3. Two elementary important properties of the Christoffel symbols are that they are symmetric in the second and third indices,

$$\Gamma_{\mu\nu\lambda} = \Gamma_{\mu\lambda\nu} \quad , \quad \Gamma^{\mu}_{\ \nu\lambda} = \Gamma^{\mu}_{\ \lambda\nu} \tag{2.67}$$

(this follows simply from the definition), and that symmetrising $\Gamma_{\mu\nu\lambda}$ over the first pair of indices one finds

$$\Gamma_{\mu\nu\lambda} + \Gamma_{\nu\mu\lambda} = g_{\mu\nu,\lambda} \tag{2.68}$$

(and this follows from noting that 4 of the 6 partial derivative terms of the metric cancel in this linear combination while 2 add up)

4. One can also consider spacelike paths that extremise (minimise) proper distance, by using the action

$$S_0 \sim \int ds \tag{2.69}$$

where

$$ds^2 = g_{\mu\nu}(x)dx^{\mu}dx^{\nu} \tag{2.70}$$

is the proper distance (or arc-length in the traditional terminology of the differnetial geometry of curves).

One should also consider massless particles, whose worldlines will be null (or lightlike) paths. However, in that case one can evidently not use proper time or proper distance, since these are by definition zero along a null path, $ds^2 = 0$. We will come back to this special case, and a unified description of the massive and massless case, below (section 2.5). In all cases, we will refer to the resulting paths as *geodesics*. If required, we add the qualifier "timelike", "spacelike" or "null", and this is meaningful and unambiguous since, as we will see below, a geodesic that is initially timelike will always remain timelike etc.

5. By definition, massive test particles are those particles that satisfy the above geodesic equation, i.e. that follow timelike geodesics in space-time. However, it needs to be borne in mind that this notion of a test particle is a fiction, in particular as it neglects the backreaction, i.e. the change in the background gravitational field due to the mass of the particle. Moreover, real particles either have a finite extent (in which case this finite size should play a role in their equations of motion) or are considered to be point-like. However, the notion of a point-like particle is extremely dangerous and delicate in general relativity: as we will see later, if a given total mass is concentrated in a sufficiently small region of space-time (and "point-like" certainly qualifies as "sufficiently small"), then one will end up with a black hole rather than with the description of a particle. The correct description of point particles in general relativity is a complicated issue and an active area of research.[5]

[5]See e.g. E. Poisson, A. Pound, I. Vega, *The Motion of Point Particles in Curved Spacetime*,

2.4 Geodesic Equation and Coordinate Transformations

As we have seen, in a coordinate system x^μ, with metric $g_{\mu\nu}$ and Christoffel symbols $\Gamma^\mu_{\nu\lambda}$ the geodesic equation takes the form (2.66)

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \ . \tag{2.71}$$

Since the coordinate system x^μ was arbitrary, evidently this equation is true for an arbitrary coordinate system. Nevertheless, when one changes from coordinates x^μ to y^α, the components of the metric and the Christoffel symbols will change (the latter in a somewhat complicated way). It is therefore of interest to find out explicitly how it comes about that the validity of (2.71) is independent of the coordinate system. It turns out that the geodesic equations in the two coordinate systems are simply related by the Jacobi matrix,

$$\frac{d^2y^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = \frac{\partial y^\alpha}{\partial x^\mu}\left[\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \ . \tag{2.72}$$

This generalises the result (1.112) that we had obtained before regarding coordinate transformations in Minkowski space, and the same remarks about covariance and tensors etc. apply. In particular, the invertibility of the Jacobi matrix allows us to conclude that

$$\frac{d^2y^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = 0 \quad \Leftrightarrow \quad \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \ , \tag{2.73}$$

as anticipated. I will now give 3 arguments to establish (2.72).

1. General Argument at the Level of the Variation of the Action

   That the geodesic equation transforms in this simple way (namely as a vector) should not come as a surprise. We obtained this equation as a variational equation. The Lagrangian itself is a scalar (invariant under coordinate transformations),

   $$d\tau = \sqrt{-g_{\alpha\beta}dy^\alpha dy^\beta} = \sqrt{-g_{\mu\nu}dx^\mu dx^\nu} \ . \tag{2.74}$$

   Moreover, the variations of y^α and x^μ are related by

   $$\delta y^\alpha = \frac{\partial y^\alpha}{\partial x^\mu}\delta x^\mu = J^\alpha_\mu \delta x^\mu \ . \tag{2.75}$$

   Putting these pieces together, one finds the desired result (2.72).

2. General Argument at the Level of the Euler-Lagrange Equations

   A general feature of the Lagrange formalism and the Euler-Lagrange equations is their covariance under what are known as "point-transformations" in classical

`arXiv:1102.0529 [gr-qc]` for a detailed discussion and many references (but you will need to acquire a solid understanding of tensor analysis first).

mechanics, arbitrary (possibly even time-dependent) coordinate transformations of the configuration space variables. In the case at hand, this reduces to the statement that the Euler-Lagrange equations with respect to x^μ and y^α respectively are related by the Jacobi matrix,

$$\frac{d}{d\tau}\frac{\partial L}{\partial x'^\mu} - \frac{\partial L}{\partial x^\mu} = J^\alpha_\mu \left[\frac{d}{d\tau}\frac{\partial L}{\partial y'^\alpha} - \frac{\partial L}{\partial y^\alpha}\right] \tag{2.76}$$

where $x'^\mu = dx^\mu/d\lambda$ etc. This also implies (2.72).

3. Specific Argument at the Level of the Geodesic Equation

   Knowing how the metric transforms under coordinate transformations,

$$g_{\alpha\beta} = J^\mu_\alpha J^\nu_\beta g_{\mu\nu} \quad , \tag{2.77}$$

   we can now also determine how the Christoffel symbols (2.64) and the geodesic equation transform. A straightforward but not particularly inspiring calculation (which you should nevertheless do) shows that under $x^\mu \to y^\alpha$ the Christoffel symbols are related by

$$\Gamma^\alpha_{\ \beta\gamma} = \Gamma^\mu_{\ \nu\lambda}\frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial x^\lambda}{\partial y^\gamma} + \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial y^\beta \partial y^\gamma} \quad , \tag{2.78}$$

   or

$$\Gamma^\alpha_{\ \beta\gamma} = J^\alpha_\mu J^\nu_\beta J^\lambda_\gamma \Gamma^\mu_{\ \nu\lambda} + J^\alpha_\mu \partial_\beta J^\mu_\gamma \quad . \tag{2.79}$$

   Thus, $\Gamma^\mu_{\ \nu\lambda}$ transforms inhomogenously under coordinate transformations. If only the first term on the right hand side were present, then $\Gamma^\mu_{\ \nu\lambda}$ would be a tensor. However, the second term is there precisely to compensate for the fact that $\ddot{x}^\mu$ is also not a tensor - the combined geodesic equation turns out to transform in a nice way under coordinate transformations.

   Namely, after another not terribly inspiring calculation (which you should nevertheless also do at least once in your life) , one finds precisely (2.72),

$$\frac{d^2 y^\alpha}{d\tau^2} + \Gamma^\alpha_{\ \beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = \frac{\partial y^\alpha}{\partial x^\mu}\left[\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \quad . \tag{2.80}$$

   An explicit proof of (2.79) and (2.80) is given at the end of this subsection. A more general result along these lines will be established in section 5.1 below, when we introduce the covariant derivative of a vector field.

Remarks:

1. Argument 3 in particular highlights the importance and significance of the Christoffel symbols: they are not a nuisance but a blessing, because they allow one to modify the definition of the acceleration $\ddot{x}^\mu = (d/d\tau)\dot{x}^\mu$ in such a way that it

behaves nicely under coordinate transformations, namely just like the velocity $\dot{x}^\mu$ itself. We will see later on, in section 5, that the Christoffel symbols allow us to accomplish this much more generally. Namely, with their aid we can define a (generally covariant) notion of a derivative that generalises the ordinary (or partial) derivative and maps tensors to tensors.

2. General covariance, i.e. invariance of the equations of motion under general coordinate transformations, as exhibited e.g. by the geodesic equation, is of course a desirable feature regardless of whether or not one is attempting to describe gravity. After all, the particle could not care less which coordinates we use to describe its motion, and therefore we should also formulate the equations of motion for a particle in a way that does not single out some preferred coordinate system or class of coordinate systems. This is precisely what is achieved by general covariance.

   However, here general covariance seems to have arisen somewhat coincidentally and spontaneously, and the relation between general covariance and gravity, or general covariance and the equivalence principle, may still appear to be somewhat mysterious at this point. The precise relation between the two concepts will be explained in section 4.1.

3. There is of course a very good physical reason for why the force term in the geodesic equation (quadratic in the 4-velocities) is not tensorial. This simply reflects the equivalence principle that locally, at a point (or in a sufficiently small neighbourhood of a point) you can eliminate the gravitational force by going to a freely falling (inertial) coordinate system. This would not be possible if the gravitational force term in the equation of motion for a particle were tensorial.

---

1. Proof of (2.79)

   For partial derivatives one has the chain rule $\partial_\gamma = J^\lambda_\gamma \partial_\lambda$ ("∂_λ is a covector"). Therefore for the partial derivatives of the metric one has

$$g_{\alpha\beta,\gamma} = (J^\mu_\alpha J^\nu_\beta g_{\mu\nu})_{,\gamma} = g_{\mu\nu,\lambda} J^\mu_\alpha J^\nu_\beta J^\lambda_\gamma + (J^\mu_{\alpha\gamma} J^\nu_\beta + J^\mu_\alpha J^\nu_{\beta\gamma}) g_{\mu\nu} \quad . \tag{2.81}$$

   Adding up the 3 terms comprising the Christoffel symbol $\Gamma_{\alpha\beta\gamma}$, one obtains

$$\begin{aligned} 2\Gamma_{\alpha\beta\gamma} &= g_{\alpha\beta,\gamma} + g_{\alpha\gamma,\beta} - g_{\beta\gamma,\alpha} \\ &= 2J^\mu_\alpha J^\nu_\beta J^\lambda_\gamma \Gamma_{\mu\nu\lambda} \\ &\quad + (J^\mu_{\alpha\gamma} J^\nu_\beta + J^\mu_\alpha J^\nu_{\beta\gamma} + J^\mu_{\alpha\beta} J^\nu_\gamma + J^\mu_\alpha J^\nu_{\gamma\beta} - J^\mu_{\beta\alpha} J^\nu_\gamma - J^\mu_\beta J^\nu_{\gamma\alpha}) g_{\mu\nu} \quad . \end{aligned} \tag{2.82}$$

   In the last line, the 3rd term cancels against the 5th (because $J^\mu_{\alpha\beta}$ is symmetric), the 1st term cancels against the 6th (because $J^\mu_{\alpha\gamma}$ and $g_{\mu\nu}$ are symmetric), while the 2nd and 4th term add up, so that one finds

$$\Gamma_{\alpha\beta\gamma} = J^\mu_\alpha J^\nu_\beta J^\lambda_\gamma \Gamma_{\mu\nu\lambda} + J^\mu_\alpha J^\nu_{\beta\gamma} g_{\mu\nu} \quad . \tag{2.83}$$

Now the hard work has been done. Raising the 1st index of the Christoffel symbol, using the inverse metric

$$g^{\alpha\delta} = g^{\sigma\rho} J^\alpha_\sigma J^\delta_\rho \quad , \tag{2.84}$$

it is now simple to see that one obtains the claimed result (4),

$$\Gamma^\alpha_{\ \beta\gamma} = g^{\alpha\delta}\Gamma_{\delta\beta\gamma} = J^\alpha_\mu J^\nu_\beta J^\lambda_\gamma \Gamma^\mu_{\ \nu\lambda} + J^\alpha_\mu J^\mu_{\beta\gamma} \quad . \tag{2.85}$$

For example, for the 2nd term one has (just using properties of inverse Jacobi matrices and metrics)

$$\begin{aligned} g^{\alpha\delta} J^\mu_\delta J^\nu_{\beta\gamma} g_{\mu\nu} &= g^{\sigma\rho} J^\alpha_\sigma J^\delta_\rho J^\mu_\delta J^\nu_{\beta\gamma} g_{\mu\nu} = g^{\sigma\rho} J^\alpha_\sigma \delta^\mu_\rho J^\nu_{\beta\gamma} g_{\mu\nu} \\ &= g^{\sigma\mu} J^\alpha_\sigma J^\nu_{\beta\gamma} g_{\mu\nu} = \delta^\sigma_\nu J^\alpha_\sigma J^\nu_{\beta\gamma} = J^\alpha_\nu J^\nu_{\beta\gamma} \end{aligned} \tag{2.86}$$

2. Proof of (2.72)

   The 4-velocities transform as vectors (the chain rule again), $\dot{y}^\alpha = J^\alpha_\mu \dot{x}^\mu$. Therefore for the acceleration one has

$$\ddot{y}^\alpha = J^\alpha_\mu \ddot{x}^\mu + J^\alpha_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \quad . \tag{2.87}$$

   Therefore

$$\begin{aligned} \ddot{y}^\alpha + \Gamma^\alpha_{\ \beta\gamma}\dot{y}^\beta\dot{y}^\gamma &= J^\alpha_\mu(\ddot{x}^\mu + \Gamma^\mu_{\ \nu\lambda} J^\nu_\beta J^\lambda_\gamma \dot{y}^\beta\dot{y}^\gamma) + J^\alpha_\mu J^\mu_{\beta\gamma}\dot{y}^\beta\dot{y}^\gamma + J^\alpha_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \\ &= J^\alpha_\mu(\ddot{x}^\mu + \Gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda) + (J^\alpha_\mu J^\mu_{\beta\gamma} + J^\alpha_{\mu\nu} J^\mu_\beta J^\nu_\gamma)\dot{y}^\beta\dot{y}^\gamma \end{aligned} \tag{2.88}$$

   The 1st term will give us the desired result, and cooperatively the 2nd term is identically zero because (use $\partial_\gamma = J^\nu_\gamma \partial_\nu$ again)

$$0 = (\delta^\alpha_\beta)_{,\gamma} = (J^\alpha_\mu J^\mu_\beta)_{,\gamma} = J^\alpha_{\mu\nu} J^\nu_\gamma J^\mu_\beta + J^\alpha_\mu J^\mu_{\beta\gamma} \quad . \tag{2.89}$$

Apology and Outlook:

You may feel that, after a promising start, some of the things that we have done subsequently (in particular in this subsection) look terribly messy. I agree, indeed they are! However, I can assure you that this is by far the messiest part of the entire lecture notes and that things will improve dramatically rather quickly.

Indeed, the main purpose and benefit of developing tensor calculus in the next couple of sections is to develop a formalism in the framework of which (among other things)

- one can avoid having to deal explicitly with objects that transform in complicated ways under coordinate transformations
- the transformation behaviour of any object is manifest (and does not have to be checked)
- it is straightforward to write down equations that are *generally covariant*, i.e. independent of the coordinate system in the sense that they are satisfied in all coordinate systems if and only if they are satisfied in one.

This tensor calculus formalism is simple, elegant and efficient and will then allow us to make rapid progress towards describing the dynamics in a (and subsequently of the) gravitational field in a way compatible with the Einstein equivalence principle.

2.5 Alternative Action Principles for Geodesics

As we have already noted in section 2.3, there is a problem with the action $S \sim m\int d\tau$ (2.56) for massless particles (null geodesics). For this reason and many other practical purposes (the square root in the action is awkward) it is much more convenient to use, instead of the action

$$S_0[x] = -m\int d\tau = -m\int d\lambda\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} \equiv \int d\lambda\, \mathcal{L}_0^\lambda \tag{2.90}$$

the simpler Lagrangian

$$\mathcal{L} = \tfrac{1}{2}g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \tag{2.91}$$

and action

$$S_1[x] = \int d\lambda\, \mathcal{L} \quad . \tag{2.92}$$

Let us first verify that S_1 really leads to the same equations of motion as S_0. Either by direct variation of the action, or by using the Euler-Lagrange equations

$$\frac{d}{d\lambda}\frac{\partial\mathcal{L}}{\partial(dx^\gamma/d\lambda)} - \frac{\partial\mathcal{L}}{\partial x^\gamma} = 0 \quad , \tag{2.93}$$

one finds that the action is extremised by solutions to the equation

$$\frac{d}{d\lambda}\left(g_{\gamma\beta}\frac{dx^\beta}{d\lambda}\right) = \tfrac{1}{2}g_{\alpha\beta,\gamma}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \quad . \tag{2.94}$$

The terms involving first derivatives of the metric cooperatively combine into the Christoffel symbols,

$$\left(\frac{d}{d\lambda}g_{\gamma\beta}\right)\frac{dx^\beta}{d\lambda} - \tfrac{1}{2}g_{\alpha\beta,\gamma}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = g_{\gamma\beta,\alpha}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} - \tfrac{1}{2}g_{\alpha\beta,\gamma}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \Gamma_{\gamma\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \quad . \tag{2.95}$$

Here we have used the fact that we can write

$$g_{\gamma\beta,\alpha}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \tfrac{1}{2}\left(g_{\gamma\beta,\alpha} + g_{\gamma\alpha,\beta}\right)\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \tag{2.96}$$

because

$$\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = \frac{dx^\beta}{d\lambda}\frac{dx^\alpha}{d\lambda} \tag{2.97}$$

is symmetric. Therefore one has

$$g_{\gamma\beta}\frac{d^2x^\beta}{d\lambda^2} + \Gamma_{\gamma\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \quad . \tag{2.98}$$

By raising the index (or multiplying with the inverse metric) one can write this as

$$\frac{d^2x^\gamma}{d\lambda^2} + \Gamma^\gamma_{\ \alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \quad . \tag{2.99}$$

This is identical to the geodesic equation derived from S_0 (with $\lambda \to \tau$, the proper time).

We will make extensive use of this simpler Lagrangian for geodesics throughout these notes. In particular, in practice the version (2.94) of the geodesic action is much more efficient and user-friendly than the standard form, because everything is expressed directly in terms of the metric and its first derivatives (neither does one need the inverse metric, nor does one have to assemble the derivatives of the metric into Christoffel symbols first).

Moreover, as will be explained in section 3.1 below, (2.94) actually also provides one with a fairly efficient method to determine (essentially read off) the Christoffel symbols, simply by comparing (2.94) with (2.98) or (2.99).

One important consequence of (2.99) is that the quantity $\mathcal{L}$ is a constant of motion, i.e. constant along the geodesic,

$$\frac{d^2x^\gamma}{d\lambda^2} + \Gamma^\gamma_{\ \alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \quad \Rightarrow \quad \frac{d}{d\lambda}\left(g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right) = 0 \quad . \tag{2.100}$$

This useful result can be understood and derived in a variety of ways:

- The least insightful way is just direct calculation. Nevertheless, this is straightforward and a good exercise in Γ-ology (and as such is left as an exercise).
- Alternatively, noting that $\mathcal{L}$ does not depend explicitly on λ, this result can be derived (as the corresponding conserved "energy") from Noether's theorem (cf. section 2.6 below for this argument).
- Yet another derivation will be given in section 5.8, using the concept of "parallel transport".

One obvious consequence of (2.100) is that, if one imposes the initial condition

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\bigg|_{\lambda=0} = \epsilon \ , \tag{2.101}$$

then this condition will be satisfied for all λ. This is as it should be. After all, something that starts off as a massless particle will remain a massless particle etc. In particular, therefore, even though with this choice of Lagrangian λ is a priori unrelated to proper time, say, this shows that one can choose $\epsilon = \mp 1$ for timelike (spacelike) geodesics, and λ can then be identified with proper time (proper distance), while the choice $\epsilon = 0$ sets the initial conditions appropriate to massless particles (for which λ is then not related to proper time or proper distance).

Moreover, the constancy of $\mathcal{L}$ for solutions of the Euler-Lagrange equations is the reason (or one explanation for) why $\mathcal{L}$ and $\sqrt{\pm\mathcal{L}}$ (more generally any monotonic function $f(\mathcal{L})$ of $\mathcal{L}$) give rise to equivalent equations of motion.

Indeed, this can be seen by simply comparing the Euler-Lagrange equations for $f(\mathcal{L})$ with those for $\mathcal{L}$. Denoting the Euler-Lagrange equations for a Lagrangian $\mathcal{L}$ by

$$\mathcal{E}_\gamma^{\mathcal{L}} = \frac{d}{d\lambda}\frac{\partial\mathcal{L}}{\partial(dx^\gamma/d\lambda)} - \frac{\partial\mathcal{L}}{\partial x^\gamma} \ , \tag{2.102}$$

for any $\mathcal{L}$ and any f one has

$$\mathcal{E}_\gamma^{f(\mathcal{L})} = f'(\mathcal{L})\mathcal{E}_\gamma^{\mathcal{L}} + f''(\mathcal{L})\left(\frac{d}{d\lambda}\mathcal{L}\right)\frac{\partial\mathcal{L}}{\partial(dx^\gamma/d\lambda)} \ . \tag{2.103}$$

Thus, if

$$\mathcal{E}_\gamma^{\mathcal{L}} = 0 \quad \Rightarrow \quad \frac{d}{d\lambda}\mathcal{L} = 0 \ , \tag{2.104}$$

then one has

$$\mathcal{E}_\gamma^{\mathcal{L}} = 0 \quad \Rightarrow \quad \mathcal{E}_\gamma^{f(\mathcal{L})} = 0 \ , \tag{2.105}$$

and if f is monotonic ($f' \neq 0$ everywhere), then one also has the converse,

$$f' \neq 0 \quad \Rightarrow \quad \left(\mathcal{E}_\gamma^{\mathcal{L}} = 0 \quad \Leftrightarrow \quad \mathcal{E}_\gamma^{f(\mathcal{L})} = 0\right) \ . \tag{2.106}$$

2.6 On the Relation between the two Action Principles

Even though not strictly required in the following, it is nevertheless quite instructive in its own right to try to understand and establish the precise relation between the two actions S_0 and S_1, and this is the subject of this subsection.

The first thing to notice is that S_0 is manifestly parametrisation-invariant, i.e. independent of how one parametrises the path. The reason for this is that

$$d\tau = (d\tau/d\lambda)d\lambda \tag{2.107}$$

is evidently independent of λ. This is not the case for S_1, which changes under parametrisations or, put more positively, singles out a preferred parametrisation (more precisely, as we will see below, a special class of parametrisations).

Thus, what is the relation (if any) between the two actions? In order to explain this, it will be useful to introduce an additional field $e(\lambda)$ (i.e. in addition to the $x^\alpha(\lambda)$), and a "master action" (or parent action) S which we can relate to both S_0 and S_1. Consider the action

$$S[x,e] = \tfrac{1}{2}\int d\lambda \left(e(\lambda)^{-1} g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} - m^2 e(\lambda)\right) = \int d\lambda \left(e(\lambda)^{-1}\mathcal{L} - \tfrac{1}{2}m^2 e(\lambda)\right) \tag{2.108}$$

The crucial property of this action is that it is parametrisation invariant provided that one declares $e(\lambda)$ to transform appropriately. It is easy to see that under a transformation $\lambda \to \bar{\lambda} = f(\lambda)$, with

$$\bar{x}^\alpha(\bar{\lambda}) = x^\alpha(\lambda) \qquad d\bar{\lambda} = f'(\lambda)d\lambda \tag{2.109}$$

the action $S[x,e]$ is invariant provided that $e(\lambda)$ transforms such that $e(\lambda)d\lambda$ is invariant, i.e.

$$\bar{e}(\bar{\lambda})d\bar{\lambda} \stackrel{!}{=} e(\lambda)d\lambda \quad \Rightarrow \quad \bar{e}(\bar{\lambda}) = e(\lambda)/f'(\lambda) \ . \tag{2.110}$$

Indeed, this is evident when one writes the action (2.108) in the form

$$S[x,e] = \tfrac{1}{2}\int e(\lambda)d\lambda \left(e(\lambda)^{-2} g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} - m^2\right) \tag{2.111}$$

and notes that $d\lambda$ and $e(\lambda)$ only appear in the combinations $e(\lambda)d\lambda$ and $e(\lambda)^{-1}(d/d\lambda)$.

Now what is the relation between the action $S[x,e]$ and the two "standard" actions $S_0[x]$ and $S_1[x]$?

- Courtesy of this parametrisation invariance, we can always choose a "gauge" in which $e(\lambda) = 1$. With this choice, the action $S[x,e]$ manifestly reduces to the action $S_1[x]$ modulo an irrelevant field-independent constant,

  $$S[x, e = 1] = \int d\lambda\, \mathcal{L} - \tfrac{1}{2}m^2 \int d\lambda = S_1[x] + \text{const.} \ . \tag{2.112}$$

  Thus we can regard S_1 as a gauge-fixed version of S (no wonder it is not parametrisation invariant ...). We will come back to the small residual gauge invariance (reparametrisations that preserve the gauge condition $e(\lambda) = 1$) below.

- Alternatively, instead of fixing the gauge, we can try to eliminate $e(\lambda)$ (which appears purely algebraically, i.e. without derivatives, in the action) by its equation of motion. Varying $S[x,e]$ with respect to $e(\lambda)$, one finds the constraint

  $$g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} + m^2 e(\lambda)^2 = 0 \ . \tag{2.113}$$

This is just the usual mass-shell condition in disguise. It suggests that a better gauge fixing than $e(\lambda) = 1$ would have been $e(\lambda) = m^{-1}$. However, the sole effect of this would have been to replace $\mathcal{L}$ in (2.112) by $m\mathcal{L}$,

$$e(\lambda) = 1 \to e(\lambda) = m^{-1} \quad \Rightarrow \quad \mathcal{L} \to m\mathcal{L} \ . \tag{2.114}$$

In any case, for a massive particle, $m^2 \neq 0$, one can alternatively solve (2.113) for $e(\lambda)$,

$$e(\lambda) = m^{-1}\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} \ . \tag{2.115}$$

Using this to eliminate $e(\lambda)$ from the action, one finds

$$S[x, e = m^{-1}\sqrt{\ldots}] = -m\int d\lambda\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} = -m\int d\tau = S_0[x] \ . \tag{2.116}$$

Thus for $m^2 \neq 0$ we find exactly the original action (integral of the proper time) $S_0[x]$ (and since we have not touched or fixed the parametrisation invariance, no wonder that S_0 is parametrisation invariant).

Thus we have elucidated the common origin of the actions S_0 and S_1 for a massive particle.

The perspective provided by the parent action $S[x, e]$ also gives some further insights. For example, an added benefit of the parent action $S[x, e]$ is that it also makes perfect sense for a massless particle. For $m^2 = 0$, the mass shell condition

$$g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \tag{2.117}$$

says that these particles move along null lines, and the action reduces to

$$S[x, e] = \tfrac{1}{2}\int d\lambda\, e(\lambda)^{-1} g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \tag{2.118}$$

which is parametrisation invariant but can (as in the massive case) be fixed to $e(\lambda) = 1$, upon which the action reduces to $S_1[x]$. Thus we see that $S_1[x]$ indeed provides a simple and unified action for both massive and massless particles, and in both cases the resulting equation of motion is the (affinely parametrised) geodesic equation (2.99),

$$\frac{d^2x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\ \beta\gamma}\frac{dx^\beta}{d\lambda}\frac{dx^\gamma}{d\lambda} = 0 \ . \tag{2.119}$$

<u>REMARKS:</u>

1. The infinitesimal form of the invariance of the action $S[x, e]$ under (2.109) and (2.110) is obtained by considering the infinitesimal transformation of $x^\alpha(\lambda)$ and

$e(\lambda)$ induced by an infinitesimal transformation $\bar{\lambda} = \lambda + \epsilon(\lambda)$,

$$\begin{aligned} \delta\lambda = \epsilon(\lambda) \quad \Rightarrow \quad \delta x^\alpha(\lambda) &= \epsilon(\lambda)\frac{dx^\alpha(\lambda)}{d\lambda} \\ \delta e(\lambda) &= \frac{d\epsilon(\lambda)}{d\lambda}e(\lambda) + \epsilon(\lambda)\frac{de(\lambda)}{d\lambda} \quad . \end{aligned} \tag{2.120}$$

Here the (at first perhaps somewhat peculiar looking) transformation behaviour of the auxiliary field $e(\lambda)$ arises from the transformation behaviour (2.110) by setting

$$\delta e(\lambda) = e(\lambda) - \bar{e}(\lambda) \tag{2.121}$$

and calculating (keeping at most linear terms in $\epsilon(\lambda)$)

$$\begin{aligned} e(\lambda) = f'(\lambda)\bar{e}(\bar{\lambda}) &= (1 + \epsilon'(\lambda))\, \bar{e}(\lambda + \epsilon(\lambda)) \\ &= (1 + \epsilon'(\lambda))\, (\bar{e}(\lambda) + \epsilon(\lambda)\bar{e}'(\lambda)) \\ &= \bar{e}(\lambda) + \epsilon'(\lambda)\bar{e}(\lambda) + \epsilon(\lambda)\bar{e}'(\lambda) \quad . \end{aligned} \tag{2.122}$$

Under this infinitesimal transformation, the Lagrangian

$$\mathcal{L}_e = e(\lambda)^{-1}\mathcal{L} - \tfrac{1}{2}m^2 e(\lambda) \tag{2.123}$$

of the action $S[x,e]$ (2.108) transforms as

$$\delta\mathcal{L}_e = \frac{d}{d\lambda}(\epsilon(\lambda)\mathcal{L}_e) \quad , \tag{2.124}$$

implying the invariance of the action.

2. We saw in (2.100), that the Lagrangian $\mathcal{L}$ itself is a constant of motion,

$$\frac{d}{d\lambda}\left(g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right) = 0 \tag{2.125}$$

for a solution to the geodesic equation. From the present (action-based) perspective it is most useful to think of this as the conserved quantity associated (via Noether's theorem) to the invariance of the action $S_1[x]$ under translations in λ. Note that evidently $S_1[x]$ has this invariance (as there is no explicit dependence on λ) and that this invariance is precisely the residual parametrisation invariance $f(\lambda) = \lambda + a$, $f'(\lambda) = 1$, that leaves invariant the "gauge" condition $e(\lambda) = 1$. For an infinitesimal constant λ-translation one has $\delta x^\alpha(\lambda) = dx^\alpha/d\lambda$ etc., so that

$$\frac{\partial}{\partial\lambda}\mathcal{L} = 0 \quad \Rightarrow \quad \delta\mathcal{L} = \frac{d}{d\lambda}\mathcal{L} = \frac{d}{d\lambda}\left(\frac{\partial\mathcal{L}}{\partial(dx^\alpha/d\lambda)}\frac{dx^\alpha}{d\lambda}\right) + \quad \text{Euler} - \text{Lagrange} \quad . \tag{2.126}$$

Thus via Noether's theorem the associated conserved charge for a solution to the Euler-Lagrange equations is the Legendre transform

$$\mathcal{H} = \left(\frac{\partial\mathcal{L}}{\partial(dx^\alpha/d\lambda)}\frac{dx^\alpha}{d\lambda}\right) - \mathcal{L} \tag{2.127}$$

of the Lagrangian (also known as the Hamiltonian, once expressed in terms of the momenta). In the case at hand, with the Lagrangian $\mathcal{L}$ consisting of a purely quadratic term in the velocities (the $dx^\alpha/d\lambda$), the Hamiltonian is equal to the Lagrangian, and hence the Lagrangian $\mathcal{L}$ itself is conserved,

$$\mathcal{H} = \mathcal{L} \quad , \quad \frac{d}{d\lambda}\mathcal{L}|_{\text{solution}} = 0 \quad . \tag{2.128}$$

3. The above non-trivial Hamiltonian associated to the Lagrangian $\mathcal{L}$ should be contrasted with the Hamiltonian associated to the action S_0 with Lagrangian (2.90)

$$\mathcal{L}_0^\lambda = -m\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} \quad , \tag{2.129}$$

which turns out to be zero identically. Indeed, the canonical momenta are

$$p_\alpha = mg_{\alpha\beta}(dx^\beta/d\lambda)/\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} = mg_{\alpha\beta}(dx^\beta/d\tau) \quad , \tag{2.130}$$

which evidently satisfy the mass shell condition

$$p_\alpha p^\alpha + m^2 = 0 \tag{2.131}$$

and lead to the Hamiltonian

$$\mathcal{H}_0 = p_\alpha(dx^\alpha/d\lambda) - \mathcal{L}_0^\lambda = 0 \quad . \tag{2.132}$$

This vanishing of the Hamiltonian is strictly related to the reparametrisation invariance of the action S_0.

2.7 Affine and Non-affine Parametrisations

To understand the significance of how one parametrises the geodesic, observe that the geodesic equation itself,

$$\ddot{x}^\mu + \Gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \quad , \tag{2.133}$$

is not reparametrisation invariant. Indeed, consider a change of parametrisation $\tau \to \sigma = f(\tau)$. Then

$$\frac{dx^\mu}{d\tau} = \frac{df}{d\tau}\frac{dx^\mu}{d\sigma} \quad , \tag{2.134}$$

and therefore the geodesic equation written in terms of σ reads

$$\frac{d^2x^\mu}{d\sigma^2} + \Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\sigma}\frac{dx^\lambda}{d\sigma} = -\frac{\ddot{f}}{\dot{f}^2}\frac{dx^\mu}{d\sigma} \quad . \tag{2.135}$$

Thus the geodesic equation retains its form only under affine changes of the proper time parameter τ, $f(\tau) = a\tau + b$, and parameters $\sigma = f(\tau)$ related to τ by such an affine transformation are known as affine parameters.

From the first variational principle, based on S_0, the term on the right hand side arises in the calculation of (2.63) from the integration by parts if one does not switch back from λ to the affine parameter τ. The second variational principle, based on S_1 and the Lagrangian $\mathcal{L}$, on the other hand, always and automatically yields the geodesic equation in affine form.

Conversely, if we find a curve that satisfies

$$\frac{d^2x^\mu}{d\sigma^2} + \Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\sigma}\frac{dx^\lambda}{d\sigma} = \kappa(\sigma)\frac{dx^\mu}{d\sigma} \quad , \tag{2.136}$$

for some function $\kappa(\sigma)$ (the *inaffinity*), we can deduce that this curve is the trajectory of a geodesic, but that it is simply not parametrised by an affine parameter (like proper time in the case of a timelike curve). Comparison of (2.135) and (2.136) shows that, given $\kappa(\sigma)$, an affine parameter τ is determined by

$$\kappa(f(\tau)) = -\frac{\ddot{f}}{\dot{f}^2} \quad \Leftrightarrow \quad \kappa(\sigma) = \frac{d}{d\sigma}\ln\frac{d\tau}{d\sigma} \tag{2.137}$$

or

$$\frac{d\tau}{d\sigma} = \mathrm{e}^{\int^\sigma ds\, \kappa(s)} \quad . \tag{2.138}$$

The two integration constants, the first hidden in the lower limit of integration in the exponent and the second in the additive constant arising from integrating $d\tau/d\sigma$, are precisely the two constants a, b that parametrise the freedom in the choice of affine parameter, $\tau \to a\tau + b$.

In the following, whenever we talk about geodesics we will practically always have in mind the variational principle based on S_1 leading to the geodesic equation (2.119) in affinely parametrised form.

However, it should be kept in mind that sometimes non-affine parameters appear naturally. For instance, it is occasionally convenient to parametrise timelike geodesics in a geometry with coordinates $x^\alpha = (x^0 = t, x^k)$ not by $x^\alpha = x^\alpha(\tau)$, where τ is the proper time along the geodesic, but rather as $x^k = x^k(t)$. This is the same curve, but described with respect to coordinate time (which could for instance agree with the proper time of some other, perhaps static, observer). The curve $t \to (t, x^k(t))$ will not be an affinely parametrised curve unless t itself satisfies the geodesic equation

$$\ddot{t} = 0 \quad \Leftrightarrow \quad t = a\tau + b \ . \tag{2.139}$$

One occasion where this will play a role (and from where I have borrowed the symbol κ for the "inaffinity") is in our discussion, much later, of the horizon of a black hole, where the lack of a certain coordinate to be an affine parameter is directly related to the physical properties of black holes (see section 27.10). In this context κ is known as the surface gravity of a black hole.

2.8 Example: Geodesics in $\mathbb{R}^2$ in Polar Coordinates

It is high time to consider an example. We will consider the simplest non-trivial metric, namely the standard Euclidean metric on $\mathbb{R}^2$ in polar coordinates. Thus the line element is

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2 \tag{2.140}$$

and the non-zero components of the metric are

$$g_{xx} = g_{yy} = 1 \tag{2.141}$$

and

$$g_{rr} = 1 \quad , \quad g_{\phi\phi} = r^2 \quad . \tag{2.142}$$

respectively. Since this metric is diagonal, the non-zero components of the inverse metric $g^{\mu\nu}$ are

$$g^{xx} = g^{yy} = 1 \tag{2.143}$$

and

$$g^{rr} = 1 \quad , \quad g^{\phi\phi} = r^{-2} \tag{2.144}$$

respectively.

A reminder on notation (cf. the discussion leading to (2.13)): since μ, ν in $g_{\mu\nu}$ are coordinate indices, we should really have called $x^1 = r$, $x^2 = \phi$, say, and written $g_{11} = 1, g_{22} = r^2$, etc. However, writing g_{rr} etc. is more informative and useful since one then knows that this is the (rr)-component of the metric without having to remember if one called $r = x^1$ or $r = x^2$. In the following we will frequently use this kind of notation when dealing with a specific coordinate system, while we retain the index notation $g_{\mu\nu}$ etc. for general purposes.

Let us now look at the geodesic equations for this metric, first in the Cartesian coordinates (x, y) and then in the polar coordinates (r, ϕ).

1. Cartesian coordinates

   Since the metric in Cartesian coordinates is the constant Euclidean metric $g_{\mu\nu} = \delta_{\mu\nu}$, all the partial derivatives of the metric are zero, and therefore also all the Christoffel symbols are zero. The geodesic equations thus take the form

   $$\ddot{x} = \ddot{y} = 0 \quad . \tag{2.145}$$

   These equations could also have been obtained as the Euler-Lagrange equations of the Lagrangian

   $$\mathcal{L} = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2) \quad . \tag{2.146}$$

   The general solution is

   $$x(s) = as + b \quad , \quad y(s) = cs + d \quad . \tag{2.147}$$

Combining these two, one finds the standard representation

$$y = kx + e \tag{2.148}$$

for a straight line in $\mathbb{R}^2$, with slope k and intercept e.

2. Polar coordinates

   Now let us consider the same problem in polar coordinates. The crucial point here is that in these coordinates the geodesic equations will not simply be $\ddot{r} = \ddot{\phi} = 0$, but that there are additional terms arising

   - either from the non-linear coordinate transformation between Cartesian and polar coordinates
   - or equivalently from the fact that the coefficients of the metric are not constant in polar coordinates.

   Taking the latter point of view, the Christoffel symbols of this metric are to be calculated from

$$\Gamma_{\mu\nu\lambda} = \tfrac{1}{2}(g_{\mu\nu,\lambda} + g_{\mu\lambda,\nu} - g_{\nu\lambda,\mu}) \quad . \tag{2.149}$$

   Since the only non-trivial derivative of the metric is $g_{\phi\phi,r} = 2r$, only Christoffel symbols with exactly two ϕ's and one r are non-zero,

$$\begin{aligned} \Gamma_{r\phi\phi} &= \tfrac{1}{2}(g_{r\phi,\phi} + g_{r\phi,\phi} - g_{\phi\phi,r}) = -r \\ \Gamma_{\phi\phi r} &= \Gamma_{\phi r\phi} = r \quad . \end{aligned} \tag{2.150}$$

   Thus, since the metric is diagonal, the non-zero $\Gamma^{\mu}{}_{\nu\lambda}$ are

$$\begin{aligned} \Gamma^{r}{}_{\phi\phi} &= g^{r\mu}\Gamma_{\mu\phi\phi} = g^{rr}\Gamma_{r\phi\phi} = -r \\ \Gamma^{\phi}{}_{r\phi} &= \Gamma^{\phi}{}_{\phi r} = g^{\phi\mu}\Gamma_{\mu r\phi} = g^{\phi\phi}\Gamma_{\phi r\phi} = \frac{1}{r} \quad . \end{aligned} \tag{2.151}$$

   Note that here it was even convenient to use a hybrid notation, as in $g^{r\mu}$, where r is a coordinate and μ is a coordinate index. Once again, it is very convenient to permit oneself to use such a mixed notation.

   In any case, having assembled all the Christoffel symbols, we can now write down the geodesic equations (once again in the convenient hybrid notation). For r one has

$$\ddot{r} + \Gamma^{r}{}_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu} = 0 \quad , \tag{2.152}$$

   which, since the only non-zero $\Gamma^{r}{}_{\mu\nu}$ is $\Gamma^{r}{}_{\phi\phi}$, reduces to

$$\ddot{r} - r\dot{\phi}^2 = 0 \quad . \tag{2.153}$$

   Likewise for ϕ one finds

$$\ddot{\phi} + \frac{2}{r}\dot{\phi}\dot{r} = 0 \quad . \tag{2.154}$$

Here the factor of 2 arises because both $\Gamma^{\phi}_{\ r\phi}$ and $\Gamma^{\phi}_{\ \phi r} = \Gamma^{\phi}_{\ r\phi}$ contribute.

<u>Remarks:</u>

(a) This equation is supposed to describe geodesics in $\mathbb{R}^2$, i.e. straight lines. This can be verified in general (but, in general, polar coordinates are of course not particularly well suited to describe straight lines). However, it is easy to find a special class of solutions to the above equations, namely curves with $\dot{\phi} = \ddot{r} = 0$. These correspond to paths of the form

$$(r(s), \phi(s)) = (s, \phi_0) \ , \tag{2.155}$$

which are a special case of straight lines, namely straight lines through the origin.

(b) The geodesic equations can of course also be derived as the Euler-Lagrange equations of the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) \ . \tag{2.156}$$

Indeed, one has

$$\begin{aligned} \frac{d}{ds}\frac{\partial\mathcal{L}}{\partial\dot{r}} - \frac{\partial\mathcal{L}}{\partial r} &= \ddot{r} - r\dot{\phi}^2 = 0 \\ \frac{d}{ds}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} - \frac{\partial\mathcal{L}}{\partial\phi} &= r^2\ddot{\phi} + 2r\dot{r}\dot{\phi} = 0 \ , \end{aligned} \tag{2.157}$$

which are obviously identical to the equations derived above.

(c) You may have the impression that getting the geodesic equation in this way, rather than via calculation of the Christoffel symbols first, is much simpler. I agree wholeheartedly. Not only is the Lagrangian approach the method of choice to determine the geodesic equations. It is also frequently the most efficient method to determine the Christoffel symbols. This will be described in section 3.1.

(d) Another advantage of the Lagrangian formulation is, as in classical mechanics, that it makes it much easier to detect and exploit symmetries. Indeed, you may have already noticed that the above second-order equation for ϕ is overkill. Since the Lagrangian does not depend on ϕ (i.e. it is invariant under rotations), one has

$$\frac{d}{ds}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = 0 \ , \tag{2.158}$$

which means that $\partial\mathcal{L}/\partial\dot{\phi}$ is a constant of motion, the angular momentum L,

$$\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = r^2\dot{\phi} = L \ . \tag{2.159}$$

This equation is a first integral of the second-order equation for ϕ. We will come back to this in somewhat more generality below.

The next simplest example to discuss would be the two-sphere with its standard metric $d\theta^2 + \sin^2\theta d\phi^2$. It will appear, in bits and pieces, in section 3.1 to illustrate the general remarks.

2.9 Example: Geodesics for Ultrastatic and Direct Product Metrics

As another example, let us consider the ultrastatic metrics introduced in (2.34) with coordinates $x^\mu = (t, x^k)$ and line-element

$$ds^2 = -dt^2 + \tilde{g}_{ik}(x)dx^i dx^k \quad . \tag{2.160}$$

Because $g_{00} = -1$, $g_{0k} = 0$, and the $g_{ik} = \tilde{g}_{ik}$ are time-independent, all Christoffel symbols with at least one x^0- or t-index are zero,

$$\Gamma_{0\mu\nu} = \Gamma_{\mu 0\nu} = \Gamma_{\mu\nu 0} = 0 \quad , \tag{2.161}$$

and the purely spatial components of the Christoffel symbols agree with those of the spatial metric,

$$\Gamma^i{}_{jk} = \tilde{\Gamma}^i{}_{jk} \quad . \tag{2.162}$$

Therefore the geodesic equations read, for the t-component,

$$\ddot{t} = 0 \quad , \tag{2.163}$$

and for the spatial components

$$\ddot{x}^i + \tilde{\Gamma}^i{}_{jk}\dot{x}^j\dot{x}^k = 0 \tag{2.164}$$

where the dot denotes a derivative with respect to the affine parameter τ. The first equation tells us that

$$\ddot{t} = 0 \quad \Leftrightarrow \quad t(\tau) = a\tau + t_0 \quad . \tag{2.165}$$

Thus provided that $a \neq 0$ we can use t instead of τ to parametrise the paths (and in the present case t is then also an affine parameter, cf. the discussion in section 2.7 in connection with (2.139)), and then one can rewrite the spatial equations as equations for $x^i = x^i(t)$,

$$\frac{d^2x^i}{dt^2} + \tilde{\Gamma}^i{}_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0 \quad . \tag{2.166}$$

Therefore the solutions to the space-time geodesic equations have the form

$$x^\alpha(t) = (t, x^i(t)) \tag{2.167}$$

where $x^i(t)$ is an affinely parametrised geodesic for the metric $\tilde{g}_{ij}$. When $a = 0$, one cannot change variables from τ to t because $t = t_0$ is fixed. One is then necessarily dealing with spacelike geodesics in space-time and the solutions have the form

$$x^\alpha(\tau) = (t_0, x^i(\tau)) \tag{2.168}$$

where $x^i(\tau)$ is again an affinely parametrised geodesic for the metric $\tilde{g}_{ij}$.

These sorts of considerations evidently generalise to more general metrics of this direct product form,

$$ds^2 = g_{ab}(y)dy^a dy^b + g_{ik}(x)dx^i dx^k \quad , \tag{2.169}$$

with the conclusion that geodesics in such space-times have the form $(y^a(\tau), x^i(\tau))$ with $y^a(\tau)$ and $x^i(\tau)$ individually solutions of the geodesic equations for the metric $g_{ab}(y)$ respectively $g_{ik}(x)$.

3 Geodesics and Motion in a Gravitational Field

3.1 Consequences and Uses of the Euler-Lagrange Equations

Recall from above that the geodesic equation for a metric $g_{\mu\nu}$ can be derived from the Lagrangian $\mathcal{L} = (1/2)g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$

$$\frac{d}{d\tau}\frac{\partial\mathcal{L}}{\partial\dot{x}^\mu} - \frac{\partial\mathcal{L}}{\partial x^\mu} = 0 \quad . \tag{3.1}$$

This has several immediate consequences which are useful for the determination of Christoffel symbols and geodesics in practice.

1. Conserved charges / first integrals of the geodesic equation

   Just as in classical mechanics, a coordinate the Lagrangian does not depend on explicitly (a cyclic coordinate) leads to a conserved quantity, associated with the translation invariance of the system in that direction. In the present context this means that if, say, $\partial\mathcal{L}/\partial x^1 = 0$ (this means that the coeffcients of the metric do not depend on x^1), then the corresponding momentum

   $$p_1 = \partial\mathcal{L}/\partial\dot{x}^1 = g_{1\nu}\dot{x}^\nu \tag{3.2}$$

   is conserved along the geodesic.

   Remarks:

   (a) One might perhaps have wanted to argue that the definition (and interpretation) of conserved momenta should be based on the physical Lagrangian (2.90)

   $$\mathcal{L}_0^\lambda = -m\sqrt{-g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}} \tag{3.3}$$

   with action $S = -m\int d\tau$, but this makes no difference since the two momenta are essentially equal: one has

   $$\frac{\partial\mathcal{L}_0^\lambda}{\partial(dx^1/d\lambda)} = mp_1 \tag{3.4}$$

   with p_1 as defined in (3.2), so that this just supplies us with the additional information that the momenta obtained from the Lagrangian $\mathcal{L}$ should (for a massive particle) be interpreted as momenta per unit mass. This discrepancy could have been avoided by working with the Lagrangian $m\mathcal{L}$ (alternatively: fixing the gauge $e(\lambda) = m^{-1}$ in section 2.5, see (2.114)), but unless or until one starts coupling the particle to fields other than the gravitational field it is unnecessary (and a nuisance) to carry m around all the time.

(b) For example, on the two-sphere the Lagrangian reads

$$\mathcal{L} = \tfrac{1}{2}(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \ . \tag{3.5}$$

The angle ϕ is a cyclic variable and the angular momentum (actually angular momentum per unit mass for a massive particle)

$$p_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \sin^2\theta\dot{\phi} \tag{3.6}$$

is a conserved quantity. This generalises to conservation of angular momentum for a particle moving in an arbitrary spherically symmetric gravitational field.

(c) Likewise, if the metric is independent of the time coordinate $x^0 = t$, the corresponding conserved quantity

$$p_0 = g_{0\nu}\dot{x}^\nu \equiv -E \tag{3.7}$$

has the interpretation as minus the energy (per unit mass) of the particle, "minus" because, with our sign conventions, $p_0 = -E$ in special relativity. We will discuss the relation between this notion of energy and the notion of energy familiar from special relativity (this requires an asymptotically Minkowski-like metric) in more detail in section 25.3.

(d) We will discuss in more detail in section 3.2 (and then again in sections 9 and 10) how to detect and describe symmetries and conserved charges in coordinate systems in which the symmetries are not as manifest (via cyclic variables) as above.

2. Reading off (some) geodesics directly from the metric

Another immediate consequence is the following: consider a space or space-time with coordinates $\{y, x^\mu\}$ and a metric of the form

$$ds^2 = \pm dy^2 + g_{\mu\nu}(x,y)dx^\mu dx^\nu \ . \tag{3.8}$$

Then the coordinate lines of y are geodesics.

The quickest way to see this is (as usual) from the Lagrangian point of view. Indeed, since the Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}(\pm\dot{y}^2 + g_{\mu\nu}(y,x)\dot{x}^\mu\dot{x}^\nu) \ , \tag{3.9}$$

the Euler-Lagrange equations are of the form

$$\begin{aligned} &\pm\ddot{y} - \tfrac{1}{2}g_{\mu\nu,y}\,\dot{x}^\mu\dot{x}^\nu = 0 \\ &\ddot{x}^\mu + \text{terms proportional to } \dot{x} = 0 \ . \end{aligned} \tag{3.10}$$

Therefore $\dot{x}^\mu = 0$, $\ddot{y} = 0$ is a solution of the geodesic equation, and it describes motion along the coordinate lines of y.

Alternatively, this special form of the metric implies that any Christoffel symbol with at least two y-indices is zero, and the conclusion then follows in the same way as above.

<u>Remarks:</u>

(a) In the case of the two-sphere, with its metric $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$, this translates into the familiar statement that the great circles, the coordinate lines of $y = \theta$, are geodesics.

(b) The result is also valid when y is a timelike coordinate. For example, consider a space-time with coordinates (t, x^i) and metric (2.36)

$$ds^2 = -dt^2 + a(t)^2 \tilde{g}_{ij}(x) dx^i dx^j \quad . \tag{3.11}$$

In such a cosmological space-time, there is, according to the above result, a privileged class of freely falling (i.e. geodesic) observers, namely those that stay at fixed values of the spatial coordinates x^i. For such *comoving observers*, the coordinate-time t coincides with their proper time τ.

(c) In general, these preferred geodesics are orthogonal to the hypersurfaces of constant y, and coordinates in which the metric (locally) takes such a form in the neighbourhood of some timelike or spacelike hypersurface are occasionally called *Gaussian normal coordinates*.

3. Using the Euler-Lagrange equations to determine the Christoffel symbols

Finally, as mentioned and observed above, the Euler-Lagrange form of the geodesic equations frequently provides the most direct way of calculating Christoffel symbols - by comparing the Euler-Lagrange equations with the expected form of the geodesic equation in terms of Christoffel symbols. More precisely, by rewriting the Euler-Lagrange equations (2.94)

$$\frac{d}{d\lambda}\left(g_{\gamma\beta}\frac{dx^\beta}{d\lambda}\right) = \tfrac{1}{2} g_{\alpha\beta,\gamma}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} \quad . \tag{3.12}$$

in the form

$$\ddot{x}^\gamma + \text{terms proportional to } \dot{x}\dot{x} = 0 \quad , \tag{3.13}$$

and comparing with the geodesic equation

$$\frac{d^2 x^\gamma}{d\lambda^2} + \Gamma^\gamma_{\ \alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \quad , \tag{3.14}$$

one can read off the $\Gamma^\gamma_{\ \alpha\beta}$.

Remarks:

(a) Careful - in this and similar calculations beware of factors of 2:

$$\Gamma^{\mu}_{\ \nu\lambda}\dot{x}^{\nu}\dot{x}^{\lambda} = \Gamma^{\mu}_{\ 11}(\dot{x}^{1})^{2} + 2\Gamma^{\mu}_{\ 12}\dot{x}^{1}\dot{x}^{2} + \ldots \tag{3.15}$$

(b) For example, once again in the case of the two-sphere, for the θ-equation one has

$$\frac{d}{d\tau}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} = \ddot{\theta} \quad , \quad \frac{\partial\mathcal{L}}{\partial\theta} = \sin\theta\cos\theta\dot{\phi}^{2} \ . \tag{3.16}$$

Comparing the resulting Euler-Lagrange equation

$$\ddot{\theta} - \sin\theta\cos\theta\dot{\phi}^{2} = 0 \tag{3.17}$$

with the geodesic equation

$$\ddot{\theta} + \Gamma^{\theta}_{\ \theta\theta}\dot{\theta}^{2} + 2\Gamma^{\theta}_{\ \theta\phi}\dot{\theta}\dot{\phi} + \Gamma^{\theta}_{\ \phi\phi}\dot{\phi}^{2} = 0 \ , \tag{3.18}$$

one can immediately read off that

$$\Gamma^{\theta}_{\ \phi\phi} = -\sin\theta\cos\theta \quad , \quad \Gamma^{\theta}_{\ \theta\theta} = \Gamma^{\theta}_{\ \theta\phi} = 0 \ . \tag{3.19}$$

Likewise, from

$$\frac{d}{d\tau}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \quad \Leftrightarrow \quad \sin^{2}\theta(\ddot{\phi} + 2\cot\theta\dot{\theta}\dot{\phi}) = 0 \tag{3.20}$$

one deduces that

$$\Gamma^{\phi}_{\ \theta\phi} = \Gamma^{\phi}_{\ \phi\theta} = \cot\theta \quad , \quad \Gamma^{\phi}_{\ \theta\theta} = \Gamma^{\phi}_{\ \phi\phi} = 0 \ . \tag{3.21}$$

(c) As another example, which will turn out to be of considerable importance later on, consider a space-time metric of the form

$$ds^{2} = -A(r)dt^{2} + B(r)dr^{2} + r^{2}d\Omega^{2} \ . \tag{3.22}$$

As will be discussed in section 24.2, this is the general form of a static spherically symmetric metric, and as such will provide us with the starting point for describing the gravitational field of a star. The corresponding Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}\left(-A(r)\dot{t}^{2} + B(r)\dot{r}^{2} + r^{2}(\dot{\theta}^{2} + \sin^{2}\theta\dot{\phi}^{2})\right) \ , \tag{3.23}$$

and therefore the Euler-Lagrange equations for t (a cyclic variable) are

$$0 = \frac{d}{d\tau}(-A(r)\dot{t}) = -A(r)\left(\ddot{t} + \frac{A'(r)}{A(r)}\dot{r}\dot{t}\right) \tag{3.24}$$

(a prime denoting an r-derivative), from which one can immediately read off

$$\Gamma^{t}_{\ rt} = \Gamma^{t}_{\ tr} = \frac{A'}{2A} \quad , \quad \Gamma^{t}_{\ \mu\nu} = 0 \quad \text{otherwise} \ . \tag{3.25}$$

Likewise, the equation for r takes the form

$$\ddot{r} + \frac{B'}{2B}\dot{r}^2 + \frac{A'}{2B}\dot{t}^2 + \cdots = 0 \ , \tag{3.26}$$

and from this one can read off that

$$\Gamma^r{}_{rr} = \frac{B'}{2B} \quad , \quad \Gamma^r{}_{tt} = \frac{A'}{2B} \quad , \quad \ldots \tag{3.27}$$

As we will need them anyway in section 24.3, it is a good exercise to determine all the Christoffel symbols in this way.

3.2 Conserved Charges and (a first encounter with) Killing Vectors

In the previous section we have seen that cyclic coordinates, i.e. coordinates the metric does not depend on, lead to conserved charges, as in (3.2). As nice and useful as this may be (and it is nice and useful), it is obvioulsy somewhat unsatisfactory because it is an explicitly coordinate-dependent statement: the metric may well be independent of one coordinate in some coordinate system, but if one now performs a coordinate transformation which depends on that coordinate, then in the new coordinate system the metric will typically depend on all the new coordinates. Nevertheless,

- the statement that a metric has a certain symmetry (a translational symmetry in the first coordinate system) should be coordinate-independent, and
- thus there should be a corresponding first integral of the geodesic equation in any coordinate system.

To see how this works, let us reconsider the situation discussed in the previous section, namely a metric which in some coordinate system, we will now call it $\{y^\mu\}$, has components $g_{\mu\nu}$ which are independent of y^1, say. Translation invariance of the geodesic Lagrangian is the statement that the Lagrangian is invariant under the infinitesimal variation $\delta y^1 = \epsilon$, $\delta y^\mu = 0$ otherwise, and via Noether's theorem this leads to a conserved charge $g_{1\mu}\dot{y}^\mu$, as in (3.2).

Now we ask ourselves what this statement corresponds to in another coordinate system. Note that in the y-coordinates, invariance is the statement that the metric is invariant under the (infinitesimal) coordinate transformation $y^1 \to y^1 + \epsilon$ or $\delta y^1 = \epsilon$, $\delta y^\mu = 0$ otherwise,

$$\delta g_{\mu\nu} \equiv \epsilon \partial_{y^1} g_{\mu\nu} = 0 \ . \tag{3.28}$$

It is then clear that in another coordinate system, infinitesimal y^1-translations must also correspond to some infinitesimal coordinate transformation (but not necessarily just a translation),

$$\delta x^\alpha = \epsilon V^\alpha(x) \ . \tag{3.29}$$

In particular, if (as in the above example) in y-coordinates V^μ has the components $V^1 = 1$, $V^\mu = 0$ otherwise, then in any other coordinate system one has

$$\delta x^\alpha = (\partial x^\alpha / \partial y^\mu)\delta y^\mu = \epsilon(\partial x^\alpha / \partial y^1) \tag{3.30}$$

so that

$$V^\alpha = J_1^\alpha \tag{3.31}$$

is just the corresponding column of the Jacobi matrix.

In order to determine how to characterise the translational symmetry (3.28) of the metric in an arbitrary coordinate system, we will now proceed in two (as it turns out ultimately equivalent) ways.

1. We can investigate directly, under which conditions on the V^α the transformation (3.29) leads to an invariance of the Lagrangian (2.91). Using

$$\delta \dot{x}^\alpha = \epsilon \dot{x}^\gamma \partial_\gamma V^\alpha \tag{3.32}$$

one straightforwardly finds

$$\delta\left(g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta\right) = \epsilon\left(\delta_V g_{\alpha\beta}\right)\dot{x}^\alpha \dot{x}^\beta \tag{3.33}$$

where

$$\delta_V g_{\alpha\beta} = V^\gamma \partial_\gamma g_{\alpha\beta} + (\partial_\alpha V^\gamma) g_{\gamma\beta} + (\partial_\beta V^\gamma) g_{\alpha\gamma} \tag{3.34}$$

Thus the condition for the infinitesimal transformation (3.29) to leave the Lagrangian invariant is

$$\delta_V g_{\alpha\beta} = 0 \quad . \tag{3.35}$$

Noether's theorem then leads to the corresponding conserved charge

$$Q_V = p_\alpha V^\alpha = g_{\alpha\beta} V^\alpha \dot{x}^\beta \quad . \tag{3.36}$$

Note that for constant components V^α, (3.35) is simply the statement that the metric is constant in the direction V, $V^\gamma \partial_\gamma g_{\alpha\beta} = 0$.

2. Alternatively, we can determine the variation $\delta_V g_{\alpha\beta}$ of the components $g_{\alpha\beta}$ of the metric in x-coordinates from the variation (3.28) of the components $g_{\mu\nu}$ of the metric in y-coordinates by demanding that under a coordinate transformation the variation (3.28) of the metric transforms like the metric. Since we know how the metric transforms (2.5), and we also know how ∂_{y^1} transforms,

$$g_{\mu\nu} = J_\mu^\alpha J_\nu^\beta g_{\alpha\beta} \quad , \quad \partial_{y^1} = (\partial_{y^1} x^\alpha)\partial_\alpha \equiv J_1^\alpha \partial_\alpha \equiv V^\alpha \partial_\alpha \tag{3.37}$$

we find the condition

$$\begin{aligned} & \partial_{y^1} g_{\mu\nu} = J_1^\gamma \partial_\gamma (J_\mu^\alpha J_\nu^\beta g_{\alpha\beta}) \stackrel{!}{=} J_\mu^\alpha J_\nu^\beta \delta_V g_{\alpha\beta} \\ \Leftrightarrow \quad & \delta_V g_{\alpha\beta} = J_\alpha^\mu J_\beta^\nu J_1^\gamma \partial_\gamma (J_\mu^\delta J_\nu^\epsilon g_{\delta\epsilon}) \quad . \end{aligned} \tag{3.38}$$

In order to disentangle this, one can make use of identities such as

$$J_1^\gamma \partial_\gamma J_\mu^\delta = \partial_1 J_\mu^\delta = \partial_\mu J_1^\delta = \partial_\mu V^\delta = J_\mu^\alpha \partial_\alpha V^\delta \tag{3.39}$$

to show that this expression for $\delta_V g_{\alpha\beta}$ is identical to that given in (3.34).

All of this may seem a bit ham-handed at this point, and indeed it is. However, we will see later how these results can be written and understood in a much more pleasing and covariant way. In particular, we will see in section 5.5 how to write (3.34) in a way that makes it completely manifest that it transforms like the metric under coordinate transformations. Moreover, we will discover in section 9 that (3.34) is a special case of the *Lie derivative* of a tensor field along a vector field V, denoted by L_V. Continuous symmetries of a metric correspond to vector fields along which the Lie derivative of the metric vanishes. Such vectors are known as *Killing vectors*, and are thus vectors V^α satisfying the *Killing equation* (3.35),

$$L_V g_{\alpha\beta} \equiv \delta_V g_{\alpha\beta} = 0 \ . \tag{3.40}$$

3.3 Newtonian Limit of the Geodesic Equation

We saw that the 10 components of the metric $g_{\mu\nu}$ appear to play the role of potentials for the gravitational force. In order to substantiate this, and to show that in an appropriate limit this setting is able to reproduce the Newtonian results, we now want to find the relation of these potentials to the Newtonian potential, and the relation between the geodesic equation and the Newtonian equation of motion for a particle moving in a gravitational field.

First let us determine the conditions under which we might expect the general relativistic equation of motion (namely the non-linear coupled set of partial differential geodesic equations) to reduce to the linear equation of motion

$$\frac{d^2}{dt^2}\vec{x} = -\vec{\nabla}\phi \tag{3.41}$$

of Newtonian mechanics, with ϕ the gravitational potential, e.g.

$$\phi = -\frac{G_N M}{r} \ . \tag{3.42}$$

Thus we are trying to characterise the circumstances in which we know and can trust the validity of Newton's equations, such as those provided e.g. by the gravitational field of the earth or the sun, the gravitational fields in which Newton's laws were discovered and tested. Two of these are fairly obvious:

1. Weak Fields: our first plausible assumption is that the gravitational field is in a suitable sense sufficiently weak. We will need to make more precise by what we mean by this, and we will come back to this below.

2. Slow Motion: our second, equally reasonable and plausible, assumption is that the test particle moves at speeds at which we can neglect special relativistic effects, so "slow" should be taken to mean that its velocity is small compared to the velocity of light.

Interestingly, it turns out that one more condition is required. Note that the gravitational fields we have access to are not only quite weak but also only very slowly varying in time, and we will add this condition,

3. Stationary Fields: we will assume that the gravitational field does not vary significantly in time (over the time scale probed by our test particle).

The very fact that we have to add this condition in order to find Newton's equations (as will be borne out by the calculations below) is interesting in its own right, because it also shows that general relativity predicts phenomena deviating from the Newtonian picture even for weak fields, provided that they vary sufficiently rapidly (e.g. quickly oscillating fields), and one such phenomenon is that of gravitational waves (see section 23).

Now, having formulated in words the conditions that we wish to impose, we need to translate these conditions into equations that we can then use in conjunction with the geodesic equation.

1. In order to define a notion of weak fields, we need to keep in mind that this is not a coordinate-independent statement since we can simulate arbitrarily strong gravitational fields even in Minkowski space by going to suitably accelerated coordinates, and therefore a "weak field" condition will be a condition not only on the metric but also on the choice of coordinates. Thus we assume that we can choose coordinates $\{x^\mu\} = \{t, x^i\}$ in such a way that *in these coordinates* the metric differs from the standard constant Minkowski metric $\eta_{\mu\nu}$ only by a small amount,
$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{3.43}$$
where we will implement "by a small amount" in the calculations below by dropping all terms that are at least quadratic in $h_{\mu\nu}$ (and/or its derivatives).

2. The second condition is obviously (with the coordinates chosen above) $dx^i/dt \ll 1$ or, expressed in terms of proper time,
$$\frac{dx^i}{d\tau} \ll \frac{dt}{d\tau} \quad . \tag{3.44}$$

3. The third condition of stationarity we implement simply by considering time-independent fields,
$$g_{\alpha\beta,0} = 0 \quad \Rightarrow \quad h_{\alpha\beta,0} = 0 \tag{3.45}$$

(for a discussion and explanation of the difference betwen the term “stationary” used here and the term “static” used e.g. to describe the metric (3.22), see section 16.4 - it is not crucial here).

Before embarking on the calculation, we note that for the inverse $g^{\alpha\beta}$ of the metric $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, with $h_{\alpha\beta}$ “small”, one evidently has

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \quad \Rightarrow \quad g^{\alpha\beta} = \eta^{\alpha\beta} + \mathcal{O}(h) \ , \tag{3.46}$$

where $\eta^{\alpha\beta}$ is just the inverse Minkowski metric. The explicit expression of the order h term (which we will not need) is given in (3.65) below.

Now we look at the geodesic equation

$$\ddot{x}^\mu + \Gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \ . \tag{3.47}$$

From the decomposition $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ we see that $\Gamma^\mu_{\ \nu\lambda}$ is at least linear in $h_{\mu\nu}$, and by the weak field condition (condition 1) we will only retain the terms linear in $h_{\mu\nu}$. Then the condition of slow motion (condition 2) implies that among the quadratic terms $\dot{x}^\nu\dot{x}^\lambda$ we need to only retain the leading term, namely $\dot{t}\dot{t}$. Thus the geodesic equation can be approximated by

$$\ddot{x}^\mu + \Gamma^\mu_{\ 00}\dot{t}^2 = 0 \ . \tag{3.48}$$

Thus we need to determine

$$\Gamma^\mu_{\ 00} = \tfrac{1}{2}g^{\mu\nu}(g_{\nu 0,0} + g_{\nu 0,0} - g_{00,\nu}) \ . \tag{3.49}$$

Stationarity (condition 3) tells us that the first two terms are zero, and

$$\Gamma^\mu_{\ 00} = -\tfrac{1}{2}g^{\mu\nu}\partial_\nu g_{00} = -\tfrac{1}{2}g^{\mu i}\partial_i g_{00} \ . \tag{3.50}$$

Now $\partial_i g_{00} = \partial_i h_{00}$ is already of order h. Therefore, by the weak field condition, working to linear order in h we can can replace the inverse metric $g^{\mu i}$ by the inverse Minkowski metric $\eta^{\mu i}$, so that in this approximation

$$\Gamma^\mu_{\ 00} = -\tfrac{1}{2}\eta^{\mu i}\partial_i h_{00} \ . \tag{3.51}$$

Thus the relevant Christoffel symbols are

$$\Gamma^0_{\ 00} = 0 \ , \qquad \Gamma^i_{\ 00} = -\tfrac{1}{2}\partial^i h_{00} \ , \tag{3.52}$$

and the geodesic equation splits into

$$\begin{aligned} \ddot{t} &= 0 \\ \ddot{x}^i &= \tfrac{1}{2}\partial_i h_{00}\dot{t}^2 \ . \end{aligned} \tag{3.53}$$

The first of these just says that $\dot{t}$ is constant, or that t is also an affine parameter,

$$t(\tau) = a\tau + b \ . \tag{3.54}$$

In other words, in the Newtonian limit there is essentially (up to a choice of scale/units) no difference between coordinate time and proper time. We can use this in the second equation to convert the τ-derivatives into derivatives with respect to the coordinate time t,

$$\ddot{t} = 0 \quad \Rightarrow \quad \frac{1}{\dot{t}^2}\frac{d^2}{d\tau^2} = \frac{1}{\dot{t}}\frac{d}{d\tau}\frac{1}{\dot{t}}\frac{d}{d\tau} = \frac{d^2}{dt^2} \quad . \tag{3.55}$$

Hence we obtain

$$\frac{d^2x^i}{dt^2} = \tfrac{1}{2}h_{00,i} \tag{3.56}$$

(the spatial index i in this expression is raised or lowered with the Kronecker symbol, $\eta^{ik} = \delta^{ik}$). Comparing this with the Newtonian equation (3.41),

$$\frac{d^2x^i}{dt^2} = -\phi_{,i} \tag{3.57}$$

leads us (with the constant of integration absorbed into an arbitrary constant term in the gravitational potential) to the key identification

$$h_{00} = -2\phi \tag{3.58}$$

between the Newtonian gravitational potential and the (00)-component of the deviation of the space-time metric from the Minkowski metric. By relating this back to $g_{\alpha\beta}$,

$$g_{00} = -(1+2\phi) \quad . \tag{3.59}$$

we find the sought-for relation between the Newtonian potential and the space-time metric. Thus Newtonian gravity can be captured or described by a space-time metric of the form

$$ds^2 = -(1+2\phi(\vec{x}))dt^2 + d\vec{x}^2 \quad . \tag{3.60}$$

For a radial gravitational field, with $\phi = \phi(r)$, it is also natural to write this in terms of spatial spherical coordinates as

$$ds^2 = -(1+2\phi(r))dt^2 + dr^2 + r^2 d\Omega_2^2 \quad . \tag{3.61}$$

<u>Remarks:</u>

1. With the speed of light not set equal to $c = 1$, the dimensionally correct form of this identification is (recall that kinetic and potential energy have the same dimension so that the dimension of ϕ, the energy per unit mass, is that of a velocity-squared; thus ϕ/c^2 is dimensionless)

$$g_{00} = -(1+2\phi/c^2) \quad . \tag{3.62}$$

2. For the gravitational field of isolated systems, it makes sense to choose the integration constant in such a way that the potential goes to zero at infinity, and this choice also ensures that the metric approaches the flat Minkowski metric at infinity.

3. Restoring the appropriate units, in particular the above factor of c^2, one finds that the dimensionless factor $\phi/c^2 \sim 10^{-9}$ on the surface of the earth, 10^{-6} on the surface of the sun (see section 24.4 for some more details), so that the distortion in the space-time geometry produced by gravitation is in general quite small (justifying our approximations).

4. Just for the record, here is the explicit expression for the inverse $g^{\alpha\beta}$ of a metric of the form $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$, with $h_{\alpha\beta}$ "small". In analogy with the series expansion

$$(1+x)^{-1} = 1 - x + x^2 \mp \ldots \tag{3.63}$$

for a real number x with $|x| < 1$, the exact result for $g^{\alpha\beta}$ can be written as an infinite power series in $h_{\alpha\beta}$. We will not need the exact result here, but only the result to linear order in $h_{\alpha\beta}$.

In linear algebra notation, if I is an invertible matrix and A is sufficiently small so that $I + A$ is still invertible, one has (as a matrix generalisation of the power series (3.63) for $(1+x)^{-1}$)

$$(I + A)^{-1} = I^{-1} - I^{-1}AI^{-1} + \mathcal{O}(A^2) \quad . \tag{3.64}$$

In the case at hand (with $I \to \eta, A \to h$), this is

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \quad \Rightarrow \quad g^{\alpha\beta} = \eta^{\alpha\beta} - h^{\alpha\beta} + \mathcal{O}(h^2) \quad , \tag{3.65}$$

where

$$h^{\alpha\beta} = \eta^{\alpha\gamma}\eta^{\beta\delta}h_{\gamma\delta} \quad . \tag{3.66}$$

Indeed it is now easily verified that this satisfies

$$g^{\alpha\beta}g_{\beta\gamma} = \delta^\alpha_\gamma + \mathcal{O}(h^2) \quad . \tag{3.67}$$

5. Within this Newtonian approximation, we cannot distinguish the above result $g_{00} = -(1 + 2\phi)$ from $g_{00} = -(1+\phi)^2$, say (or a host of other possibilities).

6. Likewise, in this approximation it does not make sense to inquire about the other subleading components of the metric. As we have seen, a slowly moving particle in a weak static gravitational field is not sensitive to them, and hence can also not be used to probe or determine these components.

7. In this approximation, the modification of the space-time geometry can equivalently be described as, or attributed to, a space-time dependent speed of light in Minkowski space, along the lines of

$$ds^2 = -c(x)^2dt^2 + d\vec{x}^2 \quad , \tag{3.68}$$

with

$$c^2(x) = (1 + 2\phi(x)/c^2)c^2 \quad . \tag{3.69}$$

Einstein realised fairly early on (1911) in his search for a relativistic theory of gravity that this would have to be part of the story. However, this interpretation is neither useful nor tenable when considering gravitational fields beyond the static Newtonian approximation (which requires one to go beyond a theory with a single scalar potential).

8. Later on, we will determine the exact solution of the Einstein equations (the field equations for the gravitational field, i.e. for the metric) for the gravitational field outside a spherically symmetric mass distribution with mass M (the Schwarzschild metric). The metric turns out to have the simple form (24.31)

$$ds^2 = -\left(1 - \frac{2G_N M}{c^2 r}\right) c^2 dt^2 + \left(1 - \frac{2G_N M}{c^2 r}\right)^{-1} dr^2 + r^2 d\Omega^2 \quad . \tag{3.70}$$

From this expression one can read off that the leading correction to the flat metric indeed arises from the 00-component of the metric,

$$\begin{aligned} ds^2 &\approx -c^2 dt^2 + dr^2 + r^2 d\Omega^2 + \frac{2G_N M}{r} dt^2 + \ldots \\ &= \eta_{\alpha\beta} dx^\alpha dx^\beta + \frac{2G_N M}{rc^2} (dx^0)^2 + \ldots \quad . \end{aligned} \tag{3.71}$$

This is indeed precisely of the above Newtonian form, with the standard Newtonian potential

$$\phi(r) = -\frac{G_N M}{r} \quad . \tag{3.72}$$

One can then also determine the subleading (known as "post-Newtonian") corrections to the general relativistic gravitational field, which are evidently suppressed by additional inverse powers of c^2.

9. The key relation (3.58) can also be obtained at the level of the action. Starting with the action S_0 (the integral of the proper time), and using the time-coordinate t as the parameter, using the same approximations as above one finds that the action can be written as (keeping c explicit for a change, so that $x^0 = ct$)

$$\begin{aligned} S_0[x] &= -mc \int dt \sqrt{-g_{\mu\nu}(dx^\mu/dt)(dx^\nu/dt)} \\ &= -mc \int dt \sqrt{-\eta_{\mu\nu}(dx^\mu/dt)(dx^\nu/dt) - h_{\mu\nu}(dx^\mu/dt)(dx^\nu/dt)} \\ &= -mc \int dt \sqrt{c^2 - \eta_{ik}(dx^i/dt)(dx^k/dt) - h_{00}c^2} \\ &= -mc^2 \int dt \sqrt{1 - \vec{v}^2/c^2 - h_{00}} \quad . \end{aligned} \tag{3.73}$$

Expanding the square root and dropping the first (irrelevant) term, one finds that in this limit the action reduces to

$$S_0[x] \quad \rightarrow \quad \int dt \, \left(\frac{m}{2}\vec{v}^2 + \frac{mc^2}{2} h_{00}\right) \tag{3.74}$$

which is precisely the Newtonian action for a particle in a gravitational potential ϕ,

$$S_N[x] = \int dt \ \left(\frac{m}{2}\vec{v}^2 - m\phi\right) \tag{3.75}$$

provided that one makes the identification (3.58),

$$h_{00} = -2\phi/c^2 \ . \tag{3.76}$$

In this compact (but slightly dubious) derivation of this relation, the significance of the stationarity condition is not manifest: it enters through the condition of the equivalence of the 4-dimensional and 3-dimensional variational principles (with respect to the fields $x^\mu(\tau)$ and $x^k(t)$ respectively), guaranteed by the affine relation between t and τ implied by requiring in addition stationarity.

3.4 Rindler Coordinates Revisited

In section 1.3 we had discussed the Minkowski metric in Rindler coordinates, i.e. in coordinates adapted to a constantly accelerating observer. For an observer accelerating in the x^1-direction, the metric took the form (1.74),

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 + d\vec{y}^2 \ , \tag{3.77}$$

with $\vec{y} = (x^2, x^3)$ denoting the transverse spectator coordinates (which will again be suppressed in the following).

What is the relation, if any, between this metric and the metric describing a weak gravitational field, as derived above (after all, small accelerations should mimic weak gravitational fields)? At first sight, the only thing they appear to have in common is that the departure from what would be the Minkowski-metric in these coordinates is encoded in the time-time component of the metric, ρ^2 in one case, $(1 + 2\phi)$ in the other, but apart from that ρ^2 and $(1 + 2\phi)$ look quite different. This difference is, however, again a coordinate artefact and the Rindler metric can be made to look like the weak-field metric with the help of a suitable further redefinition of the coordinates.

For starters, it will be convenient, for this purpose and for a generalisation which we will discuss below, to introduce the acceleration a explicitly into the coordinates by redefining the coordinate transformation (1.73) to (I will now also call the Minkowski coordinates $\xi^0 = t$ and $\xi^1 = x$)

$$t(\eta, \rho) = (\rho/a) \sinh a\eta \qquad x(\eta, \rho) = (\rho/a) \cosh a\eta \tag{3.78}$$

(so this differs by $\rho \to \rho/a, \eta \to a\eta$ from the transformation given in (1.73)). Thus, now it is the observer at $\rho = 1$ who has acceleration a and whose proper time is $\tau = \eta$. The Rindler metric now has the form

$$ds^2 = -dt^2 + dx^2 = -\rho^2 d\eta^2 + a^{-2} d\rho^2 \ . \tag{3.79}$$

Now the transformation $\rho = 1 + a\hat{x}$ (reminding us that we are talking about acceleration in the $x = x^1$ direction), leads to

$$ds^2 = -(1 + a\hat{x})^2 d\eta^2 + d\hat{x}^2 \tag{3.80}$$

and it is the observer at $\hat{x} = 0$ who has proper time $\tau = \eta$ and constant acceleration a. If one now assumes that the acceleration a is sufficiently small, one can approximate

$$(1 + a\hat{x})^2 \approx 1 + 2a\hat{x} \equiv 1 + 2\phi(\hat{x}) \quad , \tag{3.81}$$

and in $\phi = a\hat{x}$ we recognise precisely the Newtonian potential for a constant force a in the $\hat{x}$-direction. Thus the Rindler and weak field form of the metric agree in this case.

<u>Remarks:</u>

1. Remarkably, this same form of the metric remains valid for an arbitrary time-dependent acceleration $a = a(\tau)$, and thus is capable of reproducing the weak field form of the metric for general potentials. To see this, consider the worldline $(t(\tau), x(\tau))$ with general 4-velocity (actually 2-velocity in this case)

   $$u^0 = \dot{t}(\tau) = \cosh v(\tau) \quad , \quad u^1 = \dot{x}(\tau) = \sinh v(\tau) \quad , \tag{3.82}$$

   which satisfies

   $$(u^0)^2 - (u^1)^2 = 1 \quad , \tag{3.83}$$

   as it should, and has the time-dependent acceleration $a(\tau) = \dot{v}(\tau)$,

   $$(\dot{u}^1)^2 - (\dot{u}^0)^2 = \ddot{x}^2 - \ddot{t}^2 = \dot{v}(\tau)^2 \equiv a(\tau)^2 \quad . \tag{3.84}$$

   We can pass to adapted coordinates $(\eta, \hat{x})$, as above, by setting

   $$(t(\eta, \hat{x}), x(\eta, \hat{x})) = (t(\eta) + \hat{x}\sinh v(\eta), x(\eta) + \hat{x}\cosh v(\eta)) \quad , \tag{3.85}$$

   leading to the metric

   $$ds^2 = -dt^2 + dx^2 = -(1 + a(\eta)\hat{x})^2 d\eta^2 + d\hat{x}^2 \quad . \tag{3.86}$$

   This is indeed the direct generalisation of (3.80) to arbitrary accelerations, the original worldline manifestly corresponding to the observer at fixed $\hat{x} = 0$, with $\eta = \tau$, and the same remarks regarding the weak field limit / small accelerations apply.

2. Another useful alternative coordinate transformation for constant a is $\rho = \exp a\xi$, leading to

   $$ds^2 = e^{2a\xi}(-d\eta^2 + d\xi^2) \quad . \tag{3.87}$$

   Note that in these coordinates the metric is *conformally flat*, i.e. it differs from the flat Minkowski metric $-d\eta^2 + d\xi^2$ in these coordinates only by an overall factor.

Moreover, for small values of a the time-component of the metric again reduces to something like the Newtonian limit expression (3.81), namely

$$\mathrm{e}^{2a\xi} \approx 1 + 2a\xi \quad . \tag{3.88}$$

For the record, and for later use, we note that the complete coordinate transformation between the Minkowski coordinates (t, x) and the conformally flat Rindler coordinates (η, ξ) is

$$(t, x) = (a^{-1}\mathrm{e}^{a\xi} \sinh a\eta, a^{-1}\mathrm{e}^{a\xi} \cosh a\eta) \quad . \tag{3.89}$$

The generator of Rindler time evolution in these coordinates is

$$\partial_\eta = a(t\partial_x + x\partial_t) \quad . \tag{3.90}$$

This is a boost in the (t, x)-plane, but the limit $a \to 0$ appears to be singular.

3. A simple and useful way to rectify this is to introduce a further constant shift of x, $x \to x - 1/a$, into the 1-parameter family (3.89) of coordinate transformations,

$$(t, x) = (a^{-1}\mathrm{e}^{a\xi} \sinh a\eta, a^{-1}\mathrm{e}^{a\xi} \cosh a\eta - a^{-1}) \quad . \tag{3.91}$$

This transformation now has the desirable property that as $a \to 0$ it continuously connects the Rindler and Minkowski coordinates,

$$a \to 0 \quad \Rightarrow \quad t \to \eta \quad , \quad x \to \xi \quad . \tag{3.92}$$

As a consequence, also the Rindler time evolution generator

$$\partial_\eta = \partial_t + a(t\partial_x + x\partial_t) \tag{3.93}$$

now has a non-singular limit as $a \to 0$, namely the Minkowski time generator ∂_t.

4. In terms of the Minkowski null (or advanced and retarded time) coordinates

$$u_M = t - x \quad , \quad v_M = t + x \quad \Rightarrow \quad ds^2 = -du_M\, dv_M \quad , \tag{3.94}$$

and their Rindler counterparts

$$u_R = \eta - \xi \quad , \quad v_R = \eta + \xi \quad \Rightarrow \quad ds^2 = -\mathrm{e}^{a(v_R - u_R)} du_R\, dv_R \quad , \tag{3.95}$$

the transformation (3.89) takes the form

$$t \mp x = a^{-1}\mathrm{e}^{a\xi} (\sinh a\eta \mp \cosh a\eta) = \mp a^{-1}\mathrm{e}^{a(\xi \mp \eta)} \tag{3.96}$$

or, compactly,

$$u_M = -a^{-1}\mathrm{e}^{-au_R} \quad , \quad v_M = +a^{-1}\mathrm{e}^{+av_R} \quad . \tag{3.97}$$

Note that the range of the coordinates is $-\infty < \eta, \xi < +\infty$ or $-\infty < u_R, v_R < +\infty$, and that the coordinates (η, ξ) or (u_R, v_R) cover (and can be used in) the right-hand quadrant $x > |t|$ of Minkowski space-time, corresponding to $-\infty < u_M = t - x < 0$ and $0 < v_M = t + x < +\infty$, the so called (right) *Rindler wedge.*

As we will see in section 7.8, these null Rindler coordinates are particularly useful for studying the solutions of the scalar wave equation in the Rindler wedge.

Let me close this section with some comments on other versions of $(3+1)$-dimensional Rindler space. First of all, instead of looking at acceleration in the x^1-direction, say, one can consider radial accelerations. To that end one first writes the metric in spatial spherical coordinates,

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega_2^2 \quad , \tag{3.98}$$

and introduces Rindler coordinates (ρ, τ), say, via

$$(t, r) = (\rho \sinh \tau, \rho \cosh \tau) \quad , \tag{3.99}$$

leading to the $(3+1)$-dimensional spherical Rindler metric

$$ds^2 = -\rho^2 d\tau^2 + d\rho^2 + \rho^2 \cosh^2 \tau \, d\Omega_2^2 \quad . \tag{3.100}$$

This form of the metric is adapted to the hyperboloids

$$r^2 - t^2 = \rho^2 \tag{3.101}$$

i.e. to a family of constantly radially accelerating observers whose worldlines asymptotically approach the lightcone through the origin $(t = 0, r = 0)$. These coordinates cover precisely the region of Minkowski space outside this lightcone, i.e. the region of events at spacelike distance from the origin while the region that is not covered is the past and future of the origin.

The "complementary" metric (adapted to the hyperboloids with $r^2 - t^2 < 0$ and covering precisely the interior of the lightcone) is the so-called *Milne metric* to be discussed in section 37.1.

A non-trivial variant of this metric[6] is obtained by shifting $r \to r - r_0$ in the above coordinate transformation,

$$(t, r) = (\rho \sinh \tau, r_0 + \rho \cosh \tau) \quad , \tag{3.102}$$

leading to the metric

$$ds^2 = -\rho^2 d\tau^2 + d\rho^2 + (r_0 + \rho \cosh \tau)^2 d\Omega_2^2 \tag{3.103}$$

[6] V. Balasubramanian, B. Czech, B. Chowdhury, J. de Boer, *The entropy of a hole in spacetime*, `arXiv:1305.0856 [hep-th]`, V. Balasubramanian, B. Chowdhury, B. Czech, J. de Boer, M. Heller, *A hole-ographic spacetime*, `arXiv:1310.4204 [hep-th]`.

(an analogous shift $x \to x - x_0$ for acceleration in the x-direction would have had no effect on the metric since such a translation is a symmetry of the Minkowski metric, whereas a translation in the radial direction is not). This form of the metric is adapted to the hyperboloids

$$(r - r_0)^2 - t^2 = \rho^2 \tag{3.104}$$

and now describes radially accelerating observers, each one asymptotically approaching the radial lightray emanating from a distance r_0 from the origin (and correspondingly the region of space-time covered by these coordinates is the complement of the past and future of the 2-sphere of radius r_0 at the origin, a "hole" in space-time).

3.5 Gravitational Redshift

Following Einstein, the gravitational redshift (i.e. the fact that photons appear to lose or gain energy when rising or falling in a gravitational field) is usually presented as a direct consequence of the Einstein Equivalence Principle (and is therefore also said to provide an experimental test of the Einstein Equivalence Principle itself). It (or, better, its Newtonian weak field limit) can indeed be derived in this way (see Remark 1 at the end of this section for one such argument, albeit not the original one). However, here we will derive this effect without any approximation within the framework that we have already adopted, inspired by the equivalence principle, namely in terms of the description of the gravitational field by a metric.

This has several advantages. It allows us to further familiarise ourselves with the formalism and to illustrate how to extract physical effects from our description of lightrays as null geodesics (much as we employed timelike geodesics above to study the Newtonian limit). Moreover, it allows us to derive formulae for this effect in quite some generality and I will actually give 3 different derivations in increasing order of generality. In conjunction with the Newtonian approximation to the gravitational field these then reduce to the result in the form in which it is usually presented, e.g. as in (3.130) or (3.131) (and as then rederived on the basis of the equivalence principle in (3.137)).

To set the stage, note that it is manifest from the expression

$$d\tau^2 = -g_{\alpha\beta}(x)dx^\alpha dx^\beta \tag{3.105}$$

for the proper time that e.g. the rate of clocks is affected by where one is in a gravitational field. However, as by the universality of gravity everything is (and in particular all ideal clocks are) affected in the same way by gravity, it is impossible to measure this effect locally, at a fixed point in a gravitational field. In order to find an observable effect, one needs to compare data from two different points in a gravitational potential.

The situation we could consider is that of two observers A and B moving on worldlines (paths) γ_A and γ_B, A sending light signals to B. In general the frequency, measured in the observers rest-frame at A (or in a locally inertial coordinate system there) will differ from the frequency measured by B upon receiving the signal.

In order to separate out Doppler-like effects due to relative velocities, we consider two observers A and B at rest radially to each other, at radii r_A and r_B, in a static spherically symmetric gravitational field. This means that the metric depends only on a radial coordinate r and we can choose it to be of the form

$$ds^2 = g_{00}(r)dt^2 + g_{rr}(r)dr^2 + r^2 d\Omega^2 \quad , \tag{3.106}$$

where $d\Omega^2$ is the standard volume element on the two-sphere (see section 24 for a more detailed justification of this ansatz for the metric).

Observer A sends out light of a given frequency ν_A, say n pulses per proper time unit $\Delta\tau_A$. Observer B receives these n pulses in his proper time $\Delta\tau_B$ and interprets this as a frequency ν_B. Thus the relation between the frequency ν_A emitted at A and the frequency ν_B observed at B is

$$\frac{\nu_A}{\nu_B} = \frac{\Delta\tau_B}{\Delta\tau_A} \quad . \tag{3.107}$$

I will now give two arguments to show that this ratio depends on the metric (i.e. the gravitational field) at r_A and r_B through

$$\frac{\nu_A}{\nu_B} = \frac{(-g_{00}(r_B))^{1/2}}{(-g_{00}(r_A))^{1/2}} \quad , \tag{3.108}$$

and then a 3rd argument establishing a slightly more general result.

1. The first argument is essentially one based on geometric optics (and is best accompanied by drawing a (1+1)-dimensional space-time diagram of the lightrays and worldlines of the observers).

   The geometry of the situation dictates that the coordinate time intervals recorded at A and B are equal, $\Delta t_A = \Delta t_B$ as nothing in the metric actually depends on t. In equations, this can be seen as follows. First of all, the equation for a radial lightray is

   $$-g_{00}(r)dt^2 = g_{rr}(r)dr^2 \quad , \tag{3.109}$$

   or

   $$\frac{dt}{dr} = \pm\left(\frac{g_{rr}(r)}{-g_{00}(r)}\right)^{1/2} \quad . \tag{3.110}$$

   From this we can calculate the coordinate time for the lightray to go from A to B. Say that the first light pulse is emitted at point A at time $t(A)_1$ and received at B at coordinate time $t(B)_1$. Then

   $$t(B)_1 - t(A)_1 = \int_{r_A}^{r_B} dr(-g_{rr}(r)/g_{00}(r))^{1/2} \tag{3.111}$$

The right hand side obviously does not depend on t, so we also have

$$t(B)_2 - t(A)_2 = \int_{r_A}^{r_B} dr(-g_{rr}(r)/g_{00}(r))^{1/2} \tag{3.112}$$

where t_2 denotes the coordinate time for the arrival of the n-th pulse. Therefore,

$$t(B)_1 - t(A)_1 = t(B)_2 - t(A)_2 \quad , \tag{3.113}$$

or

$$t(A)_2 - t(A)_1 = t(B)_2 - t(B)_1 \quad , \tag{3.114}$$

as claimed. Thus the coordinate time intervals recorded at A and B between the first and last pulse are equal. However, to convert this to proper time, we have to multiply the coordinate time intervals by an r-dependent function,

$$\Delta\tau_{A,B} = (-g_{\alpha\beta}(r_{A,B})\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt})^{1/2}\Delta t_{A,B} \quad , \tag{3.115}$$

and therefore the proper time intervals will not be equal. For observers at rest, $dx^i/dt = 0$, one has

$$\Delta\tau_{A,B} = (-g_{00}(r_{A,B}))^{1/2}\Delta t_{A,B} \quad . \tag{3.116}$$

Since $\Delta t_A = \Delta t_B$, (3.108) now follows from (3.107).

2. The second argument uses the null geodesic equation, in particular the conserved quantity associated to time-translations (recall that we have assumed that the metric (3.106) is time-independent), as well as a somewhat more covariant looking, but equivalent, notion of frequency.

   First of all, let the lightray be described by the wave vector k^α. In special relativity, we would parametrise this as $k^\alpha = (\omega, \vec{k})$ with $\omega = 2\pi\nu$ the frequency. This is the frequency observed by an inertial observer at rest, with 4-velocity $u^\alpha = (1, 0, 0, 0)$. A Lorentz-invariant, and in our context now coordinate-independent, notion of the frequency as measured by an observer with velocity u^α is thus

$$\omega = -u^\alpha k_\alpha \quad . \tag{3.117}$$

   This includes as special cases

   - the standard (special) relativistic Doppler effect (where one compares ω with $\bar{\omega} = -\bar{u}^\alpha k_\alpha$, $\bar{u}^\alpha$ the 4-velocity of a boosted observer),
   - and the gravitational redshift between static observers we want to discuss here,

   but more generally also the redshift for observers with arbitrary 4-velocity u^α. And indeed we will employ this method in section 26.5 to look at

- the redshift between a static and a freely falling observer in the Schwarzschild geometry.

Returning to the case at hand, a static observer in the spherically-symmetric and static gravitational field (3.106) is described by the 4-velocity

$$u^\alpha = (u^0, 0, 0, 0) \qquad g_{\alpha\beta}u^\alpha u^\beta = g_{00}(u^0)^2 = -1 \ . \tag{3.118}$$

Thus for the static observer at $r = r_A$, say, one has

$$u^0_A = (-g_{00}(r_A))^{-1/2} \tag{3.119}$$

(and likewise for the observer at $r = r_B$). The wave vector k^μ is a null tangent vector, $k^\mu k_\mu = 0$, to a null geodesic corresponding to the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = \tfrac{1}{2}g_{00}(r)\dot{t}^2 + \ldots \tag{3.120}$$

Since the metric is time-independent, there is (cf. the discussion in section 3.1) the corresponding conserved quantity

$$E = -\frac{\partial\mathcal{L}}{\partial\dot{t}} = -g_{00}(r)\dot{t} \tag{3.121}$$

(the minus sign serving only to make this quantity positive for $\dot{t} > 0$). Then one finds that the frequency measured by the static observer at $r = r_A$ is

$$\begin{aligned}\omega_A = -u^\alpha_A k_\alpha &= -\left(-g_{00}(r_A)\right)^{-1/2} k_0 = -\left(-g_{00}(r_A)\right)^{-1/2} g_{0\alpha}(r_A)\dot{x}^\alpha \\ &= -\left(-g_{00}(r_A)\right)^{-1/2} g_{00}(r_A)\dot{t} = E\left(-g_{00}(r_A)\right)^{-1/2}\end{aligned} \tag{3.122}$$

Since E is a conserved quantity, i.e. the same for the lightray at $r = r_A$ or $r = r_B$, one sees that $\omega_A/\omega_B = \nu_A/\nu_B$ is given by (3.108), as claimed.

Note that this derivation shows that the relation between ω and E is exactly like the relation (3.116) between $(\Delta\tau)^{-1}$ and $(\Delta t)^{-1}$, which provides us with an interpretation of the conserved quantity E for a massless particle / photon: it is the frequency measured with respect to coordinate time (as the momentum conjugate to the time-coordinate t this should not be too surprising).

3. Even with the restriction to static observers in static gravitational fields, the above derivation is not completely general, and still not completely covariant, because we used the explicit form of the metric (which is the general form of a metric with a time-translation invariance in spherical symmetry, but not in general). We can improve this somewhat by using the more general characterisation of time-translation invariance in terms of Killing vectors (section 3.2) and the associated conserved charge (3.36).

   Thus assume that we have a timelike Killing vector V^α. Then by definition a static observer is one whose 4-velocity u^α is proportional to V^α,

$$u^\alpha \sim V^\alpha \ . \tag{3.123}$$

For $V = \partial_t$ this evidently reduces to the statement that only t changes along the worldline, i.e. that the observer remains at fixed values of the spatial coordinates, and this is the sense in which we have informally used the term "static observer" so far. Denoting the norm of V by

$$\mathbb{V} = (-V^\alpha V_\alpha)^{1/2} \quad , \tag{3.124}$$

the normalisation condition $u^\alpha u_\alpha = -1$ fixes the proportionality factor between u^α and V^α to be

$$u^\alpha = V^\alpha / \mathbb{V} \quad . \tag{3.125}$$

Given the null wave vector k^α, we have the conserved energy (3.36),

$$E = -k_\alpha V^\alpha \quad . \tag{3.126}$$

Therefore, adopting the definition (3.117), the frequency observed by a static observer is

$$\omega = -u^\alpha k_\alpha = -k_\alpha V^\alpha / \mathbb{V} = E/\mathbb{V} \quad . \tag{3.127}$$

Since E is constant along the lightray, frequencies observed by two different static observers are related by

$$\frac{\omega_A}{\omega_B} = \frac{\mathbb{V}_B}{\mathbb{V}_A} \quad . \tag{3.128}$$

For this reason, the norm $\mathbb{V}$ is also known as the *redshift factor* associated with a timelike Killing vector.

Note that this result reduces to (3.108) if the metric has the form (3.106) and $V = \partial_t$ since then

$$V = \partial_t \quad \Rightarrow \quad \mathbb{V}(r) = (-g_{00}(r))^{1/2} \quad . \tag{3.129}$$

Having derived (3.108) in 3 different ways, let us now look at what the result tells us in specific situations of interest. Since on earth and in the solar system we only have access to gravitational fields that are to a reasonably high degree of precision well described by Newtonian gravity, we can use the Newtonian approximation (3.59). Then (3.108) becomes

$$g_{00} = -(1 + 2\phi) \quad \Rightarrow \quad \frac{\nu_A}{\nu_B} \sim 1 + \phi(r_B) - \phi(r_A) \quad , \tag{3.130}$$

or, with $\phi(r) = -G_N M/r$,

$$\frac{\nu_A - \nu_B}{\nu_B} = \frac{G_N M (r_B - r_A)}{r_A r_B} \tag{3.131}$$

Thus for $r_B > r_A$ one has

$$r_B > r_A \quad \Rightarrow \quad \nu_B < \nu_A \quad , \tag{3.132}$$

so that the frequency of a lightray appears to be redshifted when rising (moving upwards) in a gravitational field, and one has a gravitational blueshift effect

$$r_B < r_A \quad \Rightarrow \quad \nu_B > \nu_A \quad . \tag{3.133}$$

for photons falling (moving downwards) in a gravitational field.

So far these are the objective facts, just the results of the equation (3.108) that we have derived. Let me now turn to the more subtle question of how to interpret this result, or what words to use to best describe this result.

It is tempting (and unfortunately quite common) to interpret these results as saying that "as expected, a photon loses energy when rising in (and against the pull of) a gravitational field". This kind of non-relativistic Newtonian reasoning, wildly extrapolated to include the energy of lightrays, can to a certain extent be justified by appealing to the Einstein Equivalence Principle (see Remark 1 below), but I believe that this is misleading and provides the wrong intuition for the physics of the situation, and I will subsequently provide a different explanation that (to me) appears to be much more to the point (Remark 4).

<u>Remarks:</u>

1. The Newtonian limit (3.130) of the exact result (3.108) can to some extent also be deduced from energy conservation applied to Newtonian gravity.[7] By the Einstein Equivalence Principle a local observer at the emitter A will see a change in the internal mass of the emitter $\Delta m_A = -h\nu_A$ when a photon of frequency of ν_A is emitted. Likewise, the absorber at point B will experience an increase in inertial mass by $\Delta m_B = h\nu_B$, but the total internal plus gravitational potential energy

   $$mc^2 + m\phi \equiv m + m\phi = m(1+\phi) \tag{3.134}$$

   must be conserved, i.e.

   $$m_A(1+\phi(r_A))+m_B(1+\phi(r_B)) = (m_A+\Delta m_A)(1+\phi(r_A))+(m_B+\Delta m_B)(1+\phi(r_B)) \quad . \tag{3.135}$$

   Thus

   $$0 = \Delta m_A(1+\phi(r_A)) + \Delta m_B(1+\phi(r_B)) \quad , \tag{3.136}$$

   leading to (3.130),

   $$\frac{\nu_A}{\nu_B} = \frac{1+\phi(r_B)}{1+\phi(r_A)} \sim 1 + \phi(r_B) - \phi(r_A) \quad . \tag{3.137}$$

2. This "derivation" (in quotes, because we are wildly mixing Newtonian gravity, special relativity and quantum mechanics - do take this "derivation" with an

[7] See S. Weinberg, *Gravitation and Cosmology* section 3.5 for a careful discussion.

appropriately sized grain of salt, please) shows that gravitational redshift experiments test the Einstein Equivalence Principle in its strong form, in which the term 'laws of nature' is not restricted to mechanics (inertial = gravitational mass), but also includes quantum mechanics in the sense that it tests if in an inertial frame the relation between photon energy and frequency is unaffected by the presence of a gravitational field.

3. One significant shortcoming of the above "explanation" is that it relies on the Newtonian concept of a gravitational potential energy, and that it therefore applies at best to the Newtonian limit (3.130) of the exact result (3.108). It also seems to suggest that something is happening to the photon / lightray as it travels from A to B (namely that it "loses energy"), and it is in this sense that the Newtonian intuition is misleading.

   Indeed, note that the simple general result (3.108) depends only on the value of the gravitational field at the points r_A and r_B, not on the gravitational field inbetween. This is at it should be. After all, this gravitational time delay between clocks at r_A and r_B is all that we have used is the derivation (and not some new law of physics that causes photons to lose energy in some unexplained way). This strongly suggest the interpretation, that the gravitational redshift is only due to the different rate of clocks / proper time at the positions r_A and r_B, and not due to the fact that "something happens to the lightray as it travels through a gravitational field" (which should lead to a cumulative effect depending also on the intermediate gravitational field).

4. Indeed, let us reanalyse the gravitational redshift (without any approximation) from this point of view. We note that the observer A at r_A uses his (or her) proper time τ_A to define a notion of frequency ν_A, and hence a notion of energy $E_A = h\nu_A$. Likewise the observer B defines an energy $E_B = h\nu_B$ based on his (or her) proper time τ_B. The fact that $E_B \neq E_A$ has nothing to do with the photon or lightray having lost or gained energy, it is simply due to the fact that the observers A and B define energy in a different way!

   Moreover, there is a rigorous and well-defined sense in which the energy of a photon moving in a time-independent gravitational field is conserved. Indeed, recall that in such a case there is a conserved quantity $p_0 = -E$ (3.7). In the case at hand it takes the form (3.121),

$$E = -g_{00}(r)\dot{t} \tag{3.138}$$

   and (as the momentum conjugate to the time coordinate t) has the interpretation as (Planck's constant times) the frequency ν_t of the lightray with respect to this (observer-independent) coordinate time t,

$$E = E_t = h\nu_t \quad . \tag{3.139}$$

Putting all these facts together, one sees that the gravitational redshift can and should be attributed entirely to the gravitational time-delay between observers A and B, and thus to the different conventions for measuring frequencies and energies, and not to some kind of "loss of energy of a photon in a gravitational field".

Let me close this section with a brief comment on the observational tests of the redshift prediction (3.131) (but you should consult other sources for a more detailed discussion of experimental tests of General Relativity). While difficult to observe directly (by looking at light from the sun), the prediction (3.131) has been verified in the laboratory, first by Pound and Rebka (1960), and subsequently, with one percent accuracy, by Pound and Snider in 1964.

Here is a rough estimate of the expected effect. We first consider light reaching us (B) from the sun (A). In this case, we have $r_B \gg r_A$, where r_A is the radius of the sun, and (also inserting a so far suppressed factor of c^2) we obtain

$$\frac{\nu_A - \nu_B}{\nu_B} = \frac{G_N M(r_B - r_A)}{c^2 r_A r_B} \simeq \frac{G_N M}{c^2 r_A} \quad . \tag{3.140}$$

Using the approximate values

$$\begin{aligned} r_A &\simeq 0.7 \times 10^6 \ \mathsf{km} \\ M_{\text{sun}} &\simeq 2 \times 10^{33} \ \mathsf{g} \\ G_N &\simeq 7 \times 10^{-8} \ \mathsf{g}^{-1}\mathsf{cm}^3\mathsf{s}^{-2} \\ G_N c^{-2} &\simeq 7 \times 10^{-29} \ \mathsf{g}^{-1}\mathsf{cm} = 7 \times 10^{-34} \ \mathsf{g}^{-1}\mathsf{km} \ \ , \end{aligned} \tag{3.141}$$

one finds

$$\frac{\Delta\nu}{\nu} \simeq 2 \times 10^{-6} \quad . \tag{3.142}$$

In principle, such a frequency shift should be observable. In practice, however, the spectral lines of light emitted by the sun are strongly effected e.g. by convection in the atmosphere of the sun (Doppler effect), and this makes it difficult to measure this effect with the required precision.

In the Pound-Snider experiment, the actual value of $\Delta\nu/\nu$ is much smaller. In the original set-up one has $r_B - r_A \simeq 20\mathsf{m}$ (the distance from floor to ceiling of the laboratory), and $r_A = r_{\text{earth}} \simeq 6.4 \times 10^6\mathsf{m}$, leading to

$$\frac{\Delta\nu}{\nu} \simeq 2.5 \times 10^{-15} \quad . \tag{3.143}$$

However, here the experiment is much better controlled, very sharp spectral lines can be obtained by using the Mössbauer effect, and the gravitational redshift was verified with 1% accuracy. In the meantime, experiments with a much higher accuracy have been performed.

3.6 Equivalence Principle Revisited: Existence of Locally Inertial Coordinates

Central to our initial discussion of gravity was the Einstein Equivalence Principle which postulates the existence of locally inertial (or freely falling) coordinate systems in which locally at (or around) a point the effects of gravity are absent. Now that we have decided that the arena of gravity is a general metric space-time, we should establish that such coordinate systems indeed exist. Looking at the geodesic equation, it is clear that at least in this context "absence of gravitational effects" is tantamount to the existence of a coordinate system $\{\xi^a\}$ in which at a given point p the metric is the Minkowski metric, $g_{ab}(p) = \eta_{ab}$ and the Christoffel symbols are zero, $\Gamma^a_{\ bc}(p) = 0$,

$$g_{ab}(p) = \eta_{ab} \quad , \quad \Gamma^a_{\ bc}(p) = 0 \ . \tag{3.144}$$

Owing to the identity

$$g_{\mu\nu},_\lambda = \Gamma_{\mu\nu\lambda} + \Gamma_{\nu\mu\lambda} \ , \tag{3.145}$$

the latter condition is equivalent to $g_{ab,\,c}(p) = 0$. Below (after the Remarks) I will sketch three arguments establishing the existence of such coordinate systems, each one having its own virtues and providing its own insights into the issue.

Remarks:

1. Actually it is physically plausible (and fortuitously moreover true) that one can always find coordinates which embody the equivalence principle in the stronger sense that the metric is the flat metric η_{ab} and the Christoffel symbols are zero not just at a point but along the entire worldline of an inertial (freely falling) observer, i.e. along a geodesic γ,

$$g_{ab}|_\gamma = \eta_{ab} \quad , \quad \Gamma^a_{\ bc}|_\gamma = 0 \ . \tag{3.146}$$

   Such coordinates, based on a geodesic rather than on a point, are known as Fermi normal coordinates. The construction is similar to that of Riemann normal coordinates (based at a point) to be discussed below.[8]

2. In this mathematically idealised realisation of the equivalence principle, nothing is said about the metric and the Christoffel symbols in a neighbourhood of that point (or of the geodesic), and nothing is said about the 2nd and higher derivatives of the metric at that point.

[8] Most discussions of Fermi coordinates in the literature follow the presentation given in F. Menasse, C. Misner, *Fermi normal coordinates and some basic concepts in differential geometry*, J. Math. Phys. 4 (1963) 735-745; for a geometrically transparent treatment see also section 1.11 of E. Poisson, *A Relativist's Toolkit*; Fermi coordinates for *null* geodesics are constructed in M. Blau, D. Frank, S. Weiss, *Fermi Coordinates and Penrose Limits*, `arXiv:hep-th/0603109`.

3. In particular, thinking of the 1st derivatives of the metric as encoding the gravitational force, the 2nd derivatives of the metric must then correspond to gravitational *tidal forces*. (We will see this in more detail in section 8.4, where these tidal forces are related to components of the Riemann curvature tensor, a tensor that involves up to 2nd derivatives of the metric.) Such tidal forces are objective, i.e. physically real, as they lead to stresses in (or deformations of) extended bodies. One can and should therefore not expect to be able to eliminate such tidal forces by a suitable choice of reference system.

4. Therefore, physically what the equivalence principle says (or should say) is that in a gravitational field there is locally a reference system in which the effects of gravity are absent, provided that you choose the spacetime region to be sufficently small so that you can neglect the effect of gravitational tidal forces.

Here is a sketch of 3 arguments establishing the existence of locally inertial coordinate systems:

1. Direct Construction

   We know that given a coordinate system $\{\xi^a\}$ that is inertial at a point p, the metric and Christoffel symbols at p in a new coordinate system $\{x^\mu\}$ are determined by (1.87,1.98). Conversely, we will now see that knowledge of the metric and Christoffel symbols at a point p is sufficient to construct a locally inertial coordinate system at p.

   We will construct this coordinate system $\xi^a = \xi^a(x)$ locally around the point p (with coordinates x_0^α, say, in the original coordinate system) by a Taylor series expansion,

   $$\xi^a(x) = d^a + (x - x_0)^\alpha e^a_\alpha + \tfrac{1}{2}(x - x_0)^\beta (x - x_0)^\gamma f^a_{\beta\gamma} + \ldots \quad . \tag{3.147}$$

   Here

   $$d^a = \xi^a(x_0) \equiv \xi^a_0 \tag{3.148}$$

   are the (arbitrary) coordinate values of the point p in the new coordinates ξ^a,

   $$e^a_\alpha = \frac{\partial \xi^a}{\partial x^\alpha}(x_0) \tag{3.149}$$

   is the Jacobi matrix of the coordinate transformation at $x = x_0$, and

   $$f^a_{\beta\gamma} = \frac{\partial^2 \xi^a}{\partial x^\beta \partial x^\gamma}(x_0) \tag{3.150}$$

   is its 1st derivative at x_0.

   Form the tensorial transformation behaviour of the metric we know that

   $$g_{\alpha\beta}(x_0) = g_{ab}(\xi_0) e^a_\alpha e^b_\beta \quad . \tag{3.151}$$

Requiring that $g_{ab}(\xi_0)=\eta_{ab}$ leads to the condition

$$g_{ab}(\xi_0)=\eta_{ab} \quad \Rightarrow \quad g_{\alpha\beta}(x_0)=\eta_{ab}e^a_\alpha e^b_\beta \quad . \tag{3.152}$$

This shows that the e^a_α, thought of as a matrix, are an invertible (4×4)-matrix in $\mathsf{GL}(4,\mathbb{R})$. Denoting its inverse by e^α_a, with

$$e^\alpha_a e^a_\beta=\delta^\alpha_\beta \quad , \quad e^\alpha_a e^b_\alpha=\delta^a_b \quad , \tag{3.153}$$

we see that the inverse matrix diagonalises (and scales) the metric at the point p in such a way that

$$g_{\alpha\beta}(x_0)e^\alpha_a e^\beta_b=\eta_{ab} \quad . \tag{3.154}$$

Since $g_{\alpha\beta}(x_0)$ is a symmetric non-degenerate matrix, such matrices always exist (and are unique up to similarity transformations that leave η_{ab} invariant, i.e. up to Lorentz transformations). The notation e^a_α and e^α_a reflects the fact that these matrices are the components of an orthonormal vierbein (or vielbein) at the point p, which are traditionally denoted this way (cf. the discussion in section 4.8 below).

Taking stock, we see that the condition $g_{ab}(p)=\eta_{ab}$ determines the coordinate system to 1st order in a Taylor series expansion, up to translations (the choice of d^a) and Lorentz transformations, i.e. up to Poincaré transformation.

We now turn to the 2nd condition characterising a locally inertial coordinate system, namely $\Gamma^a_{\ bc}(p)=0$. We can write the inhomogeneous transformation behaviour of the Christoffel symbols as

$$\frac{\partial\xi^a}{\partial x^\alpha}\Gamma^\alpha_{\ \beta\gamma}=\Gamma^a_{\ bc}\frac{\partial\xi^b}{\partial x^\beta}\frac{\partial\xi^c}{\partial x^\gamma}+\frac{\partial^2\xi^a}{\partial x^\beta\partial x^\gamma} \quad . \tag{3.155}$$

Thus at the point p we have

$$e^a_\alpha\Gamma^\alpha_{\ \beta\gamma}(x_0)=\Gamma^a_{\ bc}(\xi_0)e^b_\beta e^c_\gamma+f^a_{\beta\gamma} \quad . \tag{3.156}$$

Requiring $\Gamma^a_{\ bc}(p)=0$ now uniquely determines the 2nd order Taylor coefficients,

$$\Gamma^a_{\ bc}(\xi_0)=0 \quad \Rightarrow \quad f^a_{\beta\gamma}=e^a_\alpha\Gamma^\alpha_{\ \beta\gamma}(x_0) \quad . \tag{3.157}$$

Thus to 2nd order in a Taylor series expansion, the transformation from arbitrary coordinates x^α to inertial coordinates ξ^a at the point p is given by

$$\xi^a(x)=\xi^a_0+(x-x_0)^\alpha e^a_\alpha+\tfrac{1}{2}(x-x_0)^\beta(x-x_0)^\gamma e^a_\alpha\Gamma^\alpha_{\ \beta\gamma}(x_0)+\ldots \quad . \tag{3.158}$$

To this order we can also write the inverse coordinate transformation as

$$x^\alpha(\xi)=x^\alpha_0+(\xi-\xi_0)^a e^\alpha_a-\tfrac{1}{2}(\xi-\xi_0)^b(\xi-\xi_0)^c\Gamma^\alpha_{\ \beta\gamma}(x_0)e^\beta_b e^\gamma_c+\ldots \quad . \tag{3.159}$$

We have therefore established that for an arbitrary point p in an arbitrary gravitational field one can always introduce local coordinates which are inertial at that point, and that up to 2nd order in a Taylor series expansion such a coordinate system is unique up to Poincaré transformations.

Since this leaves the infinite number of higher-order terms of the Taylor expansion undetermined, this shows that inertial coordinate systems are highly non-unique, and raises the following questions:

- Can one continue in this vein and choose the (so far undetermined) higher-order terms in the Taylor expansion such that also e.g. the 2nd derivatives of the metric at p are equal to zero,

  $$\exists ? \quad \xi^a(x): \quad g_{ab,\,cd}(p) = 0 \ \ ? \tag{3.160}$$

  The answer to this is a resounding "no", as the 3rd (numerological) argument below will show. In fact, as we will see (and study in detail) later on, the 2nd derivatives of the metric contain important coordinate-invariant information about the curvature of the metric.

- Are there nevertheless preferred inertial coordinate systems, i.e. preferred choices for the higher-order terms in the Taylor expansion? The answer to this is "yes". One such preferred and geometrically natural class of inertial coordinate systems are e.g. Riemann normal coordinates, based on geodesics at the point p, and briefly discussed below.

2. Geodesic (or Riemann Normal) Coordinates

   A slightly more insightful way of constructing a locally inertial coordinate system, rather than by directly solving the relevant differential equation, makes use of geodesics at p. Recall that in Minkowski space the metric takes the simplest possible form in coordinates whose coordinate lines are (orthogonal) geodesics. One might thus suspect that in a general metric space-time the metric will also (locally) look particularly simple when expressed in terms of such *geodesic coordinates*.

   Roughly speaking (I will give a more detailed argument below), since locally around p we can solve the geodesic equation with four linearly independent initial conditions, we can assume the existence of a coordinate system $\{\xi^a\}$ in which the coordinate lines are geodesics $\xi^a(\tau) = \xi^a\tau$. This means that in these coordinates geodesics satisfy $\ddot{\xi}^a = 0$. Comparing with the full geodesic equation in these coordinates, one sees that this implies that

   $$\Gamma^a_{\ bc}\dot{\xi}^b\dot{\xi}^c = 0 \ . \tag{3.161}$$

   As at p the $\dot{\xi}^a$ were chosen to be linearly independent, this implies $\Gamma^a_{\ bc}(p) = 0$, as desired. It is easy to see that the coordinates ξ^a can also be chosen in such a way that $g_{ab}(p) = \eta_{ab}$ (by choosing the four directions at p to be orthonormal unit vectors).

Before turning to the more detailed construction, let us look at an example. Consider the standard metric $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$ on the two-sphere. Any point is as good as any other point, and one can construct an inertial coordinate system at the north pole $\theta = 0$ in terms of geodesics shot off from the north pole into the $\phi = 0$ (ξ^1) and $\phi = \pi/2$ (ξ^2) directions. The affine parameter along a great circle (geodesic) connecting the north pole to a point (θ, ϕ) is θ, and thus θ is also the geodesic distance, and the coordinates of the point (θ, ϕ) are $(\xi^1 = \theta\cos\phi, \xi^2 = \theta\sin\phi)$. In particular, the north pole is the origin $\xi^1 = \xi^2 = 0$. Note that one could have guessed these coordinates from the fact that near $\theta = 0$ the metric is $d\theta^2 + \theta^2 d\phi^2$, which is the Euclidean metric in polar coordinates $(\theta\cos\phi, \theta\sin\phi)$.

Calculating the metric in these new components, using

$$(\xi^1 = \theta\cos\phi, \xi^2 = \theta\sin\phi) \quad \Rightarrow \quad ((\xi^1)^2 + (\xi^2)^2 = \theta^2, \xi^2/\xi^1 = \tan\phi) \tag{3.162}$$

and thus

$$d\theta = \frac{\xi^1 d\xi^1 + \xi^2 d\xi^2}{\sqrt{(\xi^1)^2 + (\xi^2)^2}} \quad , \quad d\phi = \frac{\xi^1 d\xi^2 - \xi^2 d\xi^1}{(\xi^1)^2 + (\xi^2)^2} \quad , \tag{3.163}$$

one finds

$$d\theta^2 + \sin^2\theta d\phi^2 = (d\xi^1)^2 + (d\xi^2)^2 + \mathcal{O}(\xi^2 d\xi^2) \quad , \tag{3.164}$$

i.e.

$$g_{ab}(\xi) = \delta_{ab} + \mathcal{O}(\xi^2) \quad . \tag{3.165}$$

Therefore

$$g_{ab}(\xi = 0) = \delta_{ab} \quad , \quad g_{ab,c}(\xi = 0) = 0 \quad , \tag{3.166}$$

as required.

We now (re)turn to the general construction of such coordinates, starting with the geodesic equation

$$\ddot{x}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{x}^\beta\dot{x}^\gamma = 0 \quad . \tag{3.167}$$

We consider geodesics passing through (or emanating from) the point p with coordinates x_0^α at $\tau = 0$, and with initial 4-velocity u_0^α,

$$x^\alpha(\tau = 0) = x_0^\alpha \quad , \quad \dot{x}^\alpha(\tau = 0) = u_0^\alpha \quad . \tag{3.168}$$

It then follows that the 2nd derivative at $\tau = 0$ is given by

$$\ddot{x}^\alpha(\tau = 0) = -\Gamma^\alpha_{\beta\gamma}(x_0)u_0^\beta u_0^\gamma \quad . \tag{3.169}$$

Hence in a Taylor expansion around $\tau = 0$ we can write the solution to the geodesic equation as

$$x^\alpha(\tau) = x_0^\alpha + \tau u_0^\alpha - \tfrac{1}{2}\tau^2\Gamma^\alpha_{\beta\gamma}(x_0)u_0^\beta u_0^\gamma + \ldots \quad . \tag{3.170}$$

We can expand the (arbitrary) initial 4-velocity u_0^α in terms of 4 linearly independent (and orthonormal, say) vectors at p as

$$u_0^\alpha = \lambda^a e_a^\alpha \quad , \quad g_{\alpha\beta}(x_0) e_a^\alpha e_b^\beta = \eta_{ab} \quad . \tag{3.171}$$

We can then think of the Taylor expansion (3.170) as defining a coordinate transformation

$$x^\alpha(\xi) = x_0^\alpha + (\xi - \xi_0)^a e_a^\alpha - \tfrac{1}{2}(\xi - \xi_0)^b (\xi - \xi_0)^c \Gamma^\alpha_{\beta\gamma}(x_0) e_b^\beta e_c^\gamma + \ldots \quad , \tag{3.172}$$

which has the following properties:

(a) First of all, for

$$\xi^a(\tau) = \xi_0^a + \tau \lambda^a \tag{3.173}$$

this reduces to the Taylor-expanded solution (3.170) of the geodesic equation with $u_0^\alpha = \lambda^a e_a^\alpha$. Thus in this coordinate system in particular the 4 coordinate lines

$$\xi^a_{(b)}(\tau) = \xi_0^a + \tau \delta_b^a \quad , \quad b = 0,1,2,3 \quad , \tag{3.174}$$

are (affinely parametrised) geodesics, as desired.

(b) Moreover, up to quadratic order in the Taylor expansion (3.172) is identical to the coordinate transformation (3.159). In particular, this establishes that the geodesic coordinates constructed here are a special class of inertial coordinates, with

$$g_{ab}(\xi_0) = \eta_{ab} \quad , \quad \Gamma^a_{\ bc}(\xi_0) = 0 \quad . \tag{3.175}$$

(c) From the present point of view, the 2nd condition arises from the fact (mentioned above) that in these coordinates the geodesic equation for the above geodesics reduces to

$$\ddot{\xi}^a + \Gamma^a_{\ bc} \dot{\xi}^b \dot{\xi}^c = 0 \quad \Rightarrow \quad \Gamma^a_{\ bc}(\xi_0^a + \tau \lambda^a) \lambda^b \lambda^c = 0 \quad . \tag{3.176}$$

At ξ_0, i.e. for $\tau = 0$, the Christoffel symbols are independent of the λ^a, and therefore

$$\Gamma^a_{\ bc}(\xi_0) \lambda^b \lambda^c = 0 \quad \forall \lambda^a \quad \Rightarrow \Gamma^a_{\ bc}(\xi_0) = 0 \quad , \tag{3.177}$$

as claimed.

(d) In contrast to the previous construction leading to (3.159), here the higher-order terms in the Taylor expansion of the coordinate transformation are now determined by the higher-order terms in the Taylor expansion of the solution (3.170) of the geodesic equation. These higher-order terms will depend on 2nd and higher derivatives of the metric $g_{\alpha\beta}(x)$ at x_0, and these in turn will also determine the quadratic and higher terms of the Taylor expansion of the metric in these coordinates,

$$\begin{aligned} g_{ab}(\xi) &= g_{ab}(\xi_0) + (\xi-\xi_0)^c g_{ab,\,c}(\xi_0) + \tfrac{1}{2}(\xi-\xi_0)^c(\xi-\xi_0)^d g_{ab,\,cd}(\xi_0) + \dots \\ &= \eta_{ab} + \tfrac{1}{2}(\xi-\xi_0)^c(\xi-\xi_0)^d g_{ab,\,cd}(\xi_0) + \dots \quad . \end{aligned} \tag{3.178}$$

We will determine the quadratic term in this expansion (expressed in terms of the Riemann curvature tensor) in section 8.9.

3. A Numerological Argument

This is my favourite argument because it requires no calculations and at the same time provides additional insight into the nature of curved space-times.

Assuming that the local existence of solutions to differential equations is guaranteed by some mathematical theorems, it is frequently sufficient to check that one has enough degrees of freedom to satisfy the desired initial conditions (one may also need to check integrability conditions). Here we are looking at something even more elementary, namely the functional freedom contained in the coordinate transformations to impose certain conditions at one point. In the present context, this argument is useful because it also reveals some information about the 'true' curvature hidden in the second derivatives of the metric. It works as follows:

(a) Zero'th Derivatives:

Consider a Taylor expansion of the metric around p in the sought-for new coordinates. Then the metric at p will transform with the matrix $(\partial x^\mu/\partial\xi^a)(p)$. This matrix has $(4\times 4) = 16$ independent components, precisely enough to impose the 10 conditions $g_{ab}(p) = \eta_{ab}$ up to Lorentz transformations.

(b) First Derivatives:

The derivative of the metric at p, $g_{ab,c}(p)$, will appear in conjunction with the second derivative $\partial^2 x^\mu/\partial\xi^a\partial\xi^b$. The $4\times(4\times 5)/2 = 40$ coefficients are precisely sufficient to impose the 40 conditions $g_{ab,c}(p) = 0$.

(c) Second Derivatives:

Now let us look at the second derivatives of the metric. $g_{ab,cd}$ has $(10\times 10) = 100$ independent components, while the third derivative of $x^\mu(\xi)$ at p, $\partial^3 x^\mu/\partial\xi^a\partial\xi^b\partial\xi^c$ has $4\times(4\times 5\times 6)/(2\times 3) = 80$ components. Thus 20 linear combinations of the second derivatives of the metric at p cannot in general be set to zero by a coordinate transformation. Thus these encode the information about the real curvature at p. This agrees nicely with the fact that the Riemann curvature tensor we will construct later turns out to have precisely 20 independent components.

Repeating this argument in space-time dimension $D = d + 1$, one finds that the number of 2nd derivatives of the metric modulo coordinate transformations is

$$\left(\frac{D(D+1)}{2}\right)^2 - D\frac{D(D+1)(D+2)}{6} = \frac{1}{12}D^2(D^2-1) \ . \tag{3.179}$$

Again this turns out to agree with the number of independent components (8.34) of the curvature tensor in D dimensions.

Note:

At this point in the course I find it useful to develop in parallel (and suggest to read in parallel)

- the more formal material on tensor analysis in sections 4, 5, 6, 7, 8 and 11, say (and then moving on to the Einstein equations themselves)
- and a detailed discussion of the basic properties of the Schwarzschild metric (sections 25 - 27),

since much of the latter (in particular geodesics, solar system tests of general relativity, even the issues that arise in connection with the Schwarzschild radius) can be understood just on the basis of what has been done so far (if, for the time being, one accepts on faith that the Schwarzschild metric is the unique spherically symmetric vacuum solution of the Einstein field equations).

Not only is this an interesting and physically relevant application of the machinery developed so far, it also provides an appropriate balance between physics and formalism in the lectures. More advanced material in the intervening sections can then be covered and dealt with if and when needed or desired (or, ideally, both).

4 Tensor Algebra

4.1 Principle of General Covariance

The Einstein Equivalence Principle tells us that the laws of nature (including the effects of gravity) should be such that in an inertial frame they reduce to the laws of Special Relativity.

A prime example of this is the geodesic equation

$$\ddot{x}^{\alpha} + \Gamma^{\alpha}_{\beta\gamma}\dot{x}^{\beta}\dot{x}^{\gamma} = 0 \quad , \tag{4.1}$$

describing the motion of a particle in a gravitational field. Using the Einstein Equivalence Principle, and the *General Covariance* of the geodesic equation established in section 1.6, one can argue as follows to establish that this is the "correct" equation at an aribitrary point p in an arbitrary gravitational field described by a metric $g_{\alpha\beta}(x)$:

- Choose an inertial coordinate system z^a centered at p, i.e. with $g_{ab}(p) = \eta_{ab}$ and $\Gamma^{a}_{bc}(p) = 0$. Then one has

$$\left(\ddot{x}^{\alpha} + \Gamma^{\alpha}_{\beta\gamma}\dot{x}^{\beta}\dot{x}^{\gamma}\right)(p) = \left(\frac{\partial x^{\alpha}}{\partial z^{a}}\right)\left(\ddot{z}^{a} + \Gamma^{a}_{bc}\dot{z}^{b}\dot{z}^{c}\right)(p) = \left(\frac{\partial x^{\alpha}}{\partial z^{a}}\right)\ddot{z}^{a}(p) \quad . \tag{4.2}$$

- By the Equivalence Principle, $\ddot{z}^a = 0$ is the correct equation for a particle in a freely falling reference system (such as at the origin p of an inertial coordinate system). Therefore, by the above identity, at the point p the geodesic equation is also satisfied in arbitrary coordinates x^{α}.

- Both the point p and the metric $g_{\alpha\beta}(x)$ are arbitrary, and therefore the geodesic equation holds at an arbitrary point of an arbitrary gravitational field.

This argument generalises and shows that the Einstein Equivalence Principle provides us with a general link between equations describing physics in an arbitrary gravitational field and general covariance. In other words, general covariance provides us with a concrete way of implementing the Einstein Equivalence Principle. One classical (but not entirely satisfactory - see the discussion below) textbook formulation of this *Principle of General Covariance* is[9]

Principle of General Covariance

By virtue of the Einstein Equivalence Principle, a physical equation holds in an arbitrary gravitational field if

[9]S. Weinberg, *Gravitation and Cosmology* section 4.1

1. the equation holds in the absence of gravity, i.e. when $g_{\mu\nu} = \eta_{\mu\nu}$, $\Gamma^{\mu}_{\ \nu\lambda} = 0$, and
2. the equation is *generally covariant*, i.e. preserves its form under a general coordinate transformation.

We will turn momentarily to a proof of (a slightly modified version of) this statement. First, however, I would like to add a *caveat* to the 1st and a clarification to the 2nd condition (starting with the latter):

ad 2: We first need to clarify (and then reformulate slightly) the 2nd condition, as the statement "preserves its form under a general coordinate transformation" is neither completely unambiguous (without further explanation or definitions) nor totally to the point.

Concretely, the 2nd condition means the following: assume that you have some physical equation that in some coordinate system takes the form $T = 0$, where $T = 0$ could be some multi-component (thus T is adorned with various indices) differential equation. Now perform a coordinate transformation $x \to y(x)$, and assume that the new object T' has the form

$$T' = (\ldots)T + \text{junk} \tag{4.3}$$

where the term in brackets is some invertible matrix or operator. Then clearly the presence of the junk-terms means that the equation $T' = 0$ is not equivalent to the equation $T = 0$. An example of an object that transform in this way is, as we have seen, the Christoffel symbols. On the other hand, if these junk terms are absent, so that we have

$$T' = (\ldots)T \tag{4.4}$$

then one might like to say that the equation has preserved its form under a general coordinate transformation.

As a consequence of this, one also has $T' = 0$ if and only if $T = 0$, i.e. the equation is satisfied in one coordinate system if and only if it is satisfied in any other (or all) coordinate systems. An example of this is the geodesic equation which, as we have seen, transform precisely in such a way, with the term $(\ldots)$ in brackets being the Jacobi matrix.

However, if this is what one desires (and it is), then one may as well say this directly. Thus, to be more concrete, we can replace the 2nd condition above by

2' the equation is *generally covariant*, i.e. it is satisfied in one coordinate system iff it is satisfied in all coordinate systems.

ad 1: The argument below will invoke general covariance in order to be able to look at a given equation at the origin of an inertial (freely falling) coordinate system. As we have seen in various ways in section 3.6, at that point p, the effects of gravity are absent to the extent that the metric at that point is the Minkowski metric, $g_{ab}(p) = \eta_{ab}$, and that the derivatives of the metric (or Christoffel symbols) at that point are zero, $g_{ab,c}(p) = 0$. However, as also noted there, higher derivatives of the metric can in general not also be chosen to be zero at that point (and indeed the second derivatives of the metric at that point will turn out to contain coordinate independent information about the curvature of the space-time).

In that sense, looking at an equation at the origin of an inertial coordinate system is not strictly identical to looking at that same equation in the absence of gravity (i.e. in Minkowski space), where all these higher derivatives of the metric are also zero (in inertial coordinates). For the purposes of the argument below we will ignore this difference, and thus allow for the possibility that the Einstein Equivalence Principle holds only to first order in derivatives of the metric. Indeed, as discussed in more detail in section 8.4, and as could be anticipated from our identification of the metric with the potential of the gravitational field, second derivatives of the metric encode tidal gravitational forces (which one cannot expect to be able to eliminate by passing to a freely falling reference system), so this is a plausible relaxation of some stricter interpretation of the Einstein Equivalence Principle. We will look at the implications of this in the remarks below.

With these remarks in mind, let us now establish the above statement, namely that the Einstein equivalence principle implies that an equation that satisfies the conditions 1 and 2 (or 2') is valid in an arbitrary gravitational field:

- consider some equation that satisfies these conditions, and assume that we are in an arbitrary gravitational field;
- condition 2' implies that this equation is true (or satisfied) in all coordinate systems if it is satisfied just in one coordinate system;
- now we know that we can always (locally) construct a freely falling coordinate system in which the effects of gravity are absent;
- the Einstein Equivalence Principle now posits that in such a reference system the physics is that of Minkowski space-time;
- condition 1 means that the equation is true (satisfied) there;
- thus it is valid in all coordinate systems;
- since we started off by considering an arbitrary gravitational field, it follows that the equation is now valid in an arbitrary gravitational field, as claimed in the Principle of General Covariance.

Remarks:

1. Note that general covariance alone is an empty statement since any equation (whether correct or not) can be made generally covariant simply by writing it in an arbitrary coordinate system (cf. also the discussion in section 6.4). It develops its power only when used in conjunction with the Einstein Equivalence Principle as a statement about physics in a gravitational field, namely that by virtue of its general covariance an equation will be true in a gravitational field if it is true in the absence of gravitation.

2. As alluded to above, the principle of general covariance does not fix the equations uniquely because there are generally covariant objects that one can construct e.g. from the (second) derivatives of the metric (via the Riemann curvature tensor to be introduced in section 8) that can therefore be added to an equation and which vanish for Minkowski space, i.e. in the absence of gravitation.

3. In section 6.1 we will introduce a recipe / algorithm, the *principle of minimal coupling*, that allows us to produce generally covariant equations from those of special relativity. However, as we will discuss in section 8.10, also this description is ambiguous. The upshot is that there is no unique way of implementing the principle of general covariance, but this was probably too much to hope for anyway.

4.2 Tensors and Tensor Fields

In order to construct generally covariant equations, we need objects that transform in a simple way under coordinate transformations

$$x^\alpha \to y^\mu(x) = \bar{x}^\mu(x) \tag{4.5}$$

(it will be convenient to be able to switch between denoting the new coordinates by either y^μ or $\bar{x}^\mu$, depending on what is easier to read or write in a given situation). The prime examples of such objects are *tensors*, which transform multi-linearly with the Jacobi matrices

$$J^\mu_\alpha = \frac{\partial y^\mu}{\partial x^\alpha} \quad , \quad J^\alpha_\mu = \frac{\partial x^\alpha}{\partial y^\mu} \quad , \tag{4.6}$$

satisfying

$$J^\alpha_\mu J^\mu_\beta = \delta^\alpha_\beta \quad , \quad J^\mu_\alpha J^\alpha_\nu = \delta^\mu_\nu \quad . \tag{4.7}$$

If you are already familiar with Lorentz tensors from special relativity (as briefly recalled in section 1.2, these are objects which transform in a particularly simple multi-linear way under Lorentz transformations), then hardly anything in this or the subsequent section 4.3 should be new or unexpected (but interesting new features will arise in particular when we move on from tensor algebra to tensor analysis in section 5).

1. Scalars

   The simplest example of a tensor is a function (or scalar) f which under a coordinate transformation $x^\alpha \to \bar{x}^\mu(x)$ simply transforms as

$$\bar{f}(\bar{x}(x)) = f(x) \quad . \tag{4.8}$$

   One frequently suppresses the argument, and thus writes simply $\bar{f} = f$, expressing the fact that, up to the obvious and familiar change of argument, functions are invariant under coordinate transformations.

   <u>REMARKS:</u>

   (a) Just in case this transformation behaviour is not as familiar as it should be, consider the transformation from Cartesian (x^1, x^2) to polar $(y^1 = r, y^2 = \phi)$ coordinates, and the function $f(x^1, x^2) = (x^1)^2 + (x^2)^2$. In terms of polar coordinates, this function can be written as $\bar{f}(r, \phi) = r^2$. While f and $\bar{f}$ are evidently different functions of 2 variables, they are equal up to the change of coordinates, and this is how functions or scalars transform in general.

   (b) Note that the coordinates themselves are functions (that assign coordinates to a point p). Thus the x-coordinates can be regarded as 4 functions $f^{(\alpha)}$ defined in the coordinate system x by $f^{(\alpha)}(x) = x^\alpha$. Likewise, the new coordinates y^μ can be regarded as being defined by 4 in general genuinely *different* functions of x, say $g^{(\mu)}(x) = y^\mu(x)$.

   Under changes of coordinates, these functions transform as they should, i.e. one has $\bar{f}^{(\alpha)}(y) = f^{(\alpha)}(x) = x^\alpha$, and likewise for the functions $g^{(\mu)}(x)$. In the example above, the former would simply be the Cartesian coordinates expressed in terms of polar coordinates, $\bar{f}^{(1)}(r, \phi) = r \cos \phi = x^1 = f^{(1)}(x^1, x^2)$. On the other hand, $g^{(1)}(r, \phi) = r$, which is quite distinct from $\bar{f}^{(1)}(r, \phi) = r \cos \phi$. In that sense, allowing oneself to also denote the new coordinates y^μ by $\bar{x}^\mu$ may initially be a source of confusion (but is frequently so useful that we will continue to do this anyway ...).

   For a slightly more sophisticated perspective on these things, which disentangles functions from their representatives in terms of local coordinates, see the discussion of manifolds and local charts in section 5.11.

2. Vectors

   The next simplest case are objects that transform under coordinate transformations with the Jacobi matrix J^μ_α. Such objects are called vectors and have components $V^\alpha(x)$ transforming as

$$\bar{V}^\mu(y(x)) = J^\mu_\alpha(x) V^\alpha(x) \quad . \tag{4.9}$$

These are the prototypes of tensors and indeed all the other tensorial objects that we will encounter in the following are rather straightforward generalisations of such vectors.

<u>REMARKS:</u>

(a) One comment on notation: since already the index μ on $\bar{V}^\mu$ indicates that these are the components of a vector with resepct to the coordinates $y^\mu = \bar{x}^\mu$, it is not strictly speaking necessary to indicate this as well by putting a bar on the vector. Therefore one could (and commonly just does) write simply

$$V^\mu(y) \equiv \bar{V}^\mu(y) \quad . \tag{4.10}$$

(b) One comment on terminology: it is sometimes useful to distinguish *vectors* from *vector fields* and, likewise, *tensors* from *tensor fields*. A vector is then just a vector $V^\alpha(x)$ at some point x of space-time whereas a vector field is something that assigns a vector to each point of space-time,

$$\text{vector field} : \quad x \mapsto V^\alpha(x) \tag{4.11}$$

and likewise for scalars and scalar fields, and more general tensors and tensor fields.

(c) An intermediate situation is that of a vector or tensor defined neither just at a single point nor in the entire space-time but for example along some curve (representing e.g. the worldline of a particle). A prime example is the tangent vector $\dot{x}^\alpha$ to a curve, which indeed transforms as a vector,

$$\dot{x}^\alpha \to \dot{y}^\mu = \frac{\partial y^\mu}{\partial x^\alpha}\dot{x}^\alpha = J^\mu_a \dot{x}^\alpha \quad . \tag{4.12}$$

And while the naive acceleration $\ddot{x}^\alpha$ is *not* a vector, as we have seen on several occasions before, we can define a vectorial acceleration by

$$a^\alpha = \ddot{x}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{x}^\beta \dot{x}^\gamma \quad . \tag{4.13}$$

We will rederive this expression for a covariant acceleration from a different perspective in section 5.7.

(d) One way of thinking about vector fields is as tangent vector fields to families of curves on a space or space-time which arise as the solutions to the differential equation

$$\frac{d}{d\sigma}x^\alpha(\sigma) = V^\alpha(x(\sigma)) \tag{4.14}$$

(and we take local existence and uniqueness of these solutions under suitable regularity and differentiability conditions for granted). These curves $x^\alpha(s)$ are the *integral curves* (or *orbits*) of the vector field V^α, and by by construction they are characterised by the fact that at any point x the tangent vector to the curve passing through that point is the vector $V^\alpha(x)$ at that point. Thus vector fields also generate a *flow* on the space(-time), namely the motion of points along these integral curves, $x^\alpha(\sigma) \mapsto x^\alpha(\sigma+s)$ for $s \in \mathbb{R}$.

(e) An extremely useful related way of thinking about vectors (vector fields) is as first order differential operators, via the correspondence

$$V^\alpha \Leftrightarrow V := V^\alpha \partial_\alpha \quad . \tag{4.15}$$

This operator gives the directional derivative along the flow of the vector field V, as in

$$\frac{d}{d\sigma} f(x(\sigma)) = V^\alpha \partial_\alpha f(x(\sigma)) \tag{4.16}$$

(by the chain rule). One of the advantages of this point of view is that the object V is completely invariant under coordinate transformations as the *components* V^α of V transform inversely to the *basis vectors* ∂_α. For more on this see sections 4.6 and 4.8 on the coordinate-independent interpretation of tensors below.

3. Covectors

A covector (field) is an object $U_\alpha(x)$ which under a coordinate transformation transforms inversely to a vector, i.e. as

$$\bar{U}_\mu(y(x)) \equiv U_\mu(y(x)) = J^\alpha_\mu(y(x)) U_\alpha(x) \quad . \tag{4.17}$$

A familiar example of a covector is the derivative $U_\alpha = \partial_\alpha f$ of a function (scalar) which of course transforms as

$$\partial_\mu \bar{f}(y(x)) = J^\alpha_\mu \partial_\alpha f(x) \quad , \tag{4.18}$$

because that is how partial derivatives transform.

<u>Remarks:</u>

(a) As in the case of covectors of special relativity (1.44), one should think of covectors pointwise as elements of the dual vector space $\mathbb{V}^*$ to the space of vectors $\mathbb{V}$, i.e. as linear functionals on the space of vectors, given by

$$U(x): \quad V^\alpha(x) \;\mapsto\; U_\alpha(x) V^\alpha(x) \;\in\; \mathbb{R} \quad . \tag{4.19}$$

The transformation properties of U_α and V^α guarantee that the result is a scalar (function) under coordinate transformations,

$$U_\alpha V^\alpha = \bar{U}_\mu \bar{V}^\mu \quad . \tag{4.20}$$

(b) Just as it was useful to think of a vector field $V^\alpha(x)$ in a more coordinate independent way as the components of the coordinate-independent object $V = V^\alpha\partial_\alpha$ with respect to the basis ∂_α, one can think of covector fields U_α as the components of an object

$$U = U_\alpha dx^\alpha \tag{4.21}$$

(which is invariant under coordinate transformations) with respect to the basis dx^α. The prime example is again the *differential*

$$df = \partial_\alpha f(x)dx^\alpha = \partial_\mu \bar{f}(y)dy^\mu \tag{4.22}$$

of a scalar.

(c) Combining the two points of view in the remarks above, one can thus think of df as the linear functional on vector fields that assigns to a vector field V the scalar which is the derivative of f along V,

$$\begin{aligned} df: \quad & \text{vector fields} \longrightarrow \text{scalar fields} \\ & df(V) = V^\alpha\partial_\alpha f \equiv Vf \quad . \end{aligned} \tag{4.23}$$

4. Covariant 2-Tensors

Clearly, given the above objects, we can construct more general objects which transform in a nice way under coordinate transformations by taking products of them. Tensors in general are objects which transform like (but need not be equal to) products of vectors and covectors.

In particular, a covariant 2-tensor, or (0,2)-tensor, is an object $A_{\alpha\beta}$ that transforms under coordinate transformations like the product of two covectors, i.e.

$$\bar{A}_{\mu\nu}(y(x)) \equiv A_{\mu\nu}(y(x)) = J^\alpha_\mu(y(x))J^\beta_\nu(y(x))A_{\alpha\beta}(x) \quad . \tag{4.24}$$

We already know one example of such a tensor, namely the metric tensor $g_{\alpha\beta}$ (which happens to be a symmetric tensor).

In order to avoid unnecessary clutter and to make tensorial equations more readable, I will from now on typically use a shorthand notation in which I drop not only the bars, overlines or other decorations on the transformed objects but in which I also omit the argument x or $y(x)$. In this notation, the above equation would then simply become

$$A_{\mu\nu} = J^\alpha_\mu J^\beta_\nu A_{\alpha\beta} \quad , \tag{4.25}$$

which is more user-friendly.

5. Contravariant 2-Tensors

Likewise we define a contravariant 2-tensor (or a (2,0)-tensor) to be an object $B^{\alpha\beta}$ that transforms like the product of two vectors,

$$B^{\mu\nu} = J^\mu_\alpha J^\nu_\beta B^{\alpha\beta} \quad . \tag{4.26}$$

An example is the inverse metric tensor $g^{\alpha\beta}$.

6. (p,q)-Tensors

   It should now be clear how to define a general (p,q)-tensor - namely as an object with p contravariant (upper) and q covariant (lower) indices which under a coordinate transformation transforms like a product of p vectors and q covectors,

$$T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q} = J^{\mu_1}_{\alpha_1}\ldots J^{\mu_p}_{\alpha_p}\, J^{\beta_1}_{\nu_1}\ldots J^{\beta_q}_{\nu_q}\, T^{\alpha_1\cdots\alpha_p}_{\beta_1\cdots\beta_q} \quad . \tag{4.27}$$

<u>Remarks:</u>

1. Note that, in particular, a tensor is zero (at a point) in one coordinate system if and only if the tensor is zero (at the same point) in another coordinate system.

   Thus, any law of nature (field equation, equation of motion) expressed in terms of tensors, say in the form $T^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q} = 0$, preserves its form under coordinate trasformations and is therefore automatically generally covariant,

$$T^{\alpha_1\cdots\alpha_p}_{\beta_1\cdots\beta_q}(x) = 0 \quad \forall x \quad \Leftrightarrow \quad T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}(y) = 0 \quad \forall y \quad . \tag{4.28}$$

2. An important special example of a tensor is the Kronecker tensor δ^α_β. That this is indeed a tensor follows from the fact that it can e.g. be written as a contraction of a product of tensors, namely the metric tensor and its inverse,

$$\delta^\alpha_\beta = g^{\alpha\gamma} g_{\gamma\beta} \quad , \tag{4.29}$$

   and the calculational rules of tensor algebra to be established in section 4.3 below. Calculating explicitly the components with respect to the new coordinates $y^\mu = \bar{x}^\mu$, one finds

$$\bar{\delta}^\mu_\nu = J^\mu_\alpha J^\beta_\nu\, \delta^\alpha_\beta = J^\mu_\alpha J^\alpha_\nu = \delta^\mu_\nu \quad . \tag{4.30}$$

   Thus the Kronecker tensor has the same components in all coordinate systems. This is reassuring but should not be too surprising. Together with scalars and products of scalars and Kronecker tensors it is the only tensor whose components are the same in all coordinate systems.

3. A covariant 2-tensor $T_{\alpha\beta}$, say, is said to be symmetric if $T_{\alpha\beta} = T_{\beta\alpha}$ and anti-symmetric if $T_{\alpha\beta} = -T_{\beta\alpha}$. This is well-defined because it is a generally covariant notion: a tensor is symmetric in all coordinate system iff it is symmetric in one coordinate system, etc.

This definition can be extended to any or all pairs of covariant indices or pairs of contravariant indices. Thus e.g. a tensor $T^{\alpha_1\ldots\alpha_p}$ is called totally symmetric (or totally anti-symmetric) if it is symmetric (anti-symmetric) under the exchange of any pair of indices.

On the other hand, it is not meaningful to talk of the symmetry of a (1,1)-tensor, say, as an equation like $T^\alpha_\beta = T^\beta_\alpha$ does not make any sense.

Symmetrisation and anti-symmetrisation of tensors will be discussed in section 4.3 below.

4. The number of independent components of a general (p, q)-tensor is 4^{p+q}. The number of independent components is reduced if the tensor has some symmetry properties. Thus a symmetric (0,2)- or (2,0)-tensor has $4 \times 5/2 = 10$ independent components, an anti-symmetric (0,2)- or (2,0)-tensor has $4 \times 3/2 = 6$ independent components, and a totally anti-symmetric $(0, 4)$-tensor $T_{\beta_1\ldots\beta_4}$ has only got one independent component, namely T_{0123} (all the others being determined by anti-symmetry).

5. Important examples of non-tensors are the Christoffel symbols. Another important example is the the ordinary partial derivative of a (p, q)-tensor, $\partial_\gamma T^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}$ which is *not* a $(p, q+1)$-tensor unless $p = q = 0$. This failure of the partial derivative to map tensors to tensors will motivate us below to introduce a *covariant derivative* which generalises the usual notion of a partial derivative and has the added virtue of mapping tensors to tensors.

4.3 Tensor Algebra

Tensors can be added, multiplied and contracted in certain obvious ways. The basic algebraic operations are the following:

1. Linear Combinations

   Given two (p, q)-tensors $A^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}$ and $B^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}$, their sum

$$C^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q} = A^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q} + B^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q} \tag{4.31}$$

   is also a (p, q)-tensor.

2. Direct Products

   Given a (p, q)-tensor $A^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}$ and a (p', q')-tensor $B^{\gamma_1\ldots\gamma_{p'}}_{\phantom{\gamma_1\ldots\gamma_{p'}}\delta_1\ldots\delta_{q'}}$, their direct product

$$A^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q} B^{\gamma_1\ldots\gamma_{p'}}_{\phantom{\gamma_1\ldots\gamma_{p'}}\delta_1\ldots\delta_{q'}} \tag{4.32}$$

   is a $(p + p', q + q')$-tensor,

3. Contractions

   Given a (p,q)-tensor with p and q non-zero, one can associate to it a $(p-1,q-1)$-tensor via contraction of (i.e. summation over) one covariant and one contravariant index, for example

$$A^{\alpha_1\ldots\alpha_p}{}_{\beta_1\ldots\beta_q} \quad \rightarrow \quad B^{\alpha_1\ldots\alpha_{p-1}}{}_{\beta_1\ldots\beta_{q-1}} = A^{\alpha_1\ldots\alpha_{p-1}\gamma}{}_{\beta_1\ldots\beta_{q-1}\gamma} \quad . \tag{4.33}$$

   This operation of contraction generalises the trace. In general, there are in total $(p \times q)$ distinct possibilities to do this, by contracting the 1st oder 2nd or ... or p'th upper index with the 1st or 2nd or ... or q'th lower index.

   That this indeed leads to a $(p-1,q-1)$-tensor follows from the fact that the Jacobi matrices corresponding to the two contracted indices are inverses to each other, and are multiplied (by the contraction), resulting in the identity matrix. This then leaves only $(p-1)$ Jacobi matrices and $(q-1)$ inverse Jacobi matrices, corresponding to the transformation behaviour of a $(p-1,q-1)$-tensor.

   For example, for a $(1,1)$-tensor T^{α}_{β} one has the unique contraction

$$T^{\alpha}_{\beta} \rightarrow T^{\alpha}_{\alpha} \quad . \tag{4.34}$$

   This is just the standard trace of the linear tranformation represented by the $(1,1)$-tensor T^{α}_{β}. Under the coordinate transformation $x^{\alpha} \rightarrow y^{\mu}$ this trace transforms as

$$T^{\mu}_{\mu} = J^{\mu}_{\alpha} J^{\beta}_{\mu} T^{\alpha}_{\beta} = \delta^{\beta}_{\alpha} T^{\alpha}_{\beta} = T^{\alpha}_{\alpha} \quad . \tag{4.35}$$

   Thus the trace is invariant (i.e. a scalar or $(0,0)$-tensor). Clearly this cancellation between the two Jacobi matrices corresponding to two contracted indices also takes place when the tensor T carries other non-contracted indices, and therefore the contraction of a (p,q)-tensor is a $(p-1,q-1)$-tensor, as claimed.

4. Raising and Lowering of Indices

   These operations can of course be combined in various ways. A particular important operation is, given a metric tensor, the lowering of indices with the metric, and the raising of indices with the inverse metric.

   In particular, by direct product with the metric and subsequent contraction, we can construct a covector from a vector by

$$V^{\alpha} \rightarrow g_{\alpha\beta} V^{\beta} \quad . \tag{4.36}$$

   When it is clear from the context which metric is being used, it is common and convenient to simply call this covector V_{α},

$$V_{\alpha} \equiv g_{\alpha\beta} V^{\beta} \quad . \tag{4.37}$$

Likewise, to a covector one can associate a vector by using the inverse metric,

$$A_\alpha \to g^{\alpha\beta} A_\beta \equiv A^\alpha \quad . \tag{4.38}$$

<u>REMARKS:</u>

(a) Interpreted in terms of covectors as linear functions on vectors, this perhaps somewhat obscure convention and notation has a perfectly natural interpretation. Namely, recall that even though a vector space $\mathbb{V}$ and its dual $\mathbb{V}^*$ are isomorphic (in finite dimensions), there is no natural isomorphism between them, i.e. no natural identification of vectors in $\mathbb{V}$ with covectors in $\mathbb{V}^*$. However, if one has a scalar product (metric) $< v, w >$ on $\mathbb{V}$, then this provides an identification of $\mathbb{V}$ and $\mathbb{V}^*$ through

$$v \in \mathbb{V} \mapsto \alpha_v \in \mathbb{V}^* : \quad \alpha_v(w) =< v, w > \quad . \tag{4.39}$$

Thus a metric allows one to associate a covector to a vector, and in the notation of tensor algebra favoured by physicists this is conveniently just written as $V^\alpha \mapsto V_\alpha$.

(b) This convention of using the metric and its inverse to lower and raise indices on a tensor can of course be extended to higher rank tensors, but the result will in general depend on which index is lowered or raised. One then needs to make sure that the notation is sufficiently unambiguous to keep track of this. For example, the two (in general distinct) (1,1)-tensors one obtains from the lowering of one index of a (2,0)-tensor might be denoted amd distinguished by

$$g_{\gamma\alpha} T^{\alpha\beta} = T_\gamma^{\ \beta} \quad , \quad g_{\gamma\beta} T^{\alpha\beta} = T^\alpha_{\ \gamma} \quad . \tag{4.40}$$

(c) Finally note that this notation of raising and lowering indices with the metric is consistent with denoting the inverse metric by raised indices, i.e. it is indeed true that

$$g^{\alpha\beta} = g^{\alpha\gamma} g^{\beta\delta} g_{\gamma\delta} \tag{4.41}$$

(both indices of $g_{\gamma\delta}$ raised with two inverse metrics). That this is indeed correct follows from

$$g^{\alpha\gamma} g^{\beta\delta} g_{\gamma\delta} = g^{\alpha\gamma} \delta_\gamma^\beta = g^{\alpha\beta} \quad . \tag{4.42}$$

5. Symmetrisation and anti-Symmetrisation

Given any $(0,2)$-tensor $T_{\alpha\beta}$, say, one can decompose it into its symmetric and anti-symmetric parts as

$$T_{\alpha\beta} = \tfrac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha}) + \tfrac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha}) \equiv T_{(\alpha\beta)} + T_{[\alpha\beta]} \quad . \tag{4.43}$$

The decomposition into symmetric and anti-symmetric parts is invariant under coordinate transformations. In particular, when $T_{\alpha\beta}$ is a tensor, also $T_{(\alpha\beta)}$ and

$T_{[\alpha\beta]}$ are tensors, and thus (anti-)symmetrisation is yet another linear operation that one can perform on tensors.

The factor $\frac{1}{2}$ is chosen such that the symmetrisation of a symmetric tensor is the same as the original tensor,

$$T_{\alpha\beta} = T_{\beta\alpha} \quad \Rightarrow \quad T_{(\alpha\beta)} = T_{\alpha\beta} \quad , \quad T_{[\alpha\beta]} = 0 \tag{4.44}$$

(and likewise for the anti-symmetrisation of anti-symmetric tensors).

This can be generalised to the (anti-)symmetrisation of any pair of (contravariant or covariant) indices; e.g.

$$T_{(\alpha\beta)\gamma} = \tfrac{1}{2}(T_{\alpha\beta\gamma} + T_{\beta\alpha\gamma}) \tag{4.45}$$

is the symmetrisation of $T_{\alpha\beta\gamma}$ in its first and second index. It can also be generalised to the total (anti-)symmetrisation of a higher-rank tensor; e.g.

$$T_{(\alpha\beta\gamma)} \equiv \tfrac{1}{3!}(T_{\alpha\beta\gamma} + T_{\beta\alpha\gamma} + T_{\gamma\beta\alpha} + T_{\beta\gamma\alpha} + T_{\alpha\gamma\beta} + T_{\gamma\alpha\beta}) \tag{4.46}$$

is totally symmetric, i.e. symmetric under the exchange of any pair of indices, and

$$T_{[\alpha\beta\gamma]} \equiv \tfrac{1}{3!}(T_{\alpha\beta\gamma} - T_{\beta\alpha\gamma} - T_{\gamma\beta\alpha} + T_{\beta\gamma\alpha} - T_{\alpha\gamma\beta} + T_{\gamma\alpha\beta}) \tag{4.47}$$

is totally anti-symmetric. The prefactor $\frac{1}{6}$ is again there to ensure that the total symmetrisation of a totally symmetric tensor is the original tensor (and likewise for the total anti-symmetrisation of totally anti-symmetric tensors). This generalises in an evident way to higher rank p tensors, with the combinatorial prefactor $1/p!$.

An observation we will frequently make use of to recognise when some object is a tensor is the following (occasionally known as the *quotient theorem* or quotient lemma):

For example, say that in an equation of the form

$$A_\alpha = B_{\alpha\beta}C^\beta \tag{4.48}$$

you know that A_α transforms as a covector for any vector C^β. Then it follows that $B_{\alpha\beta}$ has to be a tensor. Likewise for higher rank tensors and more contractions, as in

$$A_{\alpha\beta} = B_{\alpha\beta\gamma\delta}C^{\gamma\delta} \quad . \tag{4.49}$$

Also in that case one has the statement that if A transforms as a tensor for every tensor C, then B itself has to be a tensor.

An elementary and ham-handed proof of these kinds of statements can be obtained by contradiction: assume that B does not transform as a tensor and write its transformation, as in (4.3), as

$$B' = (\ldots)B + \text{junk} \quad . \tag{4.50}$$

If "junk" $\neq 0$, then there will be some C such that "junk" contributes to the contraction $B'C'$. That means that "junk" contributes to A', the transformed A, contradicting the premise that A is a tensor.

4.4 Generally Covariant Integration and Volume Elements

While tensors are the objects which, in a sense, transform in the nicest and simplest possible way under coordinate transformations, they are not the only relevant objects. An important class of non-tensors (but "almost" tensors) are so-called *tensor densities*. They will play a crucial role for us in order to have a generally-covariant notion of integration at our disposal, and thus ultimately also a way of writing down generally covariant action principles for fields etc.

In this section we will address the issue of generally covariant integration in a space-time equipped with a metric. This will be accomplished with the help of a particular tensor density constructed from the metric. Having thus established that tensor densities are objects of legitimate interest in their own right, we will then discuss their properties in more generality in section 4.5 below.

To set the stage, consider once again first the situation in special relativity. In that case, the integral of a Lorentz scalar $f(\xi)$ with respect to the volume element $d^4\xi$ (or $d^{27}\xi\ \ldots$) is itself a Lorentz scalar, i.e. independent of the inertial reference frame in which the integral is evaluated,

$$\int d^4\xi\ f(\xi) = \int d^4\bar{\xi}\ \bar{f}(\bar{\xi}) \qquad (\bar{\xi}^a = L^a_{\ b}\xi^b) \tag{4.51}$$

The reasons for this are that

1. f is a scalar by assumption,
$$\bar{f}(\bar{\xi}) = f(\xi)\ \ , \tag{4.52}$$
2. by the fundamental theorem of integral calculus, under an arbitrary coordinate transformation $\xi \to \bar{\xi} = \bar{\xi}(\xi)$ the volume element transforms with the *Jacobian*, the (absolute value of the) determinant of the Jacobi matrix,
$$d^4\bar{\xi} = \left|\det \frac{\partial\bar{\xi}}{\partial\xi}\right| d^4\xi\ \ , \tag{4.53}$$
3. for a Lorentz transformation one has
$$\left|\det \frac{\partial\bar{\xi}}{\partial\xi}\right| = |\det L| = 1\ \ . \tag{4.54}$$
and thus the volume element is invariant under Lorentz transformations,
$$d^4\bar{\xi} = d^4\xi\ \ . \tag{4.55}$$

Turning now to general relativity and general covariance, it is immediately apparent that the integral of a scalar $\int d^4x\ f(x)$ will *not* be generally covariant, i.e. under a general coordinate transformation $x \to y = y(x)$ generically one has

$$\int d^4x\ f(x) \neq \int d^4y\ \bar{f}(y)(y) \tag{4.56}$$

because of the non-trivial Jacobian,

$$d^4y = \left|\det\left(\frac{\partial y}{\partial x}\right)\right| d^4x \quad . \tag{4.57}$$

One way out would be to abandon the idea that one should integrate scalars and to require that the integrand $\rho(x)$ should transform in such a way that it cancels the Jacobian arising from the measure, namely as

$$\bar{\rho}(y) = \left|\det\left(\frac{\partial y}{\partial x}\right)\right|^{-1} \rho(x) \quad . \tag{4.58}$$

This is indeed an option, and we will return to this below (see remark 1 in section 4.5), but at this stage this is rather unintuitive and not particularly useful, in particular because it is not clear how one should go about finding or constructing such objects in the first place.

A first simplification arises from the fact that (4.58) implies that

- if $\rho(x)$ satisfies (4.58), then so will $f(x)\rho(x)$ for any scalar $f(x)$;
- and conversely, if $\rho_1(x)$ and $\rho_2(x)$ satisfy (4.58), then $\rho_2(x)/\rho_1(x) \equiv f(x)$ is a scalar.

Thus, in order to find all objects that satisfy (4.58), we just need to find a single one, and with respect to that one as our measure we can then integrate any scalar function $f(x)$ in a generally covariant way,

$$\bar{\rho}(y) = \left|\det\left(\frac{\partial y}{\partial x}\right)\right|^{-1} \rho(x) \quad \Rightarrow \quad \int d^4y\; \bar{\rho}(y)\bar{f}(y) = \int d^4x\; \rho(x)f(x) \tag{4.59}$$

In order to find one candidate $\rho(x)$, let us approach this question in a different way. Integrals are used to calculate or measure volumes (or areas, or lenghts, or ...). Such integrals should have a coordinate-independent meaning, but they should depend on the prescription one uses for *measuring* volumes, areas, lenghts, ... These prescriptions are concisely encoded in the *metric*. Thus it is plausible that in order to define a generally covariant notion of integration one may need to specify the metric, but that this is all that one should need to know (while the Jacobian between two coordinate systems should fundamentally be irrelevant and be considered to be a red herring).

With this in mind, let us recall the standard tensorial transformation behaviour of the metric under coordinate transformations,

$$\bar{g}_{\alpha\beta}(y) = \frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial x^\nu}{\partial y^\beta} g_{\mu\nu}(x) \quad . \tag{4.60}$$

It follows from this that the *absolute value of the determinant of the metric*

$$g := |\det(g_{\mu\nu}(x))| \tag{4.61}$$

does not transform like a scalar or some other tensor at all, but instead transforms as

$$\bar{g} = \left|\det\left(\frac{\partial x}{\partial y}\right)^2\right| g = \left|\det\left(\frac{\partial y}{\partial x}\right)\right|^{-2} g \quad . \tag{4.62}$$

In particular, its square-root $\sqrt{g}$ transforms as

$$\sqrt{\bar{g}} = \left|\det\left(\frac{\partial y}{\partial x}\right)\right|^{-1} \sqrt{g} \quad . \tag{4.63}$$

Thus it satisfies (4.58) and provides us with our required reference object

$$\rho(x) = \sqrt{g(x)} \quad . \tag{4.64}$$

In particular, the combined expression $\sqrt{g}d^4x$ is *invariant* under general coordinate transformations,

$$\sqrt{\bar{g}}d^4y = \sqrt{g}d^4x \quad , \tag{4.65}$$

and can therefore be used to define integrals of scalars in a generally covariant (but metric-dependent) way,

$$\int \sqrt{\bar{g}}d^4y\ \bar{f}(y) = \int \sqrt{g}d^4x\ f(x) \quad . \tag{4.66}$$

This will of course be important in order to formulate action principles etc. in a space-time equipped with a metric in a generally covariant way.

<u>Remarks:</u>

1. This is also frequently the quickest way to determine the volume element in non-Cartesian coordinates in Euclidean space. In Cartesian coordinates one has $\sqrt{g} = 1$, so the correct integration measure is just the familiar

$$\sqrt{g}d^3x = d^3x \quad . \tag{4.67}$$

To now determine what is the volume element in spherical coordinates $\{y^k\} = (r,\theta,\phi)$, say, instead of laboriously determining the Jacobi matrix for the coordinate transformation, and then (equally laboriously) calculating its determinant (which would be the standard uninspiring and uninspired procedure), all one needs to know is the metric in these coordinates (which one usually needs to determine anyway) to deduce

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad \Rightarrow \quad \bar{g} = r^4\sin^2\theta \tag{4.68}$$

and therefore

$$d^3x = \sqrt{\bar{g}}\ d^3y = r^2\sin\theta\ dr\ d\theta\ d\phi \quad . \tag{4.69}$$

This is of course the standard result.

2. As a variation of this theme, it is now also easy to construct a spherical (i.e. radial + angular) coordinate system z^k in which $\sqrt{g} = 1$, i.e. which is such that simply $d^3z = d^3x$ (without any Jacobian factors). To that end it suffices to let (4.69) suggest to introduce a new radial variable $\rho = \rho(r)$ by

$$d\rho = r^2 dr \quad \Rightarrow \quad \rho = r^3/3 \in (0, \infty) \tag{4.70}$$

and a new angular variable $\psi = \psi(\theta)$ by

$$d\psi = \sin\theta d\theta \quad \Rightarrow \quad \psi = -\cos\theta \in [-1, 1) \ . \tag{4.71}$$

In these coordinates, the Euclidean line element takes the form

$$\begin{aligned} ds^2 &= r(\rho)^{-4} d\rho^2 + r(\rho)^2 (d\psi^2 / \sin^2\theta(\psi) + \sin^2\theta(\psi) d\phi^2) \\ &= (3\rho)^{-4/3} d\rho^2 + (3\rho)^{2/3} \left(\frac{d\psi^2}{1-\psi^2} + (1-\psi^2) d\phi^2 \right) \end{aligned} \tag{4.72}$$

It is now manifest that in these coordinates $\{z^k\} = (\rho, \psi, \phi)$ one has

$$\sqrt{g} = 1 \quad \Rightarrow \quad d^3x = d^3z = d\rho\, d\psi\, d\phi \ . \tag{4.73}$$

The Euclidean metric in these coordinates will make a brief appearance in the discussion of the derivation of the Schwarzschild metric in section 24.3 (cf. the discussion leading to (24.33)).

3. More generally, given any metric $g_{\alpha\beta}$, one can find a coordinate system in which $g = |\det(g_{\alpha\beta})| = 1$. Such a ("*unimodular*") coordinate system is highly non-unique. In fact, given any such coordinate system x, any other coordinate system y with unit Jacobi matrix $|\det(\partial y/\partial x)| = 1$ will also satisfy this condition. Such coordinate transformations form an infinite-dimensional subgroup of the group of all coordinate transformations, known as the group of *volume-preserving coordinate transformations*.

   In particular, the above transformation from Cartesian coordinates x^k to the coordinates z^k is a non-trivial example of a volume-preserving coordinate transformation.

4.5 Tensor Densities and Volume Elements

In the previous section we have encountered certain not strictly tensorial objects which nevertheless turned out to be useful. Having thus established the basic credentials of such objects, we will now formalise this somewhat.

Thus the prime example of what we will call a tensor density is the (absolute value of the) determinant $g := |\det g_{\mu\nu}|$ of the metric tensor, which, as we have seen, transforms as

$$\bar{g} = \left| \det \left(\frac{\partial y}{\partial x} \right) \right|^{-2} g \ . \tag{4.74}$$

An object which transforms in such a way under coordinate transformations is called a *scalar tensor density of weight* $w = +2$, and the square root of the determinant $\sqrt{g}$ transforms as, and hence is, a tensor density of weight $w = +1$.

In general, a tensor density of weight w is an object that transforms as a tensor, with an additional factor of $|\det(\partial y/\partial x)|^{-w}$. It thus transforms like a tensor multiplied by $g^{w/2}$, and one has

$$\tilde{T} \quad \text{a tensor density of weight } w \quad \Leftrightarrow \quad \tilde{T} = g^{w/2} T \quad \text{with } T \text{ a tensor} \quad . \tag{4.75}$$

The algebraic rules for tensor densities are strictly analogous to those for tensors. Thus, for example, the sum of two (p,q) tensor densities of weight w (let us call this a $(p,q;w)$ tensor) is again a $(p,q;w)$ tensor, and the direct product of a $(p_1,q_1;w_1)$ and a $(p_2,q_2;w_2)$ tensor is a $(p_1+p_2, q_1+q_2; w_1+w_2)$ tensor. Contractions and the raising and lowering of indices of tensor densities can also be defined just as for ordinary tensors.

REMARKS:

1. Generalising the argument in section 4.4, we now learn that if ρ is any scalar density of weight $w = +1$, then its integral is well-defined and coordinate independent,
$$\int d^4x \, \rho(x) = \int d^4y \, \bar{\rho}(y) \quad . \tag{4.76}$$
See remark 4 at the end of this section for one way of constructing such objects without taking recourse to a metric.

2. There is one important tensor density which - like the Kronecker tensor - has the same components in all coordinate systems. This is the totally anti-symmetric Levi-Civita symbol $\in_{\alpha\beta\gamma\delta}$ (taking the values $0, \pm 1$) which is a tensor density of weight $w = -1$. Then $\sqrt{g} \in_{\alpha\beta\gamma\delta}$ is a tensor (strictly speaking it is a pseudo-tensor because of its behaviour under reversal of orientation - see below).

   To see this, recall first of all the definition of the Levi-Civita symbol: it is totally anti-symmetric,
$$\in_{\alpha\beta\gamma\delta} = \in_{[\alpha\beta\gamma\delta]} \quad , \tag{4.77}$$
   and has therefore only got one independent component which we will normalise to be
$$\in_{0123} = +1 \quad . \tag{4.78}$$
   Thus $\in_{\alpha\beta\gamma\delta} = +1$ if the indices $(\alpha\beta\gamma\delta)$ are an even permutation of (0123), $\in_{\alpha\beta\gamma\delta} = -1$ if the indices $(\alpha\beta\gamma\delta)$ are an odd permutation of (0123), and $\in_{\alpha\beta\gamma\delta} = 0$ iff any two indices are equal. This definition makes no reference to any coordinate system whatsoever, and thus tautologically the purely combinatorial object $\in_{\alpha\beta\gamma\delta}$ has the same components in all coordinate systems.

This evidently extends to other dimensions, and we will define the $D = (d+1)$-dimensional Levi-Civita symbol in the same way,

$$\in_{\alpha_1\ldots\alpha_D} = \in_{[\alpha_1\ldots\alpha_D]} \quad , \quad \in_{01\ldots d} = +1 \ . \tag{4.79}$$

Next, recall one possible definition of the determinant $\det M$ of a $(D \times D)$-matrix $M^\mu_{\ \nu}$, namely as the coefficient (proportionality factor) on the right-hand side of

$$\in_{\alpha_1\ldots\alpha_D} \ M^{\alpha_1}_{\ \beta_1} \ldots M^{\alpha_D}_{\ \beta_D} = (\det M) \ \in_{\beta_1\ldots\beta_D} \quad . \tag{4.80}$$

Now choose M to be the Jacobi matrix $J^\mu_\alpha = (\partial y^\mu / \partial x^\alpha)$. Then the above equation can be rearranged to

$$\in_{\mu_1\ldots\mu_D} = \det\left(\frac{\partial y}{\partial x}\right) J^{\alpha_1}_{\mu_1} \ldots J^{\alpha_D}_{\mu_D} \in_{\alpha_1\ldots\alpha_D} \quad . \tag{4.81}$$

This shows that $\in_{\alpha_1\ldots\alpha_D}$ transforms as a tensor density of weight $w = -1$, provided that $\det(\partial y / \partial x) > 0$. The latter condition means that the coordinate transformation preserves the orientation. Thus, $\in_{\alpha_1\ldots\alpha_D}$ transforms as a tensor density under orientation-preserving coordinate transformations but picks up a sign when the orientation is reversed. Strictly speaking $\in_{\alpha_1\ldots\alpha_D}$ is then not a tensor density but something that is called a *pseudo-tensor density*.

Going back to 4 dimensions, it follows that

$$\epsilon_{\alpha\beta\gamma\delta} \equiv \sqrt{g} \in_{\alpha\beta\gamma\delta} \tag{4.82}$$

is a totally anti-symmetric $(0,4)$ (pseudo-)tensor. Likewise, the totally anti-symmetric symbol $\in^{\alpha\beta\gamma\delta}$ is a tensor density of weight $w = +1$ and

$$\epsilon^{\alpha\beta\gamma\delta} \equiv \frac{1}{\sqrt{g}} \in^{\alpha\beta\gamma\delta} \tag{4.83}$$

is a totally anti-symmetric $(4,0)$ (pseudo-)tensor. Here, as usual, we have raised the indices of the tensor on the left-hand side with the metric, and $\in^{\alpha\beta\gamma\delta}$ is totally anti-symmetric, with $\in^{0123} = -1$. The minus-sign arises because the contraction with the metrics on the left-hand side produces a factor of

$$\sqrt{g}\det(g^{\alpha\beta}) = \sqrt{g}(\det(g_{\alpha\beta}))^{-1} = -\frac{1}{\sqrt{g}} \quad . \tag{4.84}$$

We could have chosen to not absorb the minus sign into the definition of $\in^{\alpha\beta\gamma\delta}$, at the expense of an explicit minus sign on the right-hand side of (4.83). The convention we have adopted is more convenient, however, in particular since it is compatible with the standard practice in special relativity to (tacitly) identify $\epsilon_{abcd} = \in_{abcd}$, the minus sign arising from raising the indices on $\in_{abcd}$ with the Minkowski metric η^{ab} with $\eta^{00} = -1$, so that $\in^{0123} = - \in_{0123}$.

3. There is an intimate relation between the preceding observations regarding the Levi-Civita symbol (remark 2) and those in section 4.4 above regarding invariant volume elements. Namely, the usual coordinate volume element d^4x can be written as
$$d^4x = \frac{1}{4!} \in_{\alpha\beta\gamma\delta} dx^\alpha dx^\beta dx^\gamma dx^\delta \quad . \tag{4.85}$$
This is not a tensor but transforms like a scalar density. On the other hand, if one works instead with the tensor $\epsilon_{\lambda\mu\nu\rho}$ one obtains a scalar, and this scalar is precisely the invariant volume element (4.65),
$$\frac{1}{4!}\epsilon_{\alpha\beta\gamma\delta} dx^\alpha dx^\beta dx^\gamma dx^\delta = \sqrt{g} d^4x \quad . \tag{4.86}$$

4. More generally (and without invoking a metric to provide the weight $w = +1$ density required for an invariant integration) this can be phrased in the following way: Let $A_{\alpha\beta\gamma\delta}$ be a totally anti-symmetric $(0,4)$-tensor. Thus it will be proportional to the metric dependent Levi-Civita tensor $\epsilon_{\alpha\beta\gamma\delta}$, but we will now not make use of this fact (which would return us to the setting of the previous remark). Rather, we consider its contraction with the contravariant Levi-Civita symbol $\in^{\alpha\beta\gamma\delta}$ (which exists independently of any additional structure like a metric),
$$A_{\alpha\beta\gamma\delta} \to \tilde{A} = \in^{\alpha\beta\gamma\delta} A_{\alpha\beta\gamma\delta} \quad . \tag{4.87}$$
This is now clearly a scalar density of weight $w = +1$. As a consequence, its integral (4.76) is well-defined and coordinate independent, without reference to any metric. Thus totally anti-symmetric $(0,4)$-forms provide natural 4-dimensional volume elements (and likewise for totally anti-symmetric $(0,p)$ tensors and p-dimensional volume elements).

4.6 Towards a Coordinate-Independent Interpretation of Tensors

There is a more invariant and coordinate-independent way of looking at tensors than we have developed so far. The purpose of this section (and the subsequent section 4.7) is to briefly explain this point of view, even though it is not indispensable for an understanding of the remainder of the course.

Consider first of all the differential
$$df = \partial_\alpha f(x) dx^\alpha \tag{4.88}$$
of a function (scalar field) $f = f(x)$. This is clearly a coordinate-independent object, because partial deriviatives and the differentials dx^α transform inversely to each other,
$$dy^\mu = J^\mu_\alpha dx^\alpha \quad , \quad \partial_\mu = J^\alpha_\mu \partial_\alpha \quad \Rightarrow \quad \partial_\alpha f(x) dx^\alpha = \partial_\mu \bar{f}(y) dy^\mu \quad . \tag{4.89}$$

This suggests that it is useful to regard the quantities $\partial_\alpha f$ as the coefficients of the coordinate independent object *df* in a particular coordinate system, namely when *df* is expanded in the basis $\{dx^\alpha\}$.

We can do the same thing for any covector A_α. If A_α is a covector (i.e. transforms like one under coordinate transformations), then $A := A_\alpha(x)dx^\alpha$ is coordinate-independent, and it is useful to think of the A_α as the coefficients of the covector A when expanded in a coordinate basis, $A = A_\alpha dx^\alpha$. Linear combinations of dx^α built in this way from covectors are known as *1-forms.*

From this point of view, we interpret the $\{A_\alpha\}$ simply as the (coordinate dependent) components of the (coordinate independent) 1-form A when expressed with respect to the (coordinate dependent) differentials $\{dx^\alpha\}$, considered as a basis of the space of covectors.

Something similar can be done for vector fields. Just as covectors transform inversely to coordinate differentials, vectors V^α transform inversely to partial derivatives ∂_α. Thus

$$V := V^\alpha(x)\frac{\partial}{\partial x^\alpha} \tag{4.90}$$

is coordinate-independent - a coordinate-independent linear first-order differential operator. One can thus always think of a vector field as a 1st order differential operator and this is a very fruitful point of view.

Acting on a function (scalar) f, V produces the derivative of f along V,

$$Vf = V^\alpha \partial_\alpha f \quad . \tag{4.91}$$

This is also a coordinate independent object, a scalar, arising from the contraction of a vector and a covector. And this is as it should be because, after all, both a function and a vector field can be specified on a space-time without having to introduce coordinates (e.g. by simply drawing the vector field and the profile of the function). Therefore also the change of the function along a vector field should be coordinate independent and, as we have seen, it is.

So far we have only discussed vectors and covectors. All this can, in principle, be extended to higher rank tensors, but at this point it would be very useful to introduce the notion (or at least the notation) of tensor products. I will briefly describe this in section 4.7 below.

For those who do not want to delve into this (and it is not required for the following): fact of the matter is that any (p,q)-tensor $T^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}$ can be thought of as the collection of components of a coordinate independent object T when expanded in a particular coordinate basis in terms of the dx^α and $(\partial/\partial x^\alpha)$.

Any choice of coordinate system $\{x^\alpha\}$ gives rise to such a basis $\{dx^\alpha\}$, and such bases are known as *coordinate bases* or *natural bases*. This is not the only possible choice of basis, however, and we will return to this issue in section 4.8.

4.7 Multilinear Algebra and Tensors

In (multi-)linear algebra, the tensor product is used to describe multilinear maps. Let V be a vector space, and V^* its dual, consisting of the linear maps $V \to \mathbb{R}$, and denote the action of $a \in V^*$ on $v \in V$ by

$$a \in V^*, v \in V \quad \to \quad a(v) \in \mathbb{R} \ . \tag{4.92}$$

In components, with respect to a basis E_k in V and its dual basis e^k in V^*,

$$e^i(E_k) = \delta^i_k \ , \tag{4.93}$$

this would be written as

$$(a_i e^i)(v^k E_k) = a_i v^k e^i(E_k) = a_i v^k \delta^i_k = a_k v^k \ . \tag{4.94}$$

Then a bilinear map on the Cartesian product $V \times V$ can be considered as an element of $V^* \otimes V^*$, the tensor product being defined by

$$(a \otimes b)(v, w) = a(v)\, b(w) \ . \tag{4.95}$$

Note that this is not symmetric, i.e.

$$a \otimes b \neq b \otimes a \ . \tag{4.96}$$

Again in components this would read

$$(a \otimes b)(v, w) = a_i b_k v^i w^k \ , \tag{4.97}$$

and a general element of $V^* \otimes V^*$ can be written as an object with components a_{ik} (a covariant tensor),

$$a = a_{ik} e^i \otimes e^k \ , \tag{4.98}$$

acting as

$$a(v, w) = a_{ik} v^i w^k \ . \tag{4.99}$$

From these definitions it follows that the tensor product is evidently linear,

$$a \otimes (b + c) = a \otimes b + a \otimes c \tag{4.100}$$

(and likewise for the first factor), and $\mathbb{R}$-linear, i.e. for $r \in \mathbb{R}$ one has

$$r(a \otimes b) = (ra) \otimes b = a \otimes (rb) \ . \tag{4.101}$$

The $\mathbb{R}$-linearity is in a sense the characteristic feature of the tensor product $V \otimes W$ of two vector spaces (here $V^* \otimes V^*$) that sets it apart from the direct (Cartesian) product $V \times W$ of vector spaces (here $V^* \times V^*$), which consists of the pairs (v, w), and for which there is obviously no identification between (rv, w) and (v, rw) since these are just distinct points of the Cartesian product.

This can be straightforwardly extended to a description of general multilinear maps on vector spaces:

- Using the canonical isomorphism $(V^*)^* \cong V$ for finite-dimensional vector spaces,

$$v \in V \to \hat{v} \in (V^*)^* : \quad \hat{v}(a) = a(v) \ \ , \tag{4.102}$$

one can also in the same way define the tensor product $V \otimes V$ as the space of bilinear functions on $V^* \times V^*$,

$$(v \otimes w)(a, b) = a(v)\, b(w) \ \ , \tag{4.103}$$

and a general elelement of $V \otimes V$ can be represented in terms of its components T^{ik},

$$T = T^{ik} E_i \otimes E_k \ \ , \tag{4.104}$$

acting as

$$T(a, b) = T^{ik} a_i b_k \ \ . \tag{4.105}$$

- By the same token, the tensor product $V \otimes W$ is the space of bilinear maps on $V^* \otimes W^*$.

- Multilinear maps from $V \times \ldots \times V$ to $\mathbb{R}$ are elements of

$$\otimes^p V^* = \underbrace{V^* \otimes \ldots \otimes V^*}_{p \text{ times}} \tag{4.106}$$

(and can be represented as covariant tensors of rank p), etc.

- These multilinear maps can be added and multipled and thus form an algebra, the *tensor algebra* of V^*, denoted by $T(V^*)$. As a vector space, it consists of the sums of all the p-linear maps,

$$T(V^*) = \oplus_{p=0} \otimes^p V^* \ \ . \tag{4.107}$$

The tensor product can also be used to describe multilinear maps between vector spaces:

- An element $a \otimes v$ of $V^* \otimes V$ can be regarded as a linear map from V to itself via

$$(a \otimes v_1)(v_2) = a(v_2) v_1 \tag{4.108}$$

and a general element of $V^* \otimes V$ (a "matrix" $M^i_{\ k}$) can be written as a linear combination of such maps,

$$M = M^i_{\ k} E_i \otimes e^k : \quad M(v) = M^i_{\ k} e^k(v) E_i = (M^i_{\ k} v^k) E_i \ \ . \tag{4.109}$$

- Likewise a linear map from V to some other vector space W can be regarded as an element of $V^* \otimes W$.

- Multilinear maps from V to W ("W-valued multilinear maps") are elements of $V^* \otimes \ldots \otimes V^* \otimes W$, etc.

Clearly, in general, given a basis of V and a dual basis of V^*, the tensor product can be used to construct a basis

$$(E_{i_1} \otimes \ldots \otimes E_{i_p}) \otimes (e^{k_1} \otimes \ldots e^{k_q}) \tag{4.110}$$

in the space

$$T^{p,q} = \underbrace{(V \otimes \ldots \otimes V)}_{p \text{ times}} \otimes \underbrace{(V^* \otimes \ldots \otimes V^*)}_{q \text{ times}} \tag{4.111}$$

of (p, q)-tensors,

$$T \in T^{p,q} : T = T^{i_1 \ldots i_p}_{k_1 \ldots k_q} \, (E_{i_1} \otimes \ldots \otimes E_{i_p}) \otimes (e^{k_1} \otimes \ldots \otimes e^{k_q}) \ . \tag{4.112}$$

This is the way we will use the tensor product notation below, as a multilinear operation providing us with a basis for higher rank tensor fields.

Now, as we have seen above, in the standard component/index formulation of general relativity, say, a (p, q)-tensor is defined as an object with components $T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q}$ which transforms multi-linearly with the Jacobi matrix under coordinate transformations, i.e. under $x^\mu \to y^\alpha$ one has

$$T^{\alpha_1 \ldots \alpha_p}_{\beta_1 \ldots \beta_q}(y) = J^{\alpha_1}_{\mu_1} \ldots J^{\alpha_p}_{\mu_p} \, J^{\nu_1}_{\beta_1} \ldots J^{\nu_q}_{\beta_q} \, T^{\mu_1 \ldots \mu_p}_{\nu_1 \ldots \nu_q}(x(y)) \tag{4.113}$$

where $J^\alpha_\mu = \partial y^\alpha / \partial x^\mu$ is the Jacobi matrix and $J^\mu_\alpha = \partial x^\mu / \partial y^\alpha$ is its inverse.

The reason for introducing and working with tensors, defined in this way, is that tensorial equations have the virtue that they are *generally covariant*, i.e. that they are satisfied in all coordinate system if and only if they are satisfied in one coordinate system. The emphasis in this formulation is thus not on tensors as multilinear maps but on how they transform under coordinate transformations. This seems to be somewhat at odds with the definition of tensors in multilinear algebra, but as we will see below this is simply due to the choice of a particular class of bases (coordinate bases), with respect to which multilinear maps indeed transform in this way under changes of the coordinate basis, i.e. under changes of coordinates.

We had already noted above, that there is a more coordinate independent way of looking at covector fields and vector fields, by associating to them the objects

$$A_\mu(x) \to A(x) = A_\mu(x) dx^\mu \quad , \quad V^\alpha(x) \to V(x) = V^\alpha(x) \partial_\alpha \ . \tag{4.114}$$

which are completely invariant under coordinate transformations, with the dx^μ and the ∂_α providing a basis for the space of covector and vector fields respectively.

This perspective can now be extended to higher-rank and mixed tensors. In particular, associated with the metric $g_{\mu\nu}(x)$ we have the coordinate independent line element

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \ . \tag{4.115}$$

which we can now also think of as the tensor

$$g = g_{\mu\nu} dx^\mu \otimes dx^\nu \quad . \tag{4.116}$$

where $\otimes$ is now again the *tensor product*.

Since we are now dealing with tensor fields rather than just with tensors (multilinear maps at a given point), the tensor product in this context is required to be multilinear not just over $\mathbb{R}$, but over functions (scalars) so that e.g.

$$dx^\mu \otimes (f(x)dx^\nu) = (f(x)dx^\mu) \otimes dx^\nu \quad . \tag{4.117}$$

Now let us return to (4.116). If one wants to emphasise that the metric is a symmetric (0,2)-tensor, one can also expand it with respect to the symmetrised basis as

$$g = g_{\mu\nu} \, (dx^\mu \otimes dx^\nu + dx^\nu \otimes dx^\mu)/2 \quad , \tag{4.118}$$

but for the metric the tensor-product is often omitted and one simply writes it as the line element (4.115).

If one has a non-symmetric $(0,2)$-tensor $T_{\mu\nu}$, say, then one can also group these coefficients into the components of a coordinate-invariant object, but now the tensor product notation

$$T_{\mu\nu} \to T = T_{\mu\nu} dx^\mu \otimes dx^\nu \tag{4.119}$$

is more useful than just writing $T_{\mu\nu}dx^\mu dx^\nu$, simply to emphasise the fact that all components of $T_{\mu\nu}$, not just the symmetric part of $T_{\mu\nu}$, contribute to T because $dx^\mu \otimes dx^\nu$ is not symmetric,

$$dx^\mu \otimes dx^\nu \neq dx^\nu \otimes dx^\mu \quad , \tag{4.120}$$

(whereas just writing $dx^\mu dx^\nu$ might lead one to believe that dx^μ and dx^ν commute).

More generally, to a $(0,p)$-tensor we can associate the object

$$T = T_{\mu_1 \ldots \mu_p} dx^{\mu_1} \otimes \ldots \otimes dx^{\mu_p} \quad . \tag{4.121}$$

If $T_{\mu_1 \cdots \mu_p}$ is totally anti-symmetric, the resulting object is also referred to as a *p-form*. As we have already seen above, such p-forms provide natural and invariant p-dimensional volume elements. In particular, applying this to the Levi-Civita tensor discussed in section 4.5, we reproduce the statement (4.65) that (4.86)

$$\frac{1}{4!}\epsilon_{\alpha\beta\gamma\delta} dx^\alpha dx^\beta dx^\gamma dx^\delta = \sqrt{g} d^4x \tag{4.122}$$

is a space-time volume element that is invariant under coordinate transformations.

The tensor product notation is also useful for higher-rank contravariant or mixed tensors. Given a $(2,0)$-tensor with components $T^{\mu\nu}$, say, one really does not want to write the corresponding coordinate-invariant object as $T^{\mu\nu}\partial_\mu\partial_\nu$, say, because this may be interpreted as a second order differential operator whereas what one really means is a bilinear first order differential operator, which one writes as

$$T = T^{\mu\nu}\, \partial_\mu \otimes \partial_\nu \quad , \tag{4.123}$$

and whose components with respect to the basis $\partial_\mu \otimes \partial_\nu$ are the $T^{\mu\nu}$.

In general, we can thus think of a (p, q)-tensor field, as given in (4.113), as the components of a coordinate-independent object

$$T = T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}(x)\, (\partial_{\mu_1} \otimes \ldots \partial_{\mu_p}) \otimes (dx^{\nu_1} \otimes \ldots \otimes dx^{\nu_q}) \quad , \tag{4.124}$$

when expanded with respect to the coordinate basis in the space of tensor fields generated by dx^μ and $\partial_\mu = \partial_{x^\mu}$.

4.8 Vielbeins and Orthonormal Frames

As we saw in section 4.6, a choice of coordinates provides one with a choice of basis for vectors, covectors and other tensors, and a quantity like V^μ is then interpreted as the collection of components of an object $V = V^\mu \partial_\mu$ with respect to the coordinate basis ∂_μ. In classical tensor calculus one always works in such a basis, and with the components of tensors with respect to such a basis. This is very convenient and natural, but this is now clearly not the only choice.

Indeed, the above point of view suggests a reformulation and generalisation that is extremely natural and useful (but that I will nevertheless hardly ever make use of in these notes).

Namley, let $\{e^m_{\ \mu}(x)\}$ be such that it is an invertible matrix for every point x. Then another possible choice of basis for the space of covectors are the linear combinations

$$e^m := e^m_{\ \mu} dx^\mu \quad . \tag{4.125}$$

A general such basis is called a *vielbein*, which is German for *multileg*, quite appropriate actually, as one should visualise this as a bunch of linearly independent (co-)vectors at every point of space-time.

In two, three, and four dimensions these are also known more specifically as *zweibeins, dreibeins* and *vierbeins* respectively. In four dimensions, the Greek word *tetrads* is also commonly used. The e^m are sometimes also referred to as *frame fields*, mostly in the context of *orthonormal frames* (see below).

In general, this new basis is *not* a coordinate basis, i.e. there does not exist a coordinate system $\{y^m\}$ such that $e^m = dy^m$. If such a coordinate system does exist, then one has

$$\begin{aligned} e^m = dy^m \quad &\Rightarrow \quad e^m_{\ \mu} = \frac{\partial y^m}{\partial x^\mu} \\ &\Rightarrow \quad \partial_\nu e^m_{\ \mu} = \partial_\mu e^m_{\ \nu} \quad , \end{aligned} \tag{4.126}$$

and locally also the converse is true. In particular, if

$$\partial_\nu e^m_{\ \mu} - \partial_\mu e^m_{\ \nu} \neq 0 \quad \Rightarrow \quad e^m \quad \text{is not a coordinate basis} \ . \tag{4.127}$$

For many purposes, bases other than coordinate bases can also be extremely useful and natural, in particular the orthonormal bases we will introduce below.

The inverse relation to (4.125) is

$$dx^\mu = e^\mu_{\ m} e^m \ , \tag{4.128}$$

where $e^\mu_{\ m}(x)$ is (pointwise) the inverse matrix of $e^m_{\ \mu}(x)$,

$$e^m_{\ \mu} e^\mu_{\ n} = \delta^m_n \qquad e^\mu_{\ m} e^m_{\ \nu} = \delta^\mu_\nu \ . \tag{4.129}$$

With respect to this basis, one can expand a covector A as

$$A = A_\mu dx^\mu = A_\mu e^\mu_{\ m} e^m \equiv A_m e^m \ , \tag{4.130}$$

so that the components of A with respect to the new basis $\{e^m\}$ are

$$A_m = A_\mu e^\mu_{\ m} \ . \tag{4.131}$$

Likewise, the vielbeins allow us to pass from a natural (or coordinate) basis for vector fields, the $\{\partial_\mu\}$, to another basis

$$E_m = e^\mu_{\ m} \partial_\mu \ , \tag{4.132}$$

allowing us to write the coordinate independent vector field

$$V = V^\mu(x)\partial_\mu = V^m E_m \tag{4.133}$$

with

$$V^m = e^m_{\ \mu} V^\mu \ . \tag{4.134}$$

Note that, unlike the ∂_μ, the E_m do not commute in general, i.e.

$$[E_m, E_n] \neq 0 \ . \tag{4.135}$$

In fact, a 'dual' characterization of a coordinate basis is that the corresponding E_m do commute. This is clearly a necessary condition and, as above, locally it is also sufficient to ensure that there is a coordinate system y^m such that $E_m = \partial/\partial y^m$.

We can apply the same reasoning to any other tensor field, e.g. to the metric tensor itself. We can write the invariant line element as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\mu = g_{\mu\nu} e^\mu_{\ m} e^\nu_{\ n} e^m e^n \equiv g_{mn} e^m e^n \ , \tag{4.136}$$

so that the components of the metric with respect to the new basis are

$$g_{mn} = g_{\mu\nu} e^\mu_{\ m} e^\nu_{\ n} \ . \tag{4.137}$$

Given a metric, there is a preferred class of bases $\{e^a\}$ which are such that the corresponding matrices $e^a_{\ \mu}(x)$ diagonalise (and normalise) the metric at every point x, i.e. which are such that $g_{ab} = \eta_{ab}$ or

$$g_{ab} = \eta_{ab} \quad \Leftrightarrow \quad g_{\mu\nu}(x) e^\mu_{\ a}(x) e^\nu_{\ b}(x) = \eta_{ab} \ . \tag{4.138}$$

Such a basis e^a, with respect to which the components of the metric are the Minkowski metric η_{ab}, is known as an *orthonormal basis* or *orthonormal frame*.

In the more mathematical literature, the e^a are also referred to as *soldering forms* because they identify (solder, glue) an abstract space of (co-)vectors at each point x, labelled by $a, b, \ldots$ with the concrete space of (co-)vectors tangent to the space-time at the point x, labelled e.g. by the indices $\mu, \nu, \ldots$.

For a general metric, a basis which achieves this cannot be a coordinate basis (because this would mean that the metric is equivalent to the Minkowski metric by a coordinate transformation). However, clearly there is no obstacle to finding a more general basis which will do this: for every point x we can find a matrix $e^a_{\ \mu}(x)$ which achieves (4.138) As the metric varies smoothly with x, we can also choose the matrices $e^a_{\ \mu}(x)$ to vary smoothly with x, and hence we can put them together to define the smooth matrix-valued function $e^a_{\ \mu}(x)$ for all x. [I am ignoring some global (topological) issues here. We will not need to worry about them here.]

The reason why I referred to a "class of bases" above is that, clearly, such an orthonormal basis is not unique. At every point x it is determined up to a Lorentz transformation

$$\begin{aligned} & e^a(x) \to \Lambda^a_{\ b}(x) e^b(x) \\ & \Lambda^a_{\ b}(x) \Lambda^c_{\ d}(x) \eta_{ac} = \eta_{bd} \ . \end{aligned} \tag{4.139}$$

Thus a given metric does not determine a unique orthonormal basis, but only an orthonormal basis up to Lorentz transformations

$$e^a(x) \to \Lambda^a_{\ b}(x) e^b(x) \ . \tag{4.140}$$

Conversely, however, an orthonormal basis uniquely determines a metric via

$$ds^2 = \eta_{ab} e^a(x) e^b(x) \ . \tag{4.141}$$

If one wants the components of the metric in a given coordinate system $\{x^\mu\}$, one expands the orthonormal basis e^a in terms of the natural basis dx^μ as above as

$$e^a(x) = e^a_{\ \mu}(x) dx^\mu \ , \tag{4.142}$$

to find, as above,

$$g_{\mu\nu}(x) = e^a_{\ \mu}(x) e^b_{\ \nu}(x) \eta_{ab} \ . \tag{4.143}$$

Thus instead of the metric one can choose orthonormal vielbeins as the basic variables of General Relativity. In that case one has to demand not only general covariance but also invariance under local Lorentz transformations (acting on the orthonormal indices $a, b, \ldots$). [One could also allow for general vielbeins, in which case one would have to replace Lorentz transformations by the larger group of general linear transformations.]

<u>EXAMPLES:</u>

Here are a few examples to illustrate that orthonormal frames are not something mysterious but can usually be read off very easily from the metric in a coordinate basis.

1. The 2-Sphere Metric (2.16)

   The standard metric on a sphere of radius R is

   $$ds^2 = R^2(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{4.144}$$

   Now define

   $$e^1 = Rd\theta \quad , \quad e^2 = R\sin\theta d\phi \quad , \tag{4.145}$$

   i.e.

   $$e^a = e^a{}_\alpha dx^\alpha \tag{4.146}$$

   with

   $$e^1_\theta = R \ , \ e^1_\phi = 0 \ , \ e^2_\theta = 0 \ , \ e^2_\phi = R\sin\theta \quad . \tag{4.147}$$

   Then the metric can be written as

   $$ds^2 = e^1e^1 + e^2e^2 = \delta_{ab}e^ae^b \quad , \tag{4.148}$$

   so the e^a are an orthonormal basis. They are obviously not a coordinate basis because (4.127)

   $$\partial_\theta e^2_\phi = R\cos\theta \neq \partial_\phi e^2_\theta = 0 \quad . \tag{4.149}$$

   Likewise, we can introduce an orthonormal basis

   $$E_a = E^\alpha_a \partial_\alpha \tag{4.150}$$

   for vectors. A simple choice is

   $$E_1 = R^{-1}\partial_\theta \quad , \quad E_2 = (R\sin\theta)^{-1}\partial_\phi \quad , \tag{4.151}$$

   which satisfies

   $$g_{\alpha\beta}E^\alpha_a E^\beta_b = \delta_{ab} \quad . \tag{4.152}$$

   That this is not a coordinate basis is reflected in the fact that the commutator $[E_1, E_2] \neq 0$,

   $$[E_1, E_2] = R^{-2}(\partial_\theta(\sin\theta)^{-1})\partial_\phi = -R^{-1}\cot\theta \ E_2 \quad . \tag{4.153}$$

2. The Schwarzschild Metric (2.38)

   The metric is

$$ds^2 = -(1-2m/r)dt^2 + (1-2m/r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{4.154}$$

   With

$$(e^0, e^1, e^2, e^3) = \left((1-\frac{2m}{r})^{1/2}dt, (1-\frac{2m}{r})^{-1/2}dr, rd\theta, r\sin\theta d\phi\right) \tag{4.155}$$

   the metric can be written as

$$ds^2 = \eta_{ab}e^a e^b \quad , \tag{4.156}$$

   so the e^a are an orthonormal basis for the Schwarzschild metric.

3. The Kaluza-Klein Metric (section 44)

   Here is an example of a non-diagonal metric. The five-dimensional Kaluza-Klein metric is

$$d\hat{s}^2_{KK} = g_{\mu\nu}dx^\mu dx^\nu + (dx^5 + A_\mu dx^\mu)^2 \quad . \tag{4.157}$$

   Let $e^a_{\ \mu}$, $a = 0,1,2,3$, be vielbeins (tetrads) for the four-dimensional space-time metric $g_{\mu\nu}$. Then an orthonormal frame $\hat{e}^A$ for the Kaluza-Klein metric is

$$\hat{e}^a = e^a_{\ \mu}dx^\mu \quad , \quad \hat{e}^5 = dx^5 + A_\mu dx^\mu \quad . \tag{4.158}$$

Remarks:

1. Consider the timelike trajectory $x^\mu = x^\nu(\tau)$ of an observer, parametrised by proper time τ. Then his 4-velocity $u^\mu = dx^\mu/d\tau$ satisfies (2.55)

$$g_{\mu\nu}u^\mu u^\nu = -1 \quad . \tag{4.159}$$

   Recalling the defining relation for an orthonormal frame,

$$g_{\mu\nu}e^\mu_{\ a}e^\nu_{\ b} = \eta_{ab} \quad \Rightarrow \quad g_{\mu\nu}e^\mu_{\ a=0}e^\nu_{\ b=0} = \eta_{00} = -1 \quad , \tag{4.160}$$

   we see that the 4-velocity u^μ can be interpreted as the timelike component $e^\mu_{a=0}$ of an orthonormal frame along the worldline,

$$u^\mu = e^\mu_{\ a=0} \quad , \tag{4.161}$$

   with the spacelike components providing an orthonormal laboratory reference system. At this point there is still considerable freeedom in the choice of the spatial components of the orthonormal frame. This freedom can be significantly reduced (to rigid τ-independent rotations) by adopting a particular "parallel transport" condition of these vectors along the worldline, such as the Fermi-Walker parallel transport to be discussed in section 5.10.

In any case, however the laboratory system is defined, the frame components

$$V^a = e^a_{\ \mu} V^\mu \tag{4.162}$$

of a vector now acquire the physical interpretation as the components of V as measured with respect to the observer's proper time and his laboratory frame. Note that this generalises the fact, already used in our second derivation of the gravitational redshift in section 3.5, that the frequency of a wave with wave vector k^μ as measured by an observer with 4-velocity u^μ is (3.117)

$$\omega = -u^\mu k_\mu = -e^\mu_{\ a=0} k_\mu = e^{a=0}_{\ \mu} k^\mu \equiv k^{a=0} \quad . \tag{4.163}$$

2. The $e^a_{\ \mu}$ can in some sense be regarded as the square-root of the metric. In particular denoting the determinant of the matrix $e^a_{\ \mu}$ by

$$e(x) := \det(e^a_{\ \mu}(x)) \quad , \tag{4.164}$$

(4.138) implies

$$g(x) := |\det(g_{\alpha\beta}(x))| = e(x)^2 \quad \Leftrightarrow \quad |e(x)| = \sqrt{g(x)} \quad . \tag{4.165}$$

3. Coordinate indices can, as usual, be raised and lowered with the space-time metric $g_{\mu\nu}$ and its inverse, and Minkowski (tangent space) indices with the Minkowski metric η_{ab} and its inverse.

Note that this is consistent with the notation for $e^a_{\ \mu}$ and its inverse $e^\mu_{\ a}$ because

$$e^\mu_{\ a} = g^{\mu\nu}\eta_{ab} e^b_{\ \nu} \quad . \tag{4.166}$$

One also has other fairly evident relations like

$$g^{\mu\nu} = \eta^{ab} e^\mu_{\ a} e^\nu_{\ b} \quad , \tag{4.167}$$

etc. The reason why I have called the basis of vector fields in a general frame E_m rather than e_m is that e^m and E_m are of course not related just by lowering or raising the indices of the metric, $E_m \neq g_{mn} e^n$. The former are linear combinations of the dx^μ, the latter linear combinations of the ∂_μ, so they are very different objects.

One could now go ahead and develop the entire machinery of tensor calculus (covariant derivatives, curvature, ...) that we are about to develop in the following sections in terms of vielbeins as the basic variables instead of the metric. This is rather straightforward. For example, given the expression for the Christoffel symbols in terms of the metric, and for the metric in terms of the vielbeins, one can express the Christoffel symbols (and hence covariant derivatives and curvatures) in terms of vielbeins, but the resulting expressions are rather unenlightning and not of much use in practice.

The real power of the vielbein formalism emerges when one combines it with the formalism of *differential forms*. And in practice the most useful and efficient alternative to working in components in a coordinate basis is working with differential forms in an orthonormal basis.

I do most of my (curvature) calculations in the latter framework (and e.g. only then translate them into coordinate components for the purposes of inserting them into these notes), but this is (for the time being) not something I will develop further here.[10]

4.9 Epilogue: Indices? Indices!

Having reached this point, you may have the impression that the notation we have introduced for tensors, $T^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}$ say, and which, as you might have noticed by looking ahead, we will continue to use in these notes, with its morass of indices, is somewhat cumbersome and unelegant. And perhaps you might prefer to at the very least see everything written in terms of the index-free coordinate-invariant objects like $V = V^\mu\partial_\mu$ or $A = A_\mu dx^\mu$ introduced in section 4.6.

I cannot disagree with the sentiment that using all these indices does not appear to be particularly elegant. Mathematicians abhor it. Physicists, however, are pragmatists by nature - they will use whatever turns out to be useful or efficient for what they want to achieve, regardless of whether or not it is considered or perceived to be beautiful or elegant according to some external criteria.

In particular, in the case at hand, the index-laden notation would not be that commonly used and widespread if it did not have some distinct advantages over other options. Indeed, this notation is an extremely useful and informative bookkeeping device that conveys a lot of information in a very compact way. In particular, as we have seen, the index notation allows one to reliably read off what kind of tensor one is dealing with, along the lines of “if it has p upper and q lower indices, it transform like, hence is, a (p,q)-tensor”. Moreover, as we will see below, it provides one with a much more concise and informative way of describing and performing algebraic manipulations of tensors than some index-free notation is capable of.

Let me first make clear what the issue is and what it is not when one writes something like V^μ or $V^\mu(x)$, as this can be interpreted in (at least) 2 distinct ways:

[10] See e.g. W. Thirring, *Classical Mathematical Physics* for a presentation of general relativity entirely in the coordinate-independent formalism of differential forms, and N. Straumann, *General Relativity*, where differential forms are used whenever it is convenient or useful (and also occasionally when it is not . . .).

1. On the one hand, V^μ may refer to the numerical values of the components of a specific vector V in a specific coordinate system.

2. On the other hand, the notation V^μ may be used to indicate that the object V transforms like a vector.

The first use of V^μ is completely uncontentious: if one wants to write down the components of some object with respect to some basis, one has to write down the components of that object with respect to that basis, there is no way around that.

It is mainly the second use and interpretation of the notation that is at stake, and it is also mainly in this sense that the index notation is used for tensor algebra and tensor calculus in general and in these notes in particular.

To a somewhat lesser extent the fact that the notation itself does not indicate whether one has in mind the first or the second interpretation is also an issue (even though this is usually clear from the context). It is actually not so much an issue (if desired this is something that can easily be remedied - I will come back to this at the end of this section) as possibly the source of a major misunderstanding between mathematicians and physicists - namely that a dislike of the index notation arises from the (false!) belief that it means that one is always writing down objects with respect to a particular basis. If this were the case, this would indeed be clumsy and silly, and quite contrary to the spirit of general covariance. However, as interpretation 2 indicates, this is absolutely not what is meant.

Returning to the use of indices as a way to indicate tensorial type and tensorial operations (like contractions), let us consider the alternatives. If one wants to indicate in symbols that some object V is a vector field, then as a mathematician one might write something like $V \in \Gamma(TM)$, stating that V is a section of the tangent bundle of the space or space-time (manifold) M. This is fine, but if the space M is clear from the context, why not declare once and for all that writing V^μ means the same thing? And perhaps use different kinds of indices to refer to tensors on different spaces?

If this were all then this would hardly be an issue and even physicists could be convinced to write "$V \in \Gamma(TM)$", at least when talking to mathematicians. Where the index notation really pays off, however, is when it comes to algebraic manipulations such as those discussed in section 4.3 (and even more so when it comes to tensor analysis, which is the subject of section 5, but tensor algebra will be enough to illustrate this).

As examples consider the contractions of a $(1,2)$ tensor T, say, with itself and with a vector V. With indices one would write $T^\mu_{\ \nu\lambda}$ and V^μ and the possible contractions would be written as

$$
\begin{array}{ccccc}
T^\mu_{\ \nu\lambda} & \rightarrow & T^\mu_{\ \mu\lambda} & , & T^\mu_{\ \nu\mu} \\
(T^\mu_{\ \nu\lambda}, V^\mu) & \rightarrow & T^\mu_{\ \nu\lambda} V^\nu & , & T^\mu_{\ \nu\lambda} V^\lambda \ ,
\end{array}
\tag{4.168}
$$

the first line indicating the two distinct covectors one obtains as contractions of T itself, and the second the two distinct possibilities of contracting T and V to obtain a $(1,1)$-tensor. In an index-free notation one would have to invent some operation like C^m_n to indicate a contraction over the m'th upper and n'th lower index.[11] In this notation, the four objects above would then be written as

$$\begin{array}{rcl} T & \rightarrow & C^1_1(T) \quad , \quad C^1_2(T) \\ (T,V) & \rightarrow & C^2_1(T\otimes V) \quad , \quad C^2_2(T\otimes V) \ \ . \end{array} \tag{4.169}$$

Is this superior? It does not even allow one to read off the tensor type of the resulting objects unless one remembers what the tensor types of T and V were to begin with, whereas this is completely manifest in (4.168).

Moreover, imagine how untransparent this would become were one to perform even the simplest sequence of such elementary operations: compare

$$A_{\alpha\beta}V^\alpha W^\beta \quad \longleftrightarrow \quad C^1_1 C^2_2(A\otimes V\otimes W) \ \ . \tag{4.170}$$

If you prefer the right-hand side, or some variant of it, feel free to use it. However, you should be aware of the fact that the left-hand side contains an equivalent amount of information, simply packaged in a more digestible way that is both more informative ("it's a scalar!") and easier to manipulate. For most intents and purposes the index notation is really extremely convenient and it is for this reason that we will continue to make use of it in these notes.

One other reason for concern may be that by exclusively working with local coordinates and coordinate bases one may be missing some global aspects of a space or space-time. This is certainly true to a certain extent but is not primarily a notational issue. Rather, it means that in addition one needs to make use of more advanced notions from topology, global analysis etc. This is not something I will attempt here (cf. the book by Hawking and Ellis in the previous footnote for a description of the groundbreaking early applications of global analysis to general relativity). One related, but more elementary, issue is the introduction and use of the term *manifold* when referring to spaces or space-times of the kind we are dealing with in these notes. This is something I will very briefly come back to in section 5.11 below.

Let me, to conclude this ~~rant~~ section, come back to the issue of the notational ambiguity when one writes something like V^μ, which can occasionally be a source of confusion. Even though, as mentioned above, usually it is clear from the context what one means, one might imagine wanting to write down a couple of equations with indices which are only valid in spherical coordinates, say, and are therefore not to be understood as tensorial equations. Then it might be helpful to have a notation which reveals that information as well.

[11] I am not making this up - see e.g. section 2.2 of *The large scale structure of space-time* by S. Hawking and G. Ellis, in all other respects a wonderful book.

This can for instance be accomplished by inventing a new notation like $\stackrel{*}{=}$ (or whatever) to indicate an equality only in a special or specified coordinate system, but while this may add clarity it does not address the fundamental issue that just writing V^α does not unambiguously specify what one has in mind.

Alternatively, and more elegantly and attractively, this can e.g. be accomplished with very little effort with the help of what is known as the *Penrose abstract index notation*. The idea is to still indicate the tensor type of an object by a certain kind of indices, but with these indices only serving that purpose and not simultaneously referring to any particular kind of basis. Thus for example, one would indicate a vector by an object V^a, where the fact that one has a single upper index a just means that this is a $(1,0)$-tensor, and nothing else (exactly as in interpretation 2 above). For the components of this vector with respect to some basis (coordinates x^μ) one could then continue to use the traditional V^μ.

The advantage of this "abstract index" notation is that for tensorial operations one never needs to specify a basis anyway, so they can all be performed at the level of the abstract indices and tensorial equations look identical when written with these abstract indices or when written with concrete component indices. Thus $V^a W_a$ is used to indicate the scalar one obtains by contraction of a vector V^a with a covector W_a. Likewise, instead of $T^\mu_{\ \mu\lambda}$ (which may look basis dependent) one would write $T^a_{\ ab}$, and this is completely equivalent to writing something like $C^1_1(T)$,

$$T^a_{\ ab} \quad \longleftrightarrow \quad C^1_1(T) \tag{4.171}$$

but much more informative and user-friendly, and all the usual rules of tensor algebra apply to these abstract indices.

Whenever one wants or needs to specify a basis or coordinate system, this can be accomplished by using other kinds of indices. Thus g_{ab} could e.g. be used to refer to the metric tensor in general, while $g_{\mu\nu}$ could then be used to refer to its components in the basis x^μ. From this we see that

> "[...] the distinction between the index notation and the component notation is much more one of spirit (i.e., how one thinks of the quantities appearing) than of substance (i.e., the physical form the equations take)."[12]

While I will not make use of the abstract index notation in these notes (with the hope that this will not cause any confusion), the use of abstract indices appears to be an ideal ("eat the cake and have it too") compromise combining the best of both worlds

[12] R. Wald, *General Relativity.* See section 2.4 of this book for a more detailed explanation of the abstract index notation, which is systematically used throughout the book. For a detailed treatment of the abstract index notation and a discussion of some minor subtleties with this notation see R. Penrose, W. Rindler, *Spinors and Space-Time*, Vol. 1: *Two-Spinor Calculus and Relativistic Fields.*

and should actually keep both camps happy. It does not yet appear to have found widespread acceptance among mathematicians, however.

An alternative compromise solution is the already mentioned use of differential forms (in an orthonormal basis, say), which is manifestly covariant and minimises clutter, displaying only the (essential and informative) Lorentz Lie algebra indices while suppressing the component indices of forms (anti-symmetric tensors).

5 Tensor Analysis (Generally Covariant Differentiation)

Tensors transform in a nice and simple way under general coordinate transformations. Thus these appear to be the right objects to construct equations from that satisfy the Principle of General Covariance.

However, the laws of physics are differential equations, so we need to know how to differentiate tensors. This is not an issue of particular concern in Special Relativity, because (cf. (1.49)) the partial derivative of a Lorentz tensor

$$T^{a_1\ldots a_p}_{c_1\ldots c_q}(\xi) \quad \rightarrow \quad \partial_a T^{a_1\ldots a_p}_{c_1\ldots c_q}(\xi) \tag{5.1}$$

is again a Lorentz tensor. This relies on 2 facts, namely first that the partial derivative transforms as a covector under Lorentz transformations and secondly that the associated Jacobi matrix of Lorentz transformations is constant.

The former generalises to arbitrary coordinate transformations and implies, in particular, that the partial derivative of a scalar field is a covector field (4.18). However, because in general the Jacobi matrix is not constant, the ordinary partial derivative does not map tensors to tensors.

This is easy to see: take for example a vector V^μ. Under a coordinate transformation $x^\mu \rightarrow y^\alpha$, its partial derivative transforms as

$$\begin{aligned}
\partial_\beta V^\alpha &= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial}{\partial x^\nu}\frac{\partial y^\alpha}{\partial x^\mu}V^\mu \\
&= \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial y^\alpha}{\partial x^\mu}\partial_\nu V^\mu + \frac{\partial x^\nu}{\partial y^\beta}\frac{\partial^2 y^\alpha}{\partial x^\mu \partial x^\nu}V^\mu \quad .
\end{aligned} \tag{5.2}$$

The appearance of the second term shows that the partial derivative of a vector is not a tensor.

As the second term is zero for linear transformations, you see that partial derivatives transform in a tensorial way e.g. under Lorentz transformations, so that partial derivatives are all one usually needs in special relativity.

We also see that the lack of covariance of the partial derivative is very similar to the lack of covariance of the equation $\ddot{x}^\mu = 0$, and this suggests that the problem can be cured in the same way - by introducing Christoffel symbols. This is indeed the case.

5.1 Covariant Derivative for Vector Fields

Let us define the *covariant derivative* $\nabla_\nu V^\mu$ of a vector field V^μ by

$$\nabla_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\ \nu\lambda} V^\lambda \quad . \tag{5.3}$$

It follows from the non-tensorial behaviour (2.78), (2.79) of the Christoffel symbols under coordinate transformations $x^\mu \to y^\alpha$ that $\nabla_\nu V^\mu$, as defined above, is indeed a $(1,1)$ tensor.

In order to establish this, we transform

$$\nabla_\beta V^\alpha = \partial_\beta V^\alpha + \Gamma^\alpha_{\beta\gamma} V^\gamma \tag{5.4}$$

to x-coordinates, using the tensorial transformation behaviour of ∂_α and V^α and the non-tensorial transformation behaviour (2.79) of the Christoffel symbols, to arrive at

$$\nabla_\beta V^\alpha = J^\mu_\beta \partial_\mu (J^\alpha_\rho V^\rho) + (J^\alpha_\mu J^\nu_\beta J^\lambda_\gamma \Gamma^\mu{}_{\nu\lambda} + J^\alpha_\mu \partial_\beta J^\mu_\gamma) J^\gamma_\rho V^\rho \ . \tag{5.5}$$

The obstructions to tensoriality are the 2 terms involving the derivatives of the Jacobi matrix, but these cooperatively combine to give

$$\begin{aligned} J^\mu_\beta \partial_\mu J^\alpha_\rho + J^\alpha_\mu (\partial_\beta J^\mu_\gamma) J^\gamma_\rho &= J^\mu_\beta \partial_\rho J^\alpha_\mu + J^\alpha_\mu (\partial_\gamma J^\mu_\beta) J^\gamma_\rho \\ &= J^\mu_\beta \partial_\rho J^\alpha_\mu - (\partial_\gamma J^\alpha_\mu) J^\mu_\beta J^\gamma_\rho \\ &= J^\mu_\beta \partial_\rho J^\alpha_\mu - (\partial_\rho J^\alpha_\mu) J^\mu_\beta = 0 \ . \end{aligned} \tag{5.6}$$

Here we have used the symmetry

$$\partial_\mu J^\alpha_\rho = \frac{\partial^2 y^\alpha}{\partial x^\mu \partial x^\rho} = \partial_\rho J^\alpha_\mu \tag{5.7}$$

and

$$J^\alpha_\mu J^\mu_\beta = \delta^\alpha_\beta \quad \Rightarrow \quad J^\alpha_\mu (\partial_\gamma J^\mu_\beta) = -(\partial_\gamma J^\alpha_\mu) J^\mu_\beta \ . \tag{5.8}$$

The remaining terms in (5.5) then just give rise to the tensorial transformation of a (1,1)-tensor. Thus we have shown that (5.3) indeed defines a tensor. Moreover, in a locally inertial coordinate system this reduces to the ordinary partial derivative, and we have thus, as desired, arrived at an appropriate tensorial generalisation of the partial derivative operator.

Remarks:

1. Analysing the above argument for the tensoriality of the covariant derivative, we see that it relies exclusively on the specific non-tensorial form of the transformation behaviour of the Christoffel symbols, not on the explicit form of the Christoffel symbols themselves.

   Thus any other object $\tilde{\Gamma}^\mu{}_{\nu\lambda}$ could also be used to define a covariant derivative (generalising the partial derivative and mapping tensors to tensors) provided that it transforms in the same way as the Christoffel symbols, i.e. provided that one has

$$\tilde{\Gamma}^\alpha_{\beta\gamma} = J^\alpha_\mu J^\nu_\beta J^\lambda_\gamma \tilde{\Gamma}^\mu{}_{\nu\lambda} + J^\alpha_\mu \partial_\beta J^\mu_\gamma \ . \tag{5.9}$$

This implies (and is equivalent to the fact) that the difference

$$C^{\mu}_{\ \nu\lambda} = \tilde{\Gamma}^{\mu}_{\ \nu\lambda} - \Gamma^{\mu}_{\ \nu\lambda} \tag{5.10}$$

transforms as a tensor. Thus, any such $\tilde{\Gamma}$ is of the form

$$\tilde{\Gamma}^{\mu}_{\ \nu\lambda} = \Gamma^{\mu}_{\ \nu\lambda} + C^{\mu}_{\ \nu\lambda} \tag{5.11}$$

where $C^{\mu}_{\ \nu\lambda}$ is a (1,2)-tensor, and could be used to define a corresponding covariant derivative $\tilde{\nabla}_{\mu}$.

Therefore the question arises if the covariant derivative defined in terms of the Christoffel symbols is somehow singled out or preferred. Indeed it is, and we will return to this question on various occasions below, in particular in section 5.4.

2. We could have arrived at the above definition of the covariant derivative (using the Christoffel symbols) in a somewhat more systematic way by appealing to the equivalence principle and/or general covariance. Namely, let $\{\xi^a\}$ be an inertial coordinate system. In an inertial coordinate system we can just use the ordinary partial derivative $\partial_b V^a$. We now define the new (improved, covariant) derivative $\nabla_\nu V^\mu$ in any other coordinate system $\{x^\mu\}$ by demanding that it transforms as a (1,1)-tensor, i.e. we define

$$\nabla_\nu V^\mu := \frac{\partial x^\mu}{\partial \xi^a}\frac{\partial \xi^b}{\partial x^\nu}\partial_b V^a \quad . \tag{5.12}$$

By a straightforward calculation one finds that

$$\nabla_\nu V^\mu = \partial_\nu V^\mu + \gamma^{\mu}_{\ \nu\lambda} V^\lambda \quad , \tag{5.13}$$

where $\gamma^{\mu}_{\ \nu\lambda}$ is our old friend (1.98)

$$\gamma^{\mu}_{\ \nu\lambda} = \frac{\partial x^\mu}{\partial \xi^a}\frac{\partial^2 \xi^a}{\partial x^\nu \partial x^\lambda} \quad . \tag{5.14}$$

One would then also be led to adopt (5.13) with $\gamma \to \Gamma$ as a definition of the *covariant derivative* in a general metric space or space-time (with the Christoffel symbols calculated from the metric in the usual way).

That $\nabla_\mu V^\nu$, defined in this way, is indeed a $(1,1)$ tensor, now follows directly from the way we arrived at the definition of the covariant derivative. Indeed, imagine transforming from inertial coordinates to another coordinate system $\{y^\alpha\}$. Then (5.12) is replaced by

$$\nabla_\beta V^\alpha := \frac{\partial y^\alpha}{\partial \xi^a}\frac{\partial \xi^b}{\partial y^\beta}\partial_b V^a \quad . \tag{5.15}$$

Comparing this with (5.12), we see that the two are related by

$$\nabla_\beta V^\alpha := \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial x^\nu}{\partial y^\beta}\nabla_\nu V^\mu \quad , \tag{5.16}$$

as required.

3. Notation: frequently, the covariant derivative $\nabla_\nu V^\mu$ is also denoted by a semi-colon,
$$\nabla_\nu V^\mu = V^\mu{}_{;\nu} \quad . \tag{5.17}$$
Since covariant derivatives do not necessarily commute (as we will discuss in detail in section 8), when using this notation one has to pay attention to the order (and reversal of the order) of indices,
$$\nabla_\lambda \nabla_\nu V^\mu = V^\mu{}_{;\nu\,;\lambda} \quad . \tag{5.18}$$
One can also define the covariant directional derivative of a vector field V along another vector field X^μ by
$$\nabla_X V^\mu \equiv X^\nu \nabla_\nu V^\mu \quad . \tag{5.19}$$

4. The appearance of the Christoffel-term in the definition of the covariant derivative may at first sight appear a bit unusual (even though it also appears when one just transforms Cartesian partial derivatives to polar coordinates etc.). There is a more invariant way of explaining the appearance of this term, related to the more coordinate-independent way of looking at tensors explained in section 4.6. Namely, since the $V^\mu(x)$ are really just the coefficients of the vector field $V(x) = V^\mu(x)\partial_\mu$ when expanded in the basis ∂_μ, a meanigful definition of the derivative of a vector field must take into account not only the change in the coefficients but must also include a prescription how bases at (infinitesimally) neighbouring points are related (or *connected*). Such a prescription is provided by the *Levi-Civita connection* $\Gamma^\mu_{\ \nu\lambda}$ (or a general *connection* $\tilde{\Gamma}^\mu_{\ \nu\lambda}$).

   Indeed, writing
$$\begin{aligned} \nabla_\nu V &= \nabla_\nu (V^\mu \partial_\mu) \\ &= (\partial_\nu V^\mu)\partial_\mu + V^\lambda (\nabla_\nu \partial_\lambda) \quad , \end{aligned} \tag{5.20}$$
   we see that the covariant derivative of the coordinate basis vector ∂_λ (i.e. $V^\lambda = 1$, $V^\mu = 0$ otherwise), is the linear transformation (a prescription for a change of basis)
$$\nabla_\nu \partial_\lambda = \Gamma^\mu_{\ \nu\lambda}\partial_\mu \quad . \tag{5.21}$$

5.2 Extension of the Covariant Derivative to Other Tensor Fields

So far we have defined the covariant derivative for vector fields, and we now want to extend the definition of the covariant derivative to other tensor fields. In order to achieve this, we now adopt a more systematic and axiomatic approach.

Our basic postulates for the covariant derivative are the following:

1. Linearity and Tensoriality

   ∇_μ is a linear operator that maps (p,q)-tensors to $(p,q+1)$-tensors

2. Generalisation of the Partial Derivative

   On scalars ϕ, the covariant derivative ∇_μ reduces to the ordinary partial derivative (since $\partial_\mu\phi$ is already a covector),

$$\nabla_\mu\phi = \partial_\mu\phi \ . \tag{5.22}$$

3. Leibniz Rule (or Product Rule)

   Acting on the direct product of tensors, ∇_μ satisfies a generalised Leibniz rule,

$$\nabla_\mu(A^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}B^{\rho_1\ldots\rho_r}_{\lambda_1\ldots\lambda_s}) = \nabla_\mu(A^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q})B^{\rho_1\ldots\rho_r}_{\lambda_1\ldots\lambda_s} + A^{\mu_1\ldots\mu_p}_{\nu_1\ldots\nu_q}\nabla_\mu B^{\rho_1\ldots\rho_r}_{\lambda_1\ldots\lambda_s} \tag{5.23}$$

We will now see that, demanding the above properties, in particular the Leibniz rule, there is a unique extension of the covariant derivative on vector fields to a differential operator on general tensor fields, mapping (p,q)- to $(p,q+1)$-tensors.

To define e.g. the covariant derivative for covectors U_μ, we note that $U_\mu V^\mu$ is a scalar for any vector V^μ so that

$$\nabla_\mu(U_\nu V^\nu) = \partial_\mu(U_\nu V^\nu) = (\partial_\mu U_\nu)V^\nu + U_\nu(\partial_\mu V^\nu) \tag{5.24}$$

(since the partial derivative satisfies the Leibniz rule), and we demand

$$\nabla_\mu(U_\nu V^\nu) = (\nabla_\mu U_\nu)V^\nu + U_\nu\nabla_\mu V^\nu \ . \tag{5.25}$$

As we know $\nabla_\mu V^\nu$, these two equations determine $\nabla_\mu U_\nu$ uniquely to be

$$\nabla_\mu U_\nu = \partial_\mu U_\nu - \Gamma^\lambda_{\mu\nu}U_\lambda \ . \tag{5.26}$$

That this is indeed a $(0,2)$-tensor can either be checked directly or, alternatively, is a consequence of the quotient theorem.

The extension to other (p,q)-tensors is now immediate. Here are two ways to proceed:

1. If the (p,q)-tensor is the direct product of p vectors and q covectors, then we already know its covariant derivative (using the Leibniz rule again). We simply adopt the same resulting formula for an arbitrary (p,q)-tensor.

2. Alternatively, contract the (p,q)-tensor with p covectors and q vectors to turn it into a scalar, and proceed as above for a covector.

Either way, the result is that the covariant derivative of a general (p,q)-tensor is the sum of the partial derivative, a Christoffel symbol with a positive sign for each of the p upper indices, and a Christoffel with a negative sign for each of the q lower indices. In equations

$$\begin{aligned}\nabla_\mu T^{\nu_1\cdots\nu_p}{}_{\rho_1\cdots\rho_q} &= \partial_\mu T^{\nu_1\cdots\nu_p}{}_{\rho_1\cdots\rho_q} \\ &+ \underbrace{\Gamma^{\nu_1}_{\mu\lambda}T^{\lambda\nu_2\cdots\nu_p}{}_{\rho_1\cdots\rho_q} + \ldots + \Gamma^{\nu_p}_{\mu\lambda}T^{\nu_1\cdots\nu_{p-1}\lambda}{}_{\rho_1\cdots\rho_q}}_{p\ \text{terms}} \\ &- \underbrace{\Gamma^{\lambda}_{\mu\rho_1}T^{\nu_1\cdots\nu_p}{}_{\lambda\rho_2\cdots\rho_q} - \ldots - \Gamma^{\lambda}_{\mu\rho_q}T^{\nu_1\cdots\nu_p}{}_{\rho_1\cdots\rho_{q-1}\lambda}}_{q\ \text{terms}} \end{aligned} \tag{5.27}$$

Having defined the covariant derivative for arbitrary tensors, we are also ready to define it for tensor densities. For this we recall that if T is a $(p,q;w)$ tensor density, then $g^{-w/2}T$ is a (p,q)-tensor. Thus $\nabla_\mu(g^{-w/2}T)$ is a $(p,q+1)$-tensor. To map this back to a tensor density of weight w, we multiply this by $g^{w/2}$, arriving at the definition

$$\nabla_\mu T := g^{+w/2}\nabla_\mu(g^{-w/2}T) \ \ . \tag{5.28}$$

Working this out explictly, one finds

$$\nabla_\mu T = -\frac{w}{2g}(\partial_\mu g)T + \nabla_\mu^{\text{tensor}}T \ \ , \tag{5.29}$$

where $\nabla_\mu^{\text{tensor}}$ just means the usual covariant derivative for (p,q)-tensors defined above. For example, for a scalar density ϕ one has

$$\nabla_\mu\phi = \partial_\mu\phi - \frac{w}{2g}(\partial_\mu g)\phi \ \ . \tag{5.30}$$

In particular, since the determinant g is a scalar density of weight $+2$, it follows that

$$\nabla_\mu g = 0 \ \ , \tag{5.31}$$

which obviously simplifies integrations by parts in integrals defined with the measure $\sqrt{g}d^4x$. However, it should be kept in mind that the crucial property that makes an integral like $\int\sqrt{g}\nabla_\alpha(\ldots)^\alpha$ a total derivative is *not* this fact but the fact that in this expression the $\sqrt{g}$ cancels and the integrand becomes an ordinary total derivative (cf. the discussion of the Gauss Theorem (5.61) below).

5.3 Main Properties of the Covariant Derivative

The main properties of the covariant derivative, in addition to those that were part of our postulates (like linearity and the Leibniz rule) are the following:

1. ∇_μ Commutes with Contraction

   This means that if A is a (p,q)-tensor and B is the $(p-1,q-1)$-tensor obtained by contraction over two particular indices, then the covariant derivative of B is the same as the covariant derivative of A followed by contraction over these two indices. This comes about because of a cancellation between the corresponding two Christoffel symbols with opposite signs. Consider e.g. a (1,1)-tensor $A^\nu{}_\rho$ and its contraction $A^\nu{}_\nu$. The latter is a scalar and hence its covariant derivative is just the partial derivative. This can also be obtained by taking first the covariant derivative of A,

$$\nabla_\mu A^\nu{}_\rho = \partial_\mu A^\nu{}_\rho + \Gamma^\nu{}_{\mu\lambda} A^\lambda{}_\rho - \Gamma^\lambda{}_{\mu\rho} A^\nu{}_\lambda \ , \tag{5.32}$$

   and then contracting:

$$\nabla_\mu A^\nu{}_\nu = \partial_\mu A^\nu{}_\nu + \Gamma^\nu{}_{\mu\lambda} A^\lambda{}_\nu - \Gamma^\lambda{}_{\mu\nu} A^\nu{}_\lambda = \partial_\mu A^\nu{}_\nu \ . \tag{5.33}$$

   The most transparent way of stating this property is that the Kronecker delta is covariantly constant, i.e. that

$$\nabla_\mu \delta^\nu_\lambda = 0 \ . \tag{5.34}$$

   To see this, we use the Leibniz rule to calculate

$$\begin{aligned} \nabla_\mu A^{\nu\dots}_{\nu\dots} &= \nabla_\mu (A^{\nu\dots}_{\rho\dots} \delta^\rho_\nu) \\ &= (\nabla_\mu A^{\nu\dots}_{\rho\dots}) \delta^\rho_\nu + A^{\nu\dots}_{\rho\dots} \nabla_\mu \delta^\rho_\nu \\ &= (\nabla_\mu A^{\nu\dots}_{\rho\dots}) \delta^\rho_\nu \end{aligned} \tag{5.35}$$

   which is precisely the statement that covariant differentiation and contraction commute. To establish that the Kronecker delta is covariantly constant, we follow the rules to find

$$\begin{aligned} \nabla_\mu \delta^\nu_\lambda &= \partial_\mu \delta^\nu_\lambda + \Gamma^\nu{}_{\mu\rho} \delta^\rho_\lambda - \Gamma^\rho{}_{\mu\lambda} \delta^\nu_\rho \\ &= \Gamma^\nu{}_{\mu\lambda} - \Gamma^\nu{}_{\mu\lambda} = 0 \ . \end{aligned} \tag{5.36}$$

   This property does not rely on the specific form of the $\Gamma^\mu{}_{\nu\lambda}$, and is thus true for any covariant derivative defined by some choice of connection $\tilde{\Gamma}^\mu{}_{\nu\lambda}$,

2. The Metric is Covariantly Constant: $\nabla_\mu g_{\nu\lambda} = 0$

   This is one of the key properties of the covariant derivative ∇_μ we have defined. I will give two arguments to establish this:

   (a) Since $\nabla_\mu g_{\nu\lambda}$ is a tensor, we can choose any coordinate system we like to establish if this tensor is zero or not at a given point x. Choose an inertial coordinate system at x. Then the partial derivatives of the metric and the Christoffel symbols are zero there. Therefore the covariant derivative of the metric is zero. Since $\nabla_\mu g_{\nu\lambda}$ is a tensor, this is then true in every coordinate system.

(b) The other argument is by direct calculation. Recalling the identity

$$\partial_\mu g_{\nu\lambda} = \Gamma_{\nu\lambda\mu} + \Gamma_{\lambda\nu\mu} \ , \tag{5.37}$$

we calculate

$$\begin{aligned} \nabla_\mu g_{\nu\lambda} &= \partial_\mu g_{\nu\lambda} - \Gamma^\rho_{\ \mu\nu} g_{\rho\lambda} - \Gamma^\rho_{\ \mu\lambda} g_{\nu\rho} \\ &= \Gamma_{\nu\lambda\mu} + \Gamma_{\lambda\nu\mu} - \Gamma_{\lambda\mu\nu} - \Gamma_{\nu\mu\lambda} \\ &= 0 \ . \end{aligned} \tag{5.38}$$

3. ∇_μ Commutes with Raising and Lowering of Indices

   This is really a direct consequence of the covariant constancy of the metric. For example, if V_μ is the covector obtained by lowering an index of the vector V^μ, $V_\mu = g_{\mu\nu} V^\nu$, then

$$\nabla_\lambda V_\mu = \nabla_\lambda (g_{\mu\nu} V^\nu) = g_{\mu\nu} \nabla_\lambda V^\nu \ . \tag{5.39}$$

4. Covariant Derivatives Commute on Scalars

   This is of course a familiar property of the ordinary partial derivative, but it is also true for the second covariant derivatives of a scalar and is a consequence of the symmetry of the Christoffel symbols in the second and third indices and is also knowns as the *no torsion* property of the covariant derivative. Namely, we have

$$\begin{aligned} \nabla_\mu \nabla_\nu \phi - \nabla_\nu \nabla_\mu \phi &= \nabla_\mu \partial_\nu \phi - \nabla_\nu \partial_\mu \phi \\ &= \partial_\mu \partial_\nu \phi - \Gamma^\lambda_{\ \mu\nu} \partial_\lambda \phi - \partial_\nu \partial_\mu \phi + \Gamma^\lambda_{\ \nu\mu} \partial_\lambda \phi = 0 \ . \end{aligned} \tag{5.40}$$

   Note that *the second covariant derivatives on higher rank tensors do not commute* - we will come back to this in our discussion of the curvature tensor later on.

5.4 Uniqueness of the Levi-Civita Connection (Christoffel symbols)

We noted before that the postulates for a covariant derivative (a linear tensorial operator reducing to the partial derivative on scalars and satisfying the Leibniz rule) do not determine it uniquely but only up to the addition of a tensor to the connection,

$$\Gamma^\mu_{\ \nu\lambda} \rightarrow \tilde{\Gamma}^\mu_{\ \nu\lambda} = \Gamma^\mu_{\ \nu\lambda} + C^\mu_{\ \nu\lambda} \ , \tag{5.41}$$

where $C^\mu_{\ \nu\lambda}$ is a (1,2)-tensor. Is there anything special or preferred about the Levi-Civita connection using the Christoffel symbols?

In some sense, the answer is an immediate yes because it is this particular covariant derivative (or connection) that enters in determining the paths of freely falling particles (the geodesics which extremise proper time).

Not unrelated to this is the fact that it is the unique connection that can be built from only the metric and its 1st derivatives (and which thus vanishes in an inertial coordinate system in Minkowski space or at the origin of an inertial coordinate system in an arbitrary gravitational field).

Moreover, as we have seen, this covariant derivative has two important properties, namely that

1. the metric is covariantly constant, $\nabla_\mu g_{\nu\lambda} = 0$, and
2. the torsion is zero, i.e. the second covariant derivatives of a scalar commute.

In fact, it turns out that these two conditions uniquely determine the $\tilde{\Gamma}$ to be the Christoffel symbols. The second condition implies that the $\tilde{\Gamma}^\mu{}_{\nu\lambda}$ are symmetric in the two lower indices,

$$[\tilde{\nabla}_\mu, \tilde{\nabla}_\nu]\phi = 0 \quad \Leftrightarrow \quad \tilde{\Gamma}^\lambda{}_{\mu\nu} = \tilde{\Gamma}^\lambda{}_{\nu\mu} \quad . \tag{5.42}$$

The first condition now allows one to express the $\tilde{\Gamma}^\lambda{}_{\mu\nu}$ in terms of the derivatives of the metric, leading uniquely to the familiar expression for the Christoffel symbols $\Gamma^\mu{}_{\nu\lambda}$: First of all, by definition / construction one has (e.g. from demanding the Leibniz rule for $\tilde{\nabla}_\mu$)

$$\tilde{\nabla}_\mu g_{\nu\lambda} = \partial_\mu g_{\nu\lambda} - \tilde{\Gamma}^\rho{}_{\nu\mu} g_{\rho\lambda} - \tilde{\Gamma}^\rho{}_{\lambda\mu} g_{\nu\rho} \equiv \partial_\mu g_{\nu\lambda} - \tilde{\Gamma}_{\lambda\nu\mu} - \tilde{\Gamma}_{\nu\lambda\mu} \quad . \tag{5.43}$$

Requiring that this be zero implies in particular that

$$\begin{aligned} 0 &= \tilde{\nabla}_\mu g_{\nu\lambda} + \tilde{\nabla}_\nu g_{\mu\lambda} - \tilde{\nabla}_\lambda g_{\mu\nu} \\ &= \partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu} - \tilde{\Gamma}_{\lambda\nu\mu} - \tilde{\Gamma}_{\nu\lambda\mu} - \tilde{\Gamma}_{\lambda\mu\nu} - \tilde{\Gamma}_{\mu\lambda\nu} + \tilde{\Gamma}_{\nu\mu\lambda} + \tilde{\Gamma}_{\mu\nu\lambda} \\ &= 2(\Gamma_{\lambda\mu\nu} - \tilde{\Gamma}_{\lambda\mu\nu}) \end{aligned} \tag{5.44}$$

(where the cancellations are entirely due to the assumed symmetry of the coefficients in the last two indices). Thus $\tilde{\Gamma} = \Gamma$. This unique metric-compatible and torsion-free connection is also known as the *Levi-Civita connection*. It is the connection canonically associated to a space-time (manifold) equipped with a metric tensor, and it is the connection used in general relativity.

It is possible to relax either of the conditions (1) or (2), or both of them and this will be discussed in section 11.5, and subsequently also in section 20.7.

5.5 Tensor Analysis: Some Special Cases

In this section we will look at some common and useful special cases of the Levi-Civita covariant derivative (simply "*the* covariant derivative" in the following), such as the covariant curl and divergence etc.

1. The Covariant Curl of a Covector

   One has

$$\nabla_\mu U_\nu - \nabla_\nu U_\mu = \partial_\mu U_\nu - \partial_\nu U_\mu \quad , \tag{5.45}$$

   because the symmetric Christoffel symbols drop out in this anti-symmetric linear combination. Thus in particular the Maxwell field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5.46}$$

   is a tensor under general coordinate transformations, no metric or covariant derivative is needed to make it a tensor in a general space-time. The fact that the (ordinary) curl of a covector is, i.e. transforms like, a tensor under general coordinate transformations can of course also be checked directly (and then "explains" the above identity). We will come back to this below.

2. The Covariant Curl of an Antisymmetric Tensor

   Let $A_{\nu\lambda\ldots}$ be completely anti-symmetric. Then, as for the curl of covectors, the metric and Christoffel symbols drop out of the expression for the curl, i.e. one has

$$\nabla_{[\mu} A_{\nu\lambda\cdots]} = \partial_{[\mu} A_{\nu\lambda\cdots]} \quad . \tag{5.47}$$

   Here the square brackets on the indices denote complete anti-symmetrisation. In particular, the Bianchi identity for the Maxwell field strength tensor is independent of the metric also in a general metric space-time.

3. The Covariant Divergence of a Vector

   By the covariant divergence of a vector field one means the scalar

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu{}_{\mu\lambda} V^\lambda \quad . \tag{5.48}$$

   Now a useful identity for the contracted Christoffel symbol is

$$\Gamma^\mu{}_{\mu\lambda} = \frac{1}{\sqrt{g}} \partial_\lambda \sqrt{g} \quad . \tag{5.49}$$

   I will give a proof of this identity in an appendix to this section (subsection 5.6).

   Thus the covariant divergence can be written compactly as

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} V^\mu) \quad , \tag{5.50}$$

   and one only needs to calculate g and its derivative, not the Christoffel symbols themselves, to calculate the covariant divergence of a vector field.

   This formula is also useful (and provides the quickest way of arriving at the result) if one just wants to write the ordinary flat space divergence of vector calculus on $\mathbb{R}^3$ in, say, polar or cylindrical coordinates.

In Cartesian coordinates (x^1, x^2, x^3), the divergence of a 3-vector $\vec{V}$ is of course given by the familiar expression

$$\mathsf{div}\vec{V} = \partial_1 V^1 + \partial_2 V^2 + \partial_3 V^3 \quad . \tag{5.51}$$

However, as you also know, e.g. in spherical coordinates (r, θ, ϕ) the divergence is not simply of this form,

$$\mathsf{div}\vec{V} \neq \partial_r V^r + \partial_\theta V^\theta + \partial_\phi V^\phi \quad . \tag{5.52}$$

Rather, going through the coordinate transformation and Jacobians etc., one finds that calculating the divergence in spherical coordinates one picks up additional terms, the result taking the somewhat unintuitive form

$$\mathsf{div}\vec{V} = \partial_r V^r + \partial_\theta V^\theta + \partial_\phi V^\phi + \frac{2}{r} V^r + \cot\theta V^\theta \quad . \tag{5.53}$$

The easy and quick way to obtain this, which provides a rationale for and explanation of the origin of these additional terms, is from the result (5.50). Using $\sqrt{g} = r^2 \sin\theta$, one has

$$\begin{aligned} \mathsf{div}\vec{V} &= \frac{1}{r^2 \sin\theta} \left[\partial_r (r^2 \sin\theta V^r) + \partial_\theta (r^2 \sin\theta V^\theta) + \partial_\phi (r^2 \sin\theta V^\phi) \right] \\ &= \partial_r V^r + \partial_\theta V^\theta + \partial_\phi V^\phi + \frac{2}{r} V^r + \cot\theta V^\theta \quad . \end{aligned} \tag{5.54}$$

This thus produces the correct result on the nose and with very little effort.

4. The Covariant Laplacian of a Scalar

How should the Laplacian be defined? Well, the obvious guess (something that is covariant and reduces to the ordinary Laplacian for the Minkowski metric) is $\Box = g^{\mu\nu} \nabla_\mu \nabla_\nu$, which can alternatively be written as

$$\Box = g^{\mu\nu} \nabla_\mu \nabla_\nu = \nabla^\mu \nabla_\mu = \nabla_\mu \nabla^\mu = \nabla_\mu g^{\mu\nu} \nabla_\nu \tag{5.55}$$

etc. Note that, even though the covariant derivative on scalars reduces to the ordinary partial derivative, so that one can write

$$\Box\phi = \nabla_\mu g^{\mu\nu} \partial_\nu \phi \quad , \tag{5.56}$$

it makes no sense to write this as $\nabla_\mu \partial^\mu \phi$: since ∂_μ does not commute with the metric in general, the notation ∂^μ is at best ambiguous as it is not clear whether this should represent $g^{\mu\nu}\partial_\nu$ or $\partial_\nu g^{\mu\nu}$ or something altogether different. This ambiguity does not arise for the Minkowski metric, but of course it is present in general.

A compact yet explicit expression for the Laplacian follows from the expression for the covariant divergence of a vector:

$$\begin{aligned}\Box\phi &:= g^{\mu\nu}\nabla_\mu\nabla_\nu\phi \\ &= \nabla_\mu(g^{\mu\nu}\partial_\nu\phi) \\ &= g^{-1/2}\partial_\mu(g^{1/2}g^{\mu\nu}\partial_\nu\phi) \ . \end{aligned} \tag{5.57}$$

Again, this formula is also useful (and provides the quickest way of arriving at the result) if one just wants to write the ordinary flat space Laplacian on $\mathbb{R}^3$ in, say, polar or cylindrical coordinates.

To illustrate this, let us calculate the Laplacian for the standard metric on $\mathbb{R}^{n+1}$ in polar coordinates. The standard procedure would be to first determine the coordinate transformation $x^i = x^i(r, \text{angles})$, then calculate $\partial/\partial x^i$, and finally assemble all the bits and pieces to calculate $\Delta = \sum_i(\partial/\partial x^i)^2$. This is a pain.

To calculate the Laplacian, we do not need to know the coordinate transformation, all we need is the metric. In polar coordinates, this metric takes the form

$$ds^2(\mathbb{R}^{n+1}) = dr^2 + r^2 d\Omega_n^2 \ , \tag{5.58}$$

where $d\Omega_n^2$ is the standard line-element on the unit n-sphere S^n. The determinant of this metric is $g \sim r^{2n}$ (times a function of the coordinates (angles) on the sphere). Thus, for $n = 1$ one has $ds^2 = dr^2 + r^2 d\phi^2$ and therefore

$$\Delta = r^{-1}\partial_\mu(rg^{\mu\nu}\partial_\nu) = r^{-1}\partial_r(r\partial_r) + r^{-2}\partial_\phi^2 = \partial_r^2 + r^{-1}\partial_r + r^{-2}\partial_\phi^2 \ . \tag{5.59}$$

In general, denoting the angular part of the Laplacian, i.e. the Laplacian of S^n, by Δ_{S^n}, one finds analogously

$$\Delta = \partial_r^2 + nr^{-1}\partial_r + r^{-2}\Delta_{S^n} \ . \tag{5.60}$$

I hope you agree that this method is superior to the standard procedure.

5. The Covariant Form of the Gauss Theorem

Let V^μ be a vector field, $\nabla_\mu V^\mu$ its divergence and recall that integrals in curved space are defined with respect to the integration measure $\sqrt{g}d^4x$. Thus one has

$$\int \sqrt{g}d^4x\nabla_\mu V^\mu = \int d^4x\partial_\mu(\sqrt{g}V^\mu) \ . \tag{5.61}$$

Now the integrand on the right-hand side is an ordinary total derivative and thus, by the usual fundamental theorem of calculus, the integral of this over some region R can be written as an integral over the boundary ∂R of that region,

$$\int_R \sqrt{g}d^4x\nabla_\mu V^\mu = \int_{\partial R} d^3x(\dots) \ . \tag{5.62}$$

A somewhat more precise statement of this theorem, including the precise boundary contribution to the integral (involving the component of V^μ normal to the boundary), will be given in section 16.3.

In particular, if

- either the integral is over all of space-time and V^μ vanishes sufficiently rapidly at infinity,
- or the integral is over some region R with (finite) boundary ∂R, and the integrand vanishes on ∂R,

one has

$$\int \sqrt{g} d^4x \nabla_\mu V^\mu = 0 \quad . \tag{5.63}$$

In these circumstances, one also has the integration by parts formula

$$\int \sqrt{g} d^4x \nabla_\mu (f V^\mu) = 0 \quad \Rightarrow \quad \int \sqrt{g} d^4x f \nabla_\mu V^\mu = -\int \sqrt{g} d^4x (\nabla_\mu f) V^\mu \tag{5.64}$$

for a scalar f and a vector field V^μ.

6. The Covariant Divergence of an Antisymmetric Tensor

   For a $(p,0)$-tensor $T^{\mu\nu\cdots}$ one has

$$\begin{aligned} \nabla_\mu T^{\mu\nu\cdots} &= \partial_\mu T^{\mu\nu\cdots} + \Gamma^\mu_{\ \mu\lambda} T^{\lambda\nu\cdots} + \Gamma^\nu_{\ \mu\lambda} T^{\mu\lambda\cdots} + \ldots \\ &= g^{-1/2}\partial_\mu (g^{1/2} T^{\mu\nu\cdots}) + \Gamma^\nu_{\ \mu\lambda} T^{\mu\lambda\cdots} + \ldots \quad . \end{aligned} \tag{5.65}$$

   In particular, if $A^{\mu\nu\cdots}$ is completely anti-symmetric, the Christoffel terms disappear and one is left with

$$\nabla_\mu A^{\mu\nu\cdots} = g^{-1/2}\partial_\mu (g^{1/2} A^{\mu\nu\cdots}) \quad . \tag{5.66}$$

7. The Lie derivative of the Metric

   In section 3.2 we had encountered the expression (3.34) for the variation of the metric under an infinitesimal coordinate transformation $\delta x^\alpha = \epsilon V^\alpha$,

$$\delta_V g_{\alpha\beta} = V^\gamma \partial_\gamma g_{\alpha\beta} + (\partial_\alpha V^\gamma) g_{\gamma\beta} + (\partial_\beta V^\gamma) g_{\alpha\gamma} \quad . \tag{5.67}$$

   While we saw that this expression could be understood and deduced from the requirement that the variation of the metric is itself a tensorial object that transforms like the metric, the tensorial nature of the above expression is far from manifest. However, it has a very nice and simple expression in terms of covariant derivatives of V, namely

$$\delta_V g_{\alpha\beta} = \nabla_\alpha V_\beta + \nabla_\beta V_\alpha \tag{5.68}$$

   (as is easily verified).

Thus a vector field K^α generates a symmetry of the metric (such vectors are called *Killing vectors*) if it satisfies the *Killing equation*

$$K \quad \text{Killing Vector} \quad \Leftrightarrow \quad \nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0 \ . \tag{5.69}$$

We can also obtain this condition as the covariantisation of the statement that in a particular coordinate system the coefficients of the metric do not depend on one of these coordinates, say y,

$$\partial_y g_{\alpha\beta} = 0 \ , \tag{5.70}$$

so that the metric is then manifestly invariant under translations in y. In such a coordinate system adapted to the symmetry at hand, these translations are generated by $K = \partial_y$, and for a vector of this form (in particular, thus, with constant coefficients) one has

$$\begin{aligned} K = \partial_y \quad &\Rightarrow \quad \nabla_\alpha K^\beta = \Gamma^\beta_{\ \alpha y} \\ &\Rightarrow \quad \nabla_\alpha K_\beta = \Gamma_{\beta\alpha y} \\ &\Rightarrow \quad \nabla_\alpha K_\beta + \nabla_\beta K_\alpha = \partial_y g_{\alpha\beta} \end{aligned} \tag{5.71}$$

(where in the last step the basic relation (2.68) was used). Thus we find that the fact that the metric is y-translation invariant can be characterised covariantly as the statement that $K = \partial_y$ satisfies

$$\partial_y g_{\alpha\beta} = 0 \quad \Leftrightarrow \quad \nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0 \ . \tag{5.72}$$

This is again the Killing equation (5.69). As this equation is now tensorial it is valid in any coordinate system, in particular independently of whether or not the coordinate system is adapted to K in the way described above.

The expressions (5.68) and (5.72) will be rederived (and placed into the general context of Lie derivatives and Killing vectors) in section 9 - see in particular section 9.5.

You will have noticed that many equations simplify considerably for completely anti-symmetric tensors. In particular, their curl can be defined in a tensorial way without reference to any metric. This observation is at the heart of the coordinate independent calculus of *differential forms*. In this context, the curl is known as the *exterior derivative*.

Indeed, it is also straightforward to show directly, i.e. without going through the illogical loop of introducing the covariant derivative in order to obtain something manifestly tensorial only to find it disappear again from the final expression, that $\partial_{[\mu} A_{\mu_1 \ldots \mu_p]}$ is a tensor, i.e. transforms as a tensor under coordinate transformations: what happens is that the possible obstructions to the tensorial behaviour, namely derivatives of Jacobians, drop out after anti-symmetrisations because they are are really 2nd partial derivatives of the coordinates, which are symmetric and thus do not survive the anti-symmetrisation.

To see this completely explicitly, consider a covector $A_\mu(x)$ and a coordinate transformation $x^\mu = x^\mu(y^\alpha)$, with Jacobi matrix

$$J^\mu_\alpha = \frac{\partial x^\mu}{\partial y^\alpha} \quad . \tag{5.73}$$

As a covector, A_μ transforms as $A_\alpha = J^\mu_\alpha A_\mu$, and therefore its derivative transforms as (using $\partial_\beta = J^\nu_\beta \partial_\nu$)

$$A_\alpha = J^\mu_\alpha A_\mu \quad \Rightarrow \quad \partial_\beta A_\alpha = J^\mu_\alpha J^\nu_\beta \partial_\nu A_\mu + (\partial_\beta J^\mu_\alpha) A_\mu \quad . \tag{5.74}$$

Because of

$$\partial_\beta J^\mu_\alpha = \frac{\partial^2 x^\mu}{\partial y^\alpha \partial y^\beta} = \partial_\alpha J^\mu_\beta \quad , \tag{5.75}$$

for the anti-symmetrised derivative one finds the tensorial transformation behaviour

$$\partial_\beta A_\alpha - \partial_\alpha A_\beta = J^\mu_\alpha J^\nu_\beta (\partial_\nu A_\mu - \partial_\mu A_\nu) \quad . \tag{5.76}$$

Likewise, Lie derivatives of tensors in general (section 9) are, as the special case of the Lie derivative of the metric mentioned above - see (5.68), automatically tensorial objects (and one can, but need not, make their tensorial nature manifest by writing these derivatives in terms of covariant derivatives).

5.6 Appendix: A Formula for the Variation of the Determinant

Here is an elementary proof of the identity (5.49), and a useful more general formula for the variation of the determinant of the metric, namely

$$\delta g = g g^{\mu\nu} \delta g_{\mu\nu} \quad \text{or} \quad g^{-1} \delta g = g^{\mu\nu} \delta g_{\mu\nu} \quad . \tag{5.77}$$

This proof is based on the standard cofactor or minor expansion of the determinant of a matrix (an alternative standard proof can, as also outlined below, be based on the identity $\det G = \exp \operatorname{tr} \log G$ and its derivative or variation). The cofactor expansion formula for the determinant is

$$g = \sum_\nu (-1)^{\mu+\nu} g_{\mu\nu} |m_{\mu\nu}| \quad , \tag{5.78}$$

for a fixed (but arbitrary) value of the index μ. Here $|m_{\mu\nu}|$ is the determinant of the minor of $g_{\mu\nu}$, i.e. of the matrix one obtains by removing the μ'th row and ν'th column from $g_{\mu\nu}$.

As a consequence of (5.78) one also has

$$\sum_\nu (-1)^{\mu+\nu} g_{\lambda\nu} |m_{\mu\nu}| = 0 \qquad \lambda \neq \mu \quad , \tag{5.79}$$

since this is, in particular, the determinant of a matrix with $g_{\mu\nu} = g_{\lambda\nu}$, i.e. of a matrix with two equal rows. Together, these two results can be written as

$$\sum_\nu (-1)^{\mu+\nu} g_{\lambda\nu} |m_{\mu\nu}| = \delta_{\lambda\mu} g \quad . \tag{5.80}$$

This shows that the coefficients of the inverse metric $g^{\mu\nu}$ are given by

$$g^{\nu\mu} = (-1)^{\mu+\nu} \frac{|m_{\mu\nu}|}{g} \quad , \tag{5.81}$$

a formula that should also be familiar from linear algebra.

In order to now determine the variation (or derivative) of the determinant with respect to a variation of the matrix elements $g_{\mu\nu}$, we first consider the effect of a variation $\delta g_{\mu_0\nu_0}$ for a fixed value (μ_0, ν_0) of the index pair $(\mu\nu)$, and then sum over all (μ_0, ν_0). In order to determine the variation with respect to $\delta g_{\mu_0\nu_0}$, note the following: in the above expansion (5.78), for a fixed μ_0 none of the $m_{\mu_0\nu}$ depend on $g_{\mu_0\nu_0}$, and the only dependence of g on $g_{\mu_0\nu_0}$ is via the explicit appearance of $g_{\mu_0\nu_0}$ in the summand $g_{\mu_0\nu_0}|m_{\mu_0\nu_0}|$. Thus the variation of the determinant under a variation $\delta g_{\mu_0,\nu_0}$ of $g_{\mu_0\nu_0}$ is simply

$$g_{\mu_0\nu_0} \rightarrow g_{\mu_0\nu_0} + \delta g_{\mu_0\nu_0} \quad \Rightarrow \quad \delta g = (-1)^{\mu_0+\nu_0}(\delta g_{\mu_0\nu_0})|m_{\mu_0\nu_0}| \quad . \tag{5.82}$$

Now doing this for each $g_{\mu_0\nu_0}$ and summing, one thus finds

$$\delta g = \sum_\mu \sum_\nu (-1)^{\mu+\nu} \delta g_{\mu\nu} |m_{\mu\nu}| = g g^{\nu\mu} \delta g_{\mu\nu} \quad . \tag{5.83}$$

For a symmetric matrix, in particular for the metric, this reduces to the formula (5.77) we set out to establish. Here are some variations and applications of this formula:

1. When the determinant g is viewed as a (smooth) function of the coefficients $g_{\mu\nu}$, this shows that
$$\frac{\partial g}{\partial g_{\mu\nu}} = g g^{\nu\mu} \quad . \tag{5.84}$$

2. It also implies
$$\delta\sqrt{g} = \tfrac{1}{2}\sqrt{g} g^{\nu\mu} \delta g_{\mu\nu} \quad , \tag{5.85}$$
a particularly useful result that we will repeatedly make use of.

3. An equally useful variant of this equation is an expression for the variation of $\sqrt{g}$ expressed in terms of the variations $\delta g^{\mu\nu}$ of the components of the inverse metric. As a consequence of
$$g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_{\ \lambda} \quad \Rightarrow \quad \delta g^{\mu\nu} = -g^{\mu\lambda} \delta g_{\lambda\rho} g^{\rho\nu} \tag{5.86}$$
or
$$g^{\mu\nu} g_{\mu\nu} = 4 \quad \Rightarrow \quad (\delta g^{\mu\nu}) g_{\mu\nu} = -g^{\mu\nu} \delta g_{\mu\nu} \tag{5.87}$$
one can equivalently write (5.85) as
$$\delta\sqrt{g} = \tfrac{1}{2}\sqrt{g} g^{\nu\mu} \delta g_{\mu\nu} = -\tfrac{1}{2}\sqrt{g} g_{\nu\mu} \delta g^{\mu\nu} \quad . \tag{5.88}$$

4. It follows from (5.84) that if the variation is the partial derivative ("how does the determinant $g = g(x)$ of the metric vary with x?") one has

$$\partial_\lambda g = \frac{\partial g}{\partial g_{\mu\nu}} \partial_\lambda g_{\mu\nu} = g g^{\mu\nu} \partial_\lambda g_{\mu\nu} \quad . \tag{5.89}$$

or

$$g^{-1} \partial_\lambda g = g^{\mu\nu} \partial_\lambda g_{\mu\nu} \quad , \tag{5.90}$$

and therefore also

$$\partial_\lambda \sqrt{g} = \tfrac{1}{2} \sqrt{g} g^{\mu\nu} \partial_\lambda g_{\mu\nu} \quad . \tag{5.91}$$

5. On the other hand, the contracted Christoffel symbol is

$$\Gamma^{\mu}_{\ \mu\lambda} = \frac{1}{2} g^{\mu\nu} \partial_\lambda g_{\mu\nu} \quad . \tag{5.92}$$

Therefore

$$\Gamma^{\mu}_{\ \mu\lambda} = \frac{1}{\sqrt{g}} \partial_\lambda \sqrt{g} \tag{5.93}$$

which establishes the identity (5.49).

The result (5.83) can also be written in matrix form, with G denoting the matrix with components $(G)_{\mu\nu} = g_{\mu\nu}$, as

$$\delta \log \det G = \operatorname{tr} G^{-1} \delta G \quad . \tag{5.94}$$

In this form, the result can also be derived from variation of the remarkably useful identity

$$\det G = \mathrm{e}^{\operatorname{tr} \log G} \tag{5.95}$$

This identity, in turn, can be derived in an elementary way for diagonalisable G by noting that it holds trivially for diagonal matrices, and therefore, by the conjugation invariance of det and tr, also for diagonalisable matrices (like the metric). [And if desired, this can in turn be extended to all matrices by topological arguments involving extensions of continuous functionals from the dense set of diagonalisable matrices to the space of all matrices ...]

5.7 Covariant Differentiation Along a Curve

So far, we have defined covariant differentiation for tensors defined everywhere in space-time. Frequently, however, one encounters tensors that are only defined on curves - like the momentum of a particle which is only defined along its world line. In this section we will see how to define covariant differentiation along a curve. Thus consider a curve $x^\mu(\tau)$ (where τ could be, but need not be, proper time) and the tangent vector field $X^\mu(x(\tau)) = \dot{x}^\mu(\tau)$. Now define the covariant derivative D_τ along the curve, covariantising $d/d\tau$, by

$$\frac{d}{d\tau} = \dot{x}^\mu \partial_\mu \quad \rightarrow \quad D_\tau = X^\mu \nabla_\mu = \dot{x}^\mu \nabla_\mu \quad . \tag{5.96}$$

Frequently one also uses the (suggestive, but ugly) notation

$$D_\tau = D/D\tau \quad \text{or} \quad D/d\tau \ . \tag{5.97}$$

For example, for a vector one has

$$\begin{aligned} D_\tau V^\mu &= \dot{x}^\nu \partial_\nu V^\mu + \dot{x}^\nu \Gamma^\mu_{\nu\lambda} V^\lambda \\ &= \frac{d}{d\tau} V^\mu(x(\tau)) + \Gamma^\mu_{\nu\lambda}(x(\tau)) \dot{x}^\nu(\tau) V^\lambda(x(\tau)) \ . \end{aligned} \tag{5.98}$$

For this to make sense, V^μ needs to be defined only along the curve and not necessarily everywhere in space-time.

This notion of covariant derivative along a curve permits us, in particular, to define the (covariant) acceleration a^μ of a curve $x^\mu(\tau)$ as the covariant derivative of the velocity $u^\mu = \dot{x}^\mu$ along the curve,

$$a^\mu = D_\tau \dot{x}^\mu = \ddot{x}^\mu + \Gamma^\mu_{\nu\lambda} \dot{x}^\nu \dot{x}^\lambda = u^\nu \nabla_\nu u^\mu \ . \tag{5.99}$$

Thus we can characterise (affinely parametrised) geodesics as those curves whose covariant acceleration is zero,

$$\text{Geodesics:} \quad a^\mu = u^\nu \nabla_\nu u^\mu = 0 \ , \tag{5.100}$$

a reasonable and natural statement regarding the movement of freely falling particles. If they are not affinely parametrised, as in (2.136), then instead of $u^\nu \nabla_\nu u^\mu = 0$ one has

$$u^\nu \nabla_\nu u^\mu = \kappa u^\mu \ . \tag{5.101}$$

5.8 Parallel Transport and Geodesics

We now come to the important notion of *parallel transport* of a tensor along a curve. Note that, in a general (curved) metric space-time, it does not make sense to ask if two vectors defined at points x and y are parallel to each other or not. However, given a metric and a curve connecting these two points, one can compare the two by dragging one along the curve to the other using the covariant derivative.

We say that a tensor $T^{\cdots}_{\cdots}$ is *parallel transported along the curve* $x^\mu(\tau)$ if

$$D_\tau T^{\cdots}_{\cdots} = 0 \ . \tag{5.102}$$

Here are some immediate consequences of this definition:

1. In a locally inertial coordinate system along the curve, this condition reduces to $dT/d\tau = 0$, i.e. to the statement that the tensor does not change along the curve. Thus the above is indeed an appropriate tensorial generalisation of the intuitive notion of parallel transport to a general metric space-time.

2. The parallel transport condition is a first order differential equation along the curve and thus defines $T^{\cdots}_{\cdots}(\tau)$ given an initial value $T^{\cdots}_{\cdots}(\tau_0)$.

3. Taking T to be the tangent vector $u^\mu = \dot{x}^\mu$ to the curve itself, the condition for parallel transport becomes

$$D_\tau u^\mu = 0 \quad \Leftrightarrow \quad \ddot{x}^\mu + \Gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \ , \tag{5.103}$$

i.e. precisely the geodesic equation. We have already seen that geodesics are precisely the curves with zero acceleration. We can now equivalently characterise them by the property that their tangent vectors are parallel transported (do not change) along the curve. For this reason geodesics are also known as *autoparallels*.

4. Since the metric is covariantly constant, it is parallel along any curve. Thus, in particular, if V^μ is parallel transported, also its length remains constant along the curve,

$$D_\tau V^\mu = 0 \quad \Rightarrow \quad \frac{d}{d\tau}(g_{\mu\nu}V^\mu V^\nu) = D_\tau(g_{\mu\nu}V^\mu V^\nu) = 0 \ . \tag{5.104}$$

In particular, we rediscover the fact claimed in (2.100) that the quantity $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ is constant along a geodesic,

$$D_\tau \dot{x}^\mu = 0 \quad \Rightarrow \quad \frac{d}{d\tau}(g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu) = 0 \ . \tag{5.105}$$

5. Now let $x^\mu(\tau)$ be a geodesic and V^μ parallel along this geodesic. Then, as one might intuitively expect, also the angle between V^μ and the tangent vector to the curve u^μ remains constant. This is a consequence of the fact that both the norm of V and the norm of u are constant along the curve and that

$$\frac{d}{d\tau}(g_{\mu\nu}u^\mu V^\nu) = D_\tau(g_{\mu\nu}u^\mu V^\nu) = g_{\mu\nu}(D_\tau u^\mu)V^\nu + g_{\mu\nu}u^\mu D_\tau V^\nu = 0 \tag{5.106}$$

5.9 Example: Parallel Transport on the 2-Sphere

As usual, the simplest non-trivial example is provided by the 2-sphere with its standard line element

$$ds^2 = d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2 \ , \tag{5.107}$$

with the non-zero Christoffel symbols (determined e.g. from the geodesic equation, as in (3.16) - (3.21))

$$\Gamma^\theta_{\ \phi\phi} = -\sin\theta\ \cos\theta \quad , \quad \Gamma^\phi_{\ \theta\phi} = \Gamma^\phi_{\ \phi\theta} = \cot\theta \ . \tag{5.108}$$

Let us consider parallel transport of some vector $V = V^\alpha\partial_\alpha$ along a circle with $\theta = \theta_0$ constant, choosing the angle ϕ to parametrise the curve, i.e. we consider the family of paths

$$x^\alpha(\tau) = (\theta(\tau), \phi(\tau)) = (\theta_0, \tau) \tag{5.109}$$

with tangent vector

$$\dot{x}^\alpha = (0,1) \quad . \tag{5.110}$$

Note that this is not normalised in the standard way,

$$g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta = \sin^2\theta_0 \quad , \tag{5.111}$$

so proper distance would be measured not by ϕ but, as is also pictorially evident, by $s = \phi \sin\theta_0$ (which agrees with ϕ on the equator $\theta_0 = \pi/2$).

A vector $V^\alpha \partial_\alpha$ with coordinate components (V^θ, V^ϕ) parallel transported along such a curve thus satisfies the equations

$$\begin{aligned} 0 &= \partial_\phi V^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{x}^\beta V^\gamma = \partial_\phi V^\alpha + \Gamma^\alpha_{\phi\gamma} V^\gamma \\ &= \partial_\phi V^\alpha + \Gamma^\alpha_{\phi\phi} V^\phi + \Gamma^\alpha_{\phi\theta} V^\theta \quad . \end{aligned} \tag{5.112}$$

Using the explicit form of the Christoffel symbols, the parallel transport equations are thus

$$\begin{aligned} 0 &= \partial_\phi V^\theta - \sin\theta \;\cos\theta \; V^\phi \\ 0 &= \partial_\phi V^\phi + \cot\theta \; V^\theta \quad . \end{aligned} \tag{5.113}$$

Differentiating once more, these equations can be decoupled and take the form of harmonic oscillator equations with frequency $\cos\theta_0$,

$$(\partial_\phi^2 + \cos^2\theta_0) V^\alpha = 0 \quad . \tag{5.114}$$

The general solution of this 2nd order differential equation is of course

$$V^\alpha = A^\alpha \sin(\phi\cos\theta_0) + B^\alpha \cos(\phi\cos\theta_0) \quad . \tag{5.115}$$

Plugging this into the 1st order equations to reduce the spurious 4 to 2 integration constants, and relating them to the intial values at $\phi = 0$, say,

$$V^\alpha(\theta_0, \phi = 0) = v^\alpha \quad , \tag{5.116}$$

one finally finds the result

$$\begin{aligned} V^\theta(\theta_0,\phi) &= v^\phi \sin\theta_0 \; \sin(\phi\cos\theta_0) + v^\theta \; \cos(\phi\cos\theta_0) \\ V^\phi(\theta_0,\phi) &= -(v^\theta/\sin\theta_0) \; \sin(\phi\cos\theta_0) + v^\phi \; \cos(\phi\cos\theta_0) \quad . \end{aligned} \tag{5.117}$$

<u>REMARKS:</u>

1. In the special case of parallel transport along the equator $\theta_0 = \pi/2$, one has $\cos\theta_0 = 0$, and therefore

$$\theta_0 = \pi/2 \quad \Rightarrow \quad V^\alpha(\pi/2, \phi) = v^\alpha \quad . \tag{5.118}$$

In other words, the components are constant under parallel transport along the equator. This is inuitively obvious on the basis of spherical symmetry. Since among the family of constant $\theta = \theta_0$ curves only the equator is a geodesic (great circle), this is also in agreement with the general results obtained above, which imply that upon parallel transport along the equator the angle between the vector and the equator remains constant. In 2 dimensions, this condition, together with the fact that the lenght of a vector remains invariant under parallel transport in general, is sufficient to imply that the parallel transported components are constant along the path.

2. While the above is not unexpected, perhaps the most interesting consequence of the above result (5.117) is that, in general, not only are the components not constant but that actually, after having completed the 2π-circuit along the path to return to the starting point, the parallel transported vector will not agree with the initial vector. Indeed, the components at $\phi = 2\pi$ are related to the components v^α at $\phi = 0$ by

$$\begin{aligned} V^\theta(\theta_0, \phi = 2\pi) &= v^\phi \sin\theta_0 \ \sin(2\pi \ \cos\theta_0) + v^\theta \ \cos(2\pi \ \cos\theta_0) \\ V^\phi(\theta_0, \phi = 2\pi) &= -(v^\theta / \sin\theta_0) \ \sin(2\pi \ \cos\theta_0) + v^\phi \ \cos(2\pi \ \cos\theta_0) \ . \end{aligned} \tag{5.119}$$

3. As we will see in section 11.1, this fact that parallel transport along closed paths is non-trivial (equivalently that parallel transport from one point to another depends on the path) can be directly attributed to (and is the smoking gun of) the presence of *curvature*.

4. If desired, the result can be written in terms of proper distance s along the circle, rather than the angle ϕ, by the substitution

$$\phi \ \cos\theta_0 = s \ \cot\theta_0 \ . \tag{5.120}$$

5. The result (5.117) takes on a more transparent form when written in terms of the components of V and v with respect to an orthonormal basis (section 4.8) E_α rather than the coordinate basis ∂_α. Such an orthonormal basis is provided by

$$E_\theta = \partial_\theta \quad , \quad E_\phi = (\sin\theta)^{-1}\partial_\phi \ , \tag{5.121}$$

since one evidently has

$$g_{\alpha\beta}E^\alpha_\theta E^\beta_\theta = g_{\alpha\beta}E^\alpha_\phi E^\beta_\phi = 1 \quad , \quad g_{\alpha\beta}E^\alpha_\theta E^\beta_\phi = 0 \ . \tag{5.122}$$

The components with respect to this orthonormal basis are related to the coordinate components by

$$V = V^\alpha \partial_\alpha = \hat{V}^\alpha E_\alpha \quad \Rightarrow \quad \hat{V}^\theta = V^\theta \quad , \quad \hat{V}^\phi = \sin\theta \ V^\phi \tag{5.123}$$

(and likewise for $v = v^\alpha \partial_\alpha = \hat{v}^\alpha E_\alpha$). Then (5.117) can be written in matrix form as a rotation (orthogonal transformation)

$$\begin{pmatrix} \hat{V}^\theta(\theta_0,\phi) \\ \hat{V}^\phi(\theta_0,\phi) \end{pmatrix} = \begin{pmatrix} \cos(\phi\cos\theta_0) & \sin(\phi\cos\theta_0) \\ -\sin(\phi\cos\theta_0) & \cos(\phi\cos\theta_0) \end{pmatrix} \begin{pmatrix} \hat{v}^\theta \\ \hat{v}^\phi \end{pmatrix} \tag{5.124}$$

by the angle

$$\alpha(\phi) = \phi\cos\theta_0 \quad . \tag{5.125}$$

Thus parallel transport amounts to a continuous rotation of the orthonormal components along the path.

6. In particular, the angle that one picks up after a 2π-rotation,

$$\alpha(2\pi) = 2\pi\cos\theta_0 \quad , \tag{5.126}$$

is known as the *deficit angle* or *holonomy* of the parallel transport along the given loop. With this terminology we can say that the holonomy along the equator is trivial.

7. At the other extreme, we see that there is a non-trivial holonomy as $\theta_0 \to 0$, i.e. for parallel transport along an infinitesimal loop around the north pole, along which the parallel transported vector performs a complete 2π-rotation, $\alpha(2\pi) = 2\pi$. As shown in section 11.1, parallel transport along infinitesimal loops at or around a point provides a precise measure of the curvature at that point.

8. Curiously, as shown by Rothman, Ellis and Murugan, the holonomy along circular equatorial orbits in the Schwarzschild geometry (such orbits are geodesics at the critical points of the effective potential for geodesic motion, to be discussed in section 25.6), is non-trivial, even though again intuitive reasoning based on spherical symmetry might have led one to expect a trivial result (and would thus have led one astray).[13]

5.10 Fermi-Walker Parallel Transport

The properties of parallel transport established in section 5.8 show that this is a natural prescription for transporting tensorial objects along a geodesic. However, it is important to keep in mind that this is just one possible description, obtained by imposing the differential equation (5.102), e.g. for a vector

$$D_\tau V^\alpha = 0 \quad . \tag{5.127}$$

[13] T. Rothman, G. Ellis, J. Murugan, *Holonomy in the Schwarzschild-Droste Geometry*, `arXiv:gr-qc/0008070`.

If the curve is not a geodesic,

$$a^\alpha = D_\tau u^\alpha = \dot{x}^\beta \nabla_\beta \dot{x}^\alpha \neq 0 \ \ , \tag{5.128}$$

however, this prescription has some shortcomings. For example, parallel transport of a tangent vector to the curve at a point to another point at the curve will not give rise to the tangent vector at the second point, simply because $D_\tau V^\alpha = 0$ with initial condition $V^\alpha(\tau_0) = u^\alpha(\tau_0)$, say (parallel transport) is evidently not the same as $D_\tau u^\alpha = a^\alpha$ (the equation satisfied by the tangent vector). Likewise, the scalar product between the tangent vector to the (non-geodesic) curve and some parallel-transported vector along it will not remain constant in general,

$$D_\tau u^\alpha = a^\alpha \quad , \quad D_\tau V^\alpha = 0 \quad \Rightarrow \quad \frac{d}{d\tau}(g_{\alpha\beta} u^\alpha V^\beta) = a_\alpha V^\alpha \ \ . \tag{5.129}$$

A vivid illustration of this is provided by the example of the previous section:

- As we have seen, parallel transporting a pair of orthonormal vectors along a circle $\theta = \theta_0 \neq \pi/2$ results in a continuous rotation of these two basis vectors.
- On the other hand, it is clearly possible to transport an orthonormal basis along the circle in such a way that, for example, one basis vector always points forwards along the latitude (a tangent vector to the curve), and the other always points northwards along the longitude.

The latter procedure appears to be much more natural in this case than rotating one's basis as one goes around the sphere. Analogously, for an observer along a timelike curve it would be desirable to be able to set up once and for all a local reference system on the worldline, consisting of the (unit) tangent vector $E_0 = u^\mu \partial_\mu$ in the time-direction, and three orthogonal and mutually orthogonal vectors E_k in the spatial directions (the laboratory system of the observer), regardless of whether the observer is in free fall or not (indeed, for perfectly good reasons most laboratories are not ...).

This procedure can be formalised by replacing the parallel transport condition (5.127) along a timelike curve by the *Fermi-Walker Transport* prescription

$$F_\tau V^\alpha \equiv D_\tau V^\alpha + \mathcal{F}^\alpha_{\ \beta} V^\beta = 0 \ \ , \tag{5.130}$$

with

$$\mathcal{F}^\alpha_{\ \beta} = a^\alpha u_\beta - u^\alpha a_\beta \ \ . \tag{5.131}$$

Indeed, parallel transport according to this prescription has the following desirable features:

1. Fermi-Walker transport evidently reduces to parallel transport if the curve is a geodesic, i.e. for $a^\alpha = 0$. It obligingly does this even when the geodesic is not affinely parametrised, i.e. if one has $a^\alpha \sim u^\alpha$,

$$a^\alpha = u^\beta \nabla_\beta u^\alpha \sim u^\alpha \quad \Rightarrow \quad \mathcal{F}^\alpha_{\ \beta} = 0 \ \ . \tag{5.132}$$

2. The tangent vector to a curve is Fermi-Walker transported for any curve,

$$F_\tau u^\alpha = 0 \ . \tag{5.133}$$

Proof:

$$\begin{aligned} F_\tau u^\alpha &= D_\tau u^\alpha + \mathcal{F}^\alpha_{\ \beta} u^\beta \\ &= a^\alpha + (a^\alpha u_\beta - u^\alpha a_\beta) u^\beta = a^\alpha - a^\alpha = 0 \end{aligned} \tag{5.134}$$

because $u_\beta u^\beta = -1$ and $a_\beta u^\beta = 0$. Thus the solution to the Fermi-Walker transport prescription for $V^\alpha(\tau_0) = u^\alpha(\tau_0)$ is just the tangent vector u^α,

$$F_\tau V^\alpha = 0 \quad , \quad V^\alpha(\tau_0) = u^\alpha(\tau_0) \quad \Rightarrow \quad V^\alpha(\tau) = u^\alpha(\tau) \ . \tag{5.135}$$

3. If V^α is Fermi-Walker transported along the curve, then instead of (5.129) one obtains

$$D_\tau u^\alpha = a^\alpha \quad , \quad D_\tau V^\alpha = -\mathcal{F}^\alpha_{\ \beta} V^\beta \quad \Rightarrow \quad \frac{d}{d\tau}(g_{\alpha\beta} u^\alpha V^\beta) = 0 \ . \tag{5.136}$$

Proof:

$$\begin{aligned} \frac{d}{d\tau}(g_{\alpha\beta} u^\alpha V^\beta) &= D_\tau(g_{\alpha\beta} u^\alpha V^\beta) = g_{\alpha\beta}(a^\alpha V^\beta - u^\alpha \mathcal{F}^\beta_{\ \gamma} V^\gamma) \\ &= a_\alpha V^\alpha - a_\alpha u^\alpha u_\gamma V^\gamma + u_\alpha u^\alpha a_\gamma V^\gamma = a_\alpha V^\alpha - a_\gamma V^\gamma = 0 \ , \end{aligned} \tag{5.137}$$

beause $a_\alpha u^\alpha = 0$ and $u_\alpha u^\alpha = -1$.

4. Similarly, if V and W are Fermi-Wallker transported, one has

$$F_\tau V^\alpha = F_\tau W^\alpha = 0 \quad \Rightarrow \quad \frac{d}{d\tau}(g_{\alpha\beta} V^\alpha W^\beta) = 0 \ . \tag{5.138}$$

Proof:

$$\begin{aligned} \frac{d}{d\tau}(g_{\alpha\beta} V^\alpha W^\beta) &= g_{\alpha\beta}(-\mathcal{F}^\alpha_{\ \gamma} V^\gamma W^\beta - \mathcal{F}^\beta\ \gamma V^\alpha W^\gamma) \\ &= -(\mathcal{F}_{\alpha\gamma} + \mathcal{F}_{\gamma\alpha}) V^\alpha W^\gamma = 0 \end{aligned} \tag{5.139}$$

because $\mathcal{F}_{\alpha\gamma}$ is anti-symmetric.

Remarks:

1. The signs chosen here are appropriate for timelike curves with $u^\alpha u_\alpha = -1$. As the proofs of the above statements show, in the spacelike case one needs to replace $\mathcal{F}^\alpha_{\ \beta} \to -\mathcal{F}^\alpha_{\ \beta}$.

2. The above manipulations can be formalised (and then subsequently trivialised) by

   - extending the action of F_τ to arbitrary rank tensors in the same way as the covariant derivative, i.e. by requiring that on scalars it reduces to the ordinary derivative,

$$F_\tau f = \frac{d}{d\tau} f \ , \tag{5.140}$$

- extending it to arbitrary tensors by requiring the Leibniz rule, so that e.g. on covectors one has

$$F_\tau A_\beta = D_\tau A_\beta - \mathcal{F}^\alpha_{\ \beta} A_\alpha \quad , \tag{5.141}$$

- and then showing that as a consequence

$$\begin{aligned} F_\tau g_{\alpha\beta} &= D_\tau g_{\alpha\beta} - \mathcal{F}^\gamma_{\ \alpha} g_{\gamma\beta} - \mathcal{F}^\gamma_{\ \beta} g_{\alpha\gamma} \\ &= -(\mathcal{F}_{\alpha\beta} + \mathcal{F}_{\beta\alpha}) = 0 \quad . \end{aligned} \tag{5.142}$$

Then assertions like (5.136),

$$F_\tau u^\alpha = F_\tau V^\alpha = 0 \quad \Rightarrow \quad \frac{d}{d\tau}(u_\alpha V^\alpha) = 0 \quad . \tag{5.143}$$

or (5.138) become a triviality.

3. Note that the properties 2-4 in the above list rely on the 3 properties

$$\mathcal{F}^\alpha_{\ \beta} u^\beta = -a^\alpha \quad , \quad u_\alpha \mathcal{F}^\alpha_{\ \beta} = a_\beta \quad , \quad \mathcal{F}_{\alpha\beta} + \mathcal{F}_{\beta\alpha} = 0 \tag{5.144}$$

of $\mathcal{F}^\alpha_{\ \beta}$ respectively. These conditions determine $\mathcal{F}^\alpha_{\ \beta}$ up to rotations in the plane orthogonal to u^α, i.e. up to the ambiguity

$$\mathcal{F}^\alpha_{\ \beta} \to \mathcal{F}^\alpha_{\ \beta} + \omega^\alpha_{\ \beta} \tag{5.145}$$

with

$$u^\alpha \omega_{\alpha\beta} = \omega_{\alpha\beta} u^\beta = 0 \quad , \quad \omega_{\alpha\beta} + \omega_{\beta\alpha} = 0 \quad . \tag{5.146}$$

Since there is no such rotation term in the prescription for Fermi-Walker transport, and no natural candidate for it either with only u^α and a^α at one's disposal, it is natural to think of Fermi-Walker transport as a prescription for transporting objects in a non-rotating way.

4. In particular, if one uses the Fermi-Walker prescription to construct an orthonormal basis (E_0, E_k) along the worldline, the spatial vectors can be interpreted as providing a *non-rotating* choice of axes.[14]

5.11 Epilogue: Manifolds? Think Globally, Act Locally!

In section 4.9 I had already briefly discussed some issues regarding the use of indices (and thus in some sense of local coordinates), and had advocated them as a useful bookkeeping device that also provides a transparent way of performing algebraic operations (tensor algebra). In the meantime we have seen that this extends to tensor analysis, and I can only reiterate that for most purposes and in most cases it is much more convenient to

[14] And according to S. Hawking, G. Ellis, *The large-scale structure of space-time*, section 4.1, these "could be realised physically by small gyroscopes pointing in the direction of each vector".

perform calculations in this notation than in some supposedly more elegant index-free notation.

There is one issue, however, that is worth commenting upon, and that in the end actually provides further justification for being allowed to adopt this procedure. Namely, in using local (Cartesian, say) coordinates x^μ to describe a space or space-time (I will use "space" in the following) one is implicitly assuming the following 3 things:

1. first of all, that one can always locally introduce Cartesian coordinates on that space (so as to then be able to perform tensor algebra, tensor analysis etc.);
2. secondly, that different choices of local coordinates will give compatible descriptions of that space;
3. and finally, that in principle one can obtain complete information about the space by covering it with such local coordinate systems.

When these assumptions are satisfied, then one is justified in using local coordinates to describe such a space. The point of this brief section is just to point out that (modulo some topological fine-points) these conditions amount precisely to the definition of a (differentiable or smooth) *manifold* in mathematics.

Thus while I could have started off these notes with an introduction to and definition of smooth manifolds (and numerous textbooks do), for all local intents and purposes this is then really equivalent to (consistently) working in local coordinates, as we have done and will continue to do. It is true that the notion of manifolds, of vector bundles on them etc. becomes indispensable for certain more advanced questions dealing with the global structure of a space-time, or theorems about the existence and uniqueness of solutions to differential equations on some manifold, say, but these are not topics that will be addressed in these notes.

The idea of a manifold is that an n-dimensional manifold is a sufficiently nice topological space that locally looks like (i.e. can be modelled on) the simple and nice topological space $\mathbb{R}^n$, and that this allows one to do calculus on this space by importing the relevant concepts from $\mathbb{R}^n$.

The usual textbook definition of a manifold consists essentially of the following steps:[15]

1. Topological Spaces

   A topological space is a set S together with a collection of subsets U of S (called open sets) which includes S and the empty set, and which is closed under union

[15]This presentation is adapted from the concise and clear description in S. Mukhi, N. Mukunda, *Lectures on Advanced Mathematical Methods for Physicists*, but equivalent descriptions can be found in many other places.

and finite intersection. This set of open sets defines the topology of the space and a corresponding notion of continuous maps (the inverse image of any open set is open) and homeomorphisms (bijective maps ϕ such that both ϕ and ϕ^{-1} are continuous) between topological spaces. In particular there is a notion of continuity for (real-valued, say) functions

$$f: \quad S \to \mathbb{R} \tag{5.147}$$

(with $\mathbb{R}$ equipped with its standard topology).

2. Charts

   However, in this context there is no notion of differentiability or differentiation. In order to have such things at one's disposal one needs topological spaces that locally "look like" $\mathbb{R}^n$. The essential building blocks of such a topological space are "charts":

   A chart C on a topological space S is the pair $C = (U, \phi)$ where $U \subset S$ is an open set of S and ϕ is a homeomorphism

$$U \subset S \to \phi(U) \subset \mathbb{R}^n \quad . \tag{5.148}$$

   The homeomorphism condition implies in particular that $\phi(U)$ is open in $\mathbb{R}^n$. The integer n is then known as the dimension of U (it does not depend on ϕ).

3. Topologial Manifolds

   A topological manifold is a topological space M that is locally homeomorphic to $\mathbb{R}^n$ in the sense that for each point p there is a chart $C = (U, \phi)$ with $p \in U$ (and that satisfies some further topological regularity conditions we are not interested in, such as Hausdorff and usually either second countable or paracompact). Equivalently, a topological space has the structure of a topological manifold when it possesses a covering by open sets U_a with charts $C_a = (U_a, \phi_a)$.

4. Local Coordinates and Local Coordinate Transformations

   The notion of a chart allows (and is equivalent to and the formalisation of) the introduction of local coordinates on the open set $U \subset M$. The coordinates of a point $p \in U$ in this chart are by definition simply the Cartesian coordinates $\vec{x}_p$ of the point $\phi(p) \in \mathbb{R}^n$.

   If one has two charts on M, $C_1 = (U_1, \phi_1)$ and $C_2 = (U_2, \phi_2)$, and $U_1 \cap U_2 \neq \emptyset$, then the "transition functions"

$$\begin{aligned} \phi_1 \circ \phi_2^{-1}: &\quad \phi_2(U_1 \cap U_2) \to \phi_1(U_1 \cap U_2) \\ \phi_2 \circ \phi_1^{-1}: &\quad \phi_1(U_1 \cap U_2) \to \phi_2(U_1 \cap U_2) \end{aligned} \tag{5.149}$$

   are automatically continous maps between open subsets of $\mathbb{R}^n$. These can be interpreted as local coordinate transformations,

$$(\phi_2 \circ \phi_1^{-1})(\vec{x}_p^1) = \phi_2(p) = \vec{x}_p^2 \quad . \tag{5.150}$$

5. Local Functions and Differentiation

   In particular, with the help of charts we can express functions on M in terms of "local oordinates" on $\mathbb{R}^n$. More precisely, given a (continuous) function

$$f: \quad M \to \mathbb{R} \tag{5.151}$$

   and a chart $C = (U, \phi)$, we can associate to the restriction of f to U the function

$$f_U = f \circ \phi^{-1} : \quad \phi(U) \subset \mathbb{R}^n \to \mathbb{R} \ . \tag{5.152}$$

   i.e.

$$p \in U \quad \Rightarrow \quad f(p) = f_U(\vec{x}_p) \ . \tag{5.153}$$

   For such functions on $\mathbb{R}^n$ we now not only have a notion of continuity at our disposal, but also the notions of differentiability, smoothness, differentiation etc.

   On the intersection of 2 charts we can represent the function f in 2 different ways in terms of local coordinates, namely by the functions $f_{U_a} \equiv f_a$ for $a = 1, 2$,

$$\text{on } \ U_1 \cap U_2 : \quad f = f_1 \circ \phi_1 = f_2 \circ \phi_2 \quad \Rightarrow \quad \begin{array}{l} f_2 = f_1 \circ (\phi_1 \circ \phi_2^{-1}) \\ f_1 = f_2 \circ (\phi_2 \circ \phi_1^{-1}) \end{array} \tag{5.154}$$

   This is just the change of variables formula for a function (scalar), namely

$$f_2(\vec{x}_p^2) = f_1(\vec{x}_p^1) \ . \tag{5.155}$$

6. Compatibility of Charts

   In order to be able to extend the notion of smoothness (C^∞-differentiability), say, of a function from a local chart consistently to all of M, we need to impose compatibility conditions on intersecting charts.

   It is evident from (5.154) that the notion of smoothness of a function around a point p will only be independent of the chart if the transition functions $\phi_1 \circ \phi_2^{-1}$ and $\phi_2 \circ \phi_1^{-1}$ (i.e. the coordinate transformations) are also smooth. Thus we define 2 charts to be smoothly compatible if either $U_1 \cap U_2$ is empty or, otherwise, if these maps are smooth.

   Note that for topological manifolds and the condition of continuity any 2 charts are automatically compatible since the transition functions are continuous.

7. Smooth Atlas and Compatibility and Equivalence of Atlases

   A smooth atlas $\mathcal{A}(M)$ of M is now naturally a family of charts $C_a = (U_a, \phi_a)$ which cover M and such that all charts are mutually smoothly compatible.

   2 smooth atlases $\mathcal{A}_1(M)$ and $\mathcal{A}_2(M)$ for the same topological manifold M are said to be compatible with each other if all the charts of $\mathcal{A}_1$ are compatible with all the charts of $\mathcal{A}_2$. This defines an equivalence relation on atlases.

8. Smooth Structure and Smooth Manifold

   A smooth structure on a topological manifold M is an equivalence class

$$\mathcal{S}(M) = [\mathcal{A}(M)] \tag{5.156}$$

   of smooth atlases on M. A smooth manifold is a topological manifold M equipped with a smooth structure $\mathcal{S}$.

9. Smooth Functions and Smooth Maps

   A function

$$f : \quad M \to \mathbb{R} \tag{5.157}$$

   on a smooth manifold M is then said to be smooth if all its local coordinate representatives

$$f_a = f \circ \phi_a^{-1} : \quad \phi(U_a) \to \mathbb{R} \tag{5.158}$$

   are smooth, and a map

$$\mu : \quad M \to N \tag{5.159}$$

   from a smooth manifold M (with charts (U_a, ϕ_a)) of dimension m to a smooth manifold N (with charts $(V_{b'}, \psi_{b'})$) of dimension n is said to be smooth if all of its local coordinate representatives

$$\mu_{ab'} = \psi_{b'} \circ \mu \circ \phi_a^{-1} : \quad \phi_a(U_a) \subset \mathbb{R}^m \to \psi_{b'}(V_{b'}) \subset \mathbb{R}^n \tag{5.160}$$

   are smooth. Such smooth functions can be differentiated by differentiating their local coordinate representatives and mapping the result back to M using the charts (and likewise for maps).

Topological fine-points aside we see that a smooth manifold is by definition a space on which one can consistently do calculus in local coordinates. Hence in these notes we were, are and will be dealing with (smooth) manifolds, regardless of whether or not we state this explicitly.

Analogously one can define C^k-differentiable manifolds (transition functions are required to be of degree C^k), real analytic manifolds (transition functions are required to be real analytic), complex manifolds (modelled on open subsets of $\mathbb{C}^n$, with holomorphic transition functions), etc., as well as submanifolds (modelled on subspaces of $\mathbb{R}^n$), manifolds with boundary (modelled on the half-space $\mathbb{R}^n_+$) etc.

6 Physics in a Gravitational Field and Minimal Coupling

6.1 Principle (or Algorithm) of Minimal Coupling

Recall that the Principle of General Covariance (section 4.1) says that, by virtue of the Einstein Equivalence Principle, a generally covariant equation will be valid in an arbitrary gravitational field provided that it is valid in Minkowski space in inertial coordinates (i.e. in the absence of gravity and/or acceleration).

We now have all the tools at our disposal to construct such equations. In particular, the fact that the covariant derivative ∇ maps tensors to tensors and reduces to the ordinary partial derivative in a locally inertial coordinate system suggests the following procedure or algorithm for obtaining equations that satisfy the Principle of General Covariance:

1. Write down the Lorentz invariant equations or expressions of Special Relativity you are interested in (e.g. those of relativistic mechanics, Maxwell theory, relativistic hydrodynamics, ...) in terms of inertial coordinates ξ^a, the Minkowski metric η_{ab} and other Lorentz tensors $T^{a\cdots}_{b\cdots}$.

2. Replace the coordinates ξ^a by arbitrary coordinates x^μ,
$$\xi^a \mapsto x^\mu \quad . \tag{6.1}$$

3. Wherever the Minkowski metric η_{ab} appears, replace it by the metric $g_{\mu\nu}$ describing the gravitational field,
$$\eta_{ab} \mapsto g_{\mu\nu}(x) \quad . \tag{6.2}$$

4. Promote the Lorentz tensors $T^{a\cdots}_{b\cdots}$ to tensors $T^{\mu\cdots}_{\nu\cdots}$ under general coordinate transformations,
$$T^{a\cdots}_{b\cdots}(\xi) \mapsto T^{\mu\cdots}_{\nu\cdots}(x) \quad . \tag{6.3}$$

5. Wherever a partial derivative $\partial_a = \partial_{\xi^a}$ appears, replace it by the covariant derivative ∇_μ,
$$\partial_a \mapsto \nabla_\mu \quad . \tag{6.4}$$

6. In particular, for the proper-time derivative along a curve this entails replacing $d/d\tau$ by D_τ,
$$\frac{d}{d\tau} \mapsto D_\tau = \dot{x}^\mu \nabla_\mu \quad . \tag{6.5}$$

7. Wherever an integral $\int d^4\xi$ appears, replace it by $\int \sqrt{g} d^4x$,
$$\int d^4\xi \mapsto \int \sqrt{g} d^4x \quad . \tag{6.6}$$

By construction, the resulting equations or expressions are tensorial (generally covariant) and true in the absence of gravity and hence satisfy the conditions for the Principle of General Covariance to apply. As a consequence they will be true in the presence of gravitational fields, at least on scales small compared to those of the gravitational fields. This procedure can thus be regarded as providing us with a prescription how to couple matter (particles, fields) to the gravitational field.

<u>Remarks:</u>

1. This procedure is analogous to the perhaps more familiar "minimal coupling" algorithm for the coupling of matter to gauge fields ("replace partial by gauge covariant derivatives"), and hence also in the current context this procedure is referred to as minimal coupling.

2. The reasons for the "at least on small scales" *caveat* in the paragraph above is that if one considers higher derivatives of the metric tensor then there are other equations that one can write down, involving e.g. the curvature tensor, that are tensorial but reduce to the same equations in the absence of gravity.

3. Thus "minimal coupling", as formulated here, is not a unique and unambiguous description, but it is nevertheless a pragmatic and effective procedure. We will see an example of the ambiguity in the minimal coupling prescription in the discussion of Maxwell theory in a gravitational field in section 6.6, and we will briefly return to the issue in section 8.10.

6.2 Particle Mechanics in a Gravitational Field Revisited

We can see the power of the formalism we have developed so far by rederiving the laws of particle mechanics in a general gravitational field. In Special Relativity (SR), the motion of a free particle with mass m is governed by the equation

$$\text{SR:}\quad a^a = \frac{du^a}{d\tau} = 0 \ , \tag{6.7}$$

where $u^a = d\xi^a/d\tau$ is the 4-velocity and a^a the 4-acceleration. Thus, using the principle of minimal coupling, the equation of motion of a free particle in a general gravitational field is

$$\text{GR:}\quad a^\mu = D_\tau u^\mu = 0 \quad \Leftrightarrow \quad \ddot{x}^\mu + \Gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \ , \tag{6.8}$$

where $u^\mu = dx^\mu/d\tau$. Thus we rediscover the familiar geodesic equation, but we see that it follows much faster from demanding general covariance (as made precise by the principle of minimal coupling) than from our previous somewhat more convoluted and roundabout considerations based e.g. on the equivalence principle.

We could also have arrived at this equation for a free particle in a gravitational field by applying the minimal coupling description not at the level of the equations of motion but rather (and perhaps conceptually more satisfactorily) at the level of the action, i.e. by replacing

$$S = -m \int d\tau = -m \int \sqrt{-\eta_{ab} d\xi^a d\xi^b} \quad \rightarrow \quad -m \int d\tau = -m \int \sqrt{-g_{\mu\nu} dx^\mu dx^\nu} \ , \tag{6.9}$$

and this is exactly what we already did back in section 2.3 where we showed that this also leads to the geodesic equation (6.8).

6.3 Klein-Gordon Scalar Field in a Gravitational Field

Here is where the formalism we have developed really pays off. We will see once again that, using the minimal coupling rule, we can immediately rewrite the equations for a scalar field (here) and the Maxwell equations (in section 6.6 below) in a form in which they are valid in an arbitrary gravitational field.

1. The action for a (real) free massive scalar field ϕ in Special Relativity is

$$\text{SR:} \qquad S[\phi] = \int d^4\xi \left[-\tfrac{1}{2}\eta^{ab}\partial_a\phi\partial_b\phi - \tfrac{1}{2}m^2\phi^2 \right] \ . \tag{6.10}$$

To covariantise this, we replace $d^4\xi \rightarrow \sqrt{g}d^4x$, $\eta^{ab} \rightarrow g^{\alpha\beta}$, and we can replace ∂_a by ∇_α or ∂_α (since this makes no difference on scalars). Therefore, the covariant action in a general gravitational field is

$$\text{GR:} \qquad S[\phi, g_{\alpha\beta}] = \int \sqrt{g}d^4x \left[-\tfrac{1}{2}g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}m^2\phi^2 \right] \ . \tag{6.11}$$

Here I have also indicated the dependence of the action on the metric $g_{\alpha\beta}$. This is not (yet) a dynamical field, though, just the gravitational background field.

2. The equations of motion for ϕ one derives from this are

$$\frac{\delta}{\delta\phi} S[\phi, g_{\alpha\beta}] = 0 \quad \Rightarrow \quad \left(\Box_g - m^2 \right) \phi = 0 \tag{6.12}$$

where $\Box_g = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$, the Laplacian associated to the metric $g_{\alpha\beta}$. This is precisely what one would have obtained by applying the minimal coupling description to the Minkowski Klein-Gordon equation $(\Box_\eta - m^2)\phi = 0$.

Remarks:

(a) A comment on how to derive this: if one thinks of the ∂_α in the action as covariant derivatives, $\partial_\alpha \rightarrow \nabla_\alpha$, then one can use the covariant intgration by parts formula (5.64) to conclude (dropping the boundary term, as usual in variational calculus)

$$-\int \sqrt{g}d^4x\ (\nabla_\alpha\delta\phi)g^{\alpha\beta}\nabla_\beta\phi = \int \sqrt{g}d^4x\ \delta\phi\ g^{\alpha\beta}\nabla_\alpha\nabla_\beta\phi = \int \sqrt{g}d^4x\ \delta\phi\ \Box_g\phi\ \ . \tag{6.13}$$

Thus essentially the calculation is identical to that in Minkowski space. If one sticks with the ordinary partial derivatives, then upon the usual integration by parts one picks up a term $\sim \partial_\alpha(\sqrt{g}g^{\alpha\beta}\partial_\beta\phi)$ which then evidently leads to the Laplacian in the form (5.57).

(b) If the relative sign of $\Box_\eta$ (or $\Box_g$) and m^2 in the Klein-Gordon equation looks unfamiliar to you, then this is probably due to the fact that in a course where you first encountered the Klein-Gordon equation the opposite (particle physicists') sign convention for the Minkowski metric was used, with its negative definite spatial metric.

(c) All of this generalises in a straightforward way to (self-)interacting scalar fields, described by a potential $V(\phi)$. In particular, the action is

$$S[\phi, g_{\alpha\beta}] = \int \sqrt{g}d^4x\left[-\tfrac{1}{2}g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi - V(\phi)\right]\ \ . \tag{6.14}$$

Logically the next thing to discuss would be the energy-momentum tensor, e.g. the minimally coupled counterpart of the special relativistic (Noether) energy-momentum tensor

$$\text{SR:}\quad T_{ab} = \partial_a\phi\partial_b\phi + \eta_{ab}L \tag{6.15}$$

and its properties. However, it turns out that there is more to say about this than meets the eye, and we will therefore return to this issue in more detail in section 7.

6.4 Interlude: General Covariance in Minkowski Space?

Before turning to our next example, I want to briefly comment on the issue of general covariance in Minkowski space, as this tends to generate quite a bit of confusion and unnecessary debates. I will discuss this issue in the context of the above example of a scalar field, but the discussion is valid more generally.

On the one hand, the action (6.10) is generally considered to be invariant (only) under Lorentz or Poincaré transformations, while by construction the action (6.11) is invariant under arbitrary coordinate transformations. Does this really mean that the theory of a scalar field in a non-trivial gravitational background has more invariances than that in a Minkowski background?

On the other hand, certainly nothing prevents one from using e.g. spherical (and thus in particular non-inertial) coordinates in Minkowski space to write down the Klein-Gordon equation or action. But does this mean that the action (6.10) is actually (secretly) invariant also under such non-Lorentz transformations?

Well, that depends ... While this sounds like (and generally is correctly considered to be) a somewhat unsatisfactory answer, I can be more specific:

- it depends on what one means by "invariance" (or "covariance")
- and it depends on how one treats or regards the Minkowski metric.

From the current point of view, the natural answer is that the action (6.11) is generally covariant in *any* gravitational field, in particular therefore also in the absence of a true gravitational field, i.e. in a purely fictitious gravitational field or, equivalently, in Minkowski space. If we specialise the action (6.11) to such a gravitational field, i.e. to the Minkowski metric written in some perhaps non-inertial coordinates, we get

$$S[\phi,\eta_{\alpha\beta}] = \int \sqrt{\eta}d^4x \left[-\tfrac{1}{2}\eta^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}m^2\phi^2\right] \quad . \tag{6.16}$$

Here it is important to keep in mind that $\eta_{\alpha\beta}$ refers to the components of the Minkowski metric in the not-necessarily inertial coordinates x^α, as in

$$\eta_{\alpha\beta} = \frac{\partial\xi^a}{\partial x^\alpha}\frac{\partial\xi^b}{\partial x^\beta}\eta_{ab} \quad . \tag{6.17}$$

As a consequence, also $\sqrt{\eta}$ is not necessarily equal to 1. This action is invariant under arbitrary coordinate transformations, provided that one transforms the fields *and the metric* appropriately.

If one now chooses to write this action in inertial coordinates, $x^\alpha \to \xi^a$, with $\eta_{\alpha\beta} \to \eta_{ab}$ and thus $\sqrt{\eta} \to 1$, then the action (6.16) appears to reduce to the special relativistic action (6.10). So is this action, which is simply a generally covariant action written in some particular coordinates, invariant under Lorentz (or Poincaré) transformations only or under all coordinate transformations?

1. If one looks for the transformations of the coordinates ξ^a and the fields ϕ that leave the action invariant (with fixed metric components η_{ab}) then none too surprisingly one finds that the action is invariant under Poincaré transformations of the coordinates provided that the scalar fields transform as scalars, but not under more general transformations.

2. If one looks for the transformations of the coordinates ξ^a and the fields ϕ *and the metric* η_{ab} that leave the action invariant, then one finds that the action is invariant under arbitrary coordinate transformations

    - provided that one also transforms $\eta_{ab} \to \eta_{\alpha\beta}$ like a (0,2)-tensor

- and provided that one either thinks of $d^4\xi$ as the invariant volume element $\sqrt{\eta}d^4\xi$, or equivalently one treats the Lagrangian L as a scalar density $\sqrt{\eta}L$.

Sometimes option (1) is taken to define the *invariance group* (Poincaré transformations) while option (2) refers to the *covariance group*. In this sense, special relativity is *invariant* under Poincaré transformations but is at the same time *generally covariant*. In philosophy of science or epistemological terms whether one has option (1) or option (2) is related to the question whether or not the Minkowski metric is regarded as an *absolute element* of the theory. With η_{ab} promoted to an absolute element, general covariance is reduced to Poincaré invariance (those transformations that, from the generally covariant "$\eta_{\alpha\beta}$ transforms" point of view, leave η_{ab} invariant).[16] Unfruitful discussions ensue when tacitly conflicting assumptions are made about what are considered to be the absolute elements of a theory.

6.5 Lorentz-Covariant Formulation of Maxwell Theory (Review)

In order to discuss the formulation of Maxwell theory in a gravitational field, we will need to quickly recall the Lorentz-covariant formulation of Maxwell theory in Minkowski space. This will also fix our conventions for Maxwell theory.

In the traditional non-covariant formulation one has

1. the homogeneous equations

$$\vec{\nabla}\cdot\vec{B} = 0 \quad , \quad \vec{\nabla}\times\vec{E} + \partial_t\vec{B} = 0 \tag{6.18}$$

2. the inhomogeneous equations

$$\vec{\nabla}\cdot\vec{E} = \rho/\epsilon_0 \quad , \quad \vec{\nabla}\times\vec{B} - \frac{1}{c^2}\partial_t\vec{E} = \mu_0\vec{J} \tag{6.19}$$

3. the ensuing continuity equation

$$\partial_t\rho + \vec{\nabla}.\vec{J} = 0 \tag{6.20}$$

4. the vector and scalar potentials $\vec{A}$ and ϕ,

$$\vec{B} = \vec{\nabla}\times\vec{A} \quad , \quad \vec{E} = -\vec{\nabla}\phi - \partial_t\vec{A} \tag{6.21}$$

5. and the corresponding gauge transformations leaving $\vec{E}$ and $\vec{B}$ invariant,

$$\vec{A}\to\vec{A}+\vec{\nabla}\Psi \quad , \quad \phi\to\phi-\partial_t\Psi \quad \Rightarrow \quad \vec{E}\to\vec{E} \quad , \quad \vec{B}\to\vec{B} \ . \tag{6.22}$$

[16] For more erudite discussions of these and related issues, see e.g. section 2.5 ("On Covariance and Invariance") of N. Straumann, *General Relativity* or D. Giulini, *Some remarks on the notions of general covariance and background independence*, `arXiv:gr-qc/0603087`.

The charge density and current can be packaged into a Lorentz vector

$$J^a = (c\rho, \vec{J}) \tag{6.23}$$

(note that in signature (-+++) one has to choose whether to identify J^0 or $J_0 = -J^0$ with the charge density, here we choose the former), and the continuity equation can be written in the manifestly Lorentz-invariant form

$$\partial_a J^a = 0 \quad . \tag{6.24}$$

Likewise, the scalar and vector potential can be packaged into a Lorentz covector

$$A_a = (-\phi/c, \vec{A}) \quad , \tag{6.25}$$

and the gauge transformations can be compactly written as

$$A_a \to A_a + \partial_a \Psi \quad . \tag{6.26}$$

The gauge invariant Maxwell field strength tensor F_{ab} is defined by

$$F_{ab} = \partial_a A_b - \partial_b A_a \tag{6.27}$$

and has the components

$$F_{0k} = -F_{k0} = -E_k/c \quad , \quad F_{ik} = \epsilon_{ik\ell}\, B_\ell \tag{6.28}$$

or, in matrix form,

$$(F_{ab}) = \begin{pmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ +E_1/c & 0 & +B_3 & -B_2 \\ +E_2/c & -B_3 & 0 & +B_1 \\ +E_3/c & +B_2 & -B_1 & 0 \end{pmatrix} \tag{6.29}$$

and

$$(F^{ab}) = \begin{pmatrix} 0 & +E_1/c & +E_2/c & +E_3/c \\ -E_1/c & 0 & +B_3 & -B_2 \\ -E_2/c & -B_3 & 0 & +B_1 \\ -E_3/c & +B_2 & -B_1 & 0 \end{pmatrix} \tag{6.30}$$

In terms of these Lorentz tensors, the homogeneous Maxwell equations can be written as

$$\partial_{[a} F_{bc]} = 0 \quad \Leftrightarrow \quad \partial_a F_{bc} + \partial_c F_{ab} + \partial_b F_{ca} = 0 \quad , \tag{6.31}$$

and these equations are identically satisfied if F_{ab} derives from a potential,

$$F_{ab} = \partial_a A_b - \partial_b A_a \quad \Rightarrow \quad \partial_{[a} F_{bc]} = 0 \quad . \tag{6.32}$$

The inhomogeneous Maxwell equations can (suppressing μ_0, i.e. setting $\mu_0 = 1$) be written as

$$\partial_a F^{ab} = -J^b \quad \Leftrightarrow \quad \Box A_a - \partial_a(\partial_b A^b) = -J_a \quad . \tag{6.33}$$

These equations can be derived from the Lorentz-invariant action

$$S[A,J] = S[A] + S_I[A,J] = -\tfrac{1}{4}\int d^4\xi\; F_{ab}F^{ab} + \int d^4\xi\; A_a J^a \quad , \tag{6.34}$$

with the Maxwell Lagrangian

$$-\tfrac{1}{4}F_{ab}F^{ab} = -\tfrac{1}{2}F_{0k}F^{0k} - \tfrac{1}{4}F_{ik}F^{ik} = \tfrac{1}{2}(\vec{E}^2/c^2 - \vec{B}^2) \quad . \tag{6.35}$$

This is essentially all we will need (some facts regarding the Noether versus covariant energy-momentum tensor of Maxwell theory will be recalled below).

6.6 Maxwell Theory in a Gravitational Field

Mutatis mutandis we can now proceed in the same way as for a scalar field.

1. The basic dynamical field is the vector potential A_a. Given the vector potential A_μ, the Maxwell field strength tensor in Special Relativity is

$$\text{SR:}\quad F_{ab} = \partial_a A_b - \partial_b A_a \quad . \tag{6.36}$$

   Therefore in a general metric space-time (gravitational field) one is led to (or tempted to) define the field strength tensor as

$$\text{GR:}\quad F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \tag{6.37}$$

   - A cautionary remark: Actually, this is a bit misleading. The field strength tensor (two-form) in any, Abelian or non-Abelian, gauge theory is always given in terms of the gauge-covariant exterior derivative of the vector potential (i.e. it is the curvature of the connection), and as such has nothing whatsoever to do with a metric on space-time. So you should not really regard the first equality in the above equation as the definition of $F_{\mu\nu}$, but you should regard the second equality as a proof that $F_{\mu\nu}$, *always* defined by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, is a tensor.

     The mistake of adopting $\nabla_\mu A_\nu - \nabla_\nu A_\mu$ as the definition of $F_{\mu\nu}$ in a curved space-time has led some poor souls to believe, and even claim in published papers, that in a space-time with torsion, for which the second equality does not hold, the Maxwell field strength tensor is invariably modified by the torsion. This is nonsense.

     If there is torsion, one is of course free to consider non-minimal couplings of the torsion tensor to other tensor fields (like the Maxwell field strength tensor), but this is not required by either gauge invariance or general covariance.

2. In Special Relativity, the Maxwell equations read

$$\begin{aligned} \text{SR:} \qquad & \partial_a F^{ab} = -J^b \\ & \partial_{[a}F_{bc]} = 0 \ . \end{aligned} \tag{6.38}$$

Thus in a general gravitational field (curved space-time) these equations become

$$\begin{aligned} \text{GR:} \quad & \nabla_\mu F^{\mu\nu} = -J^\nu \\ & \nabla_{[\mu}F_{\nu\lambda]} = 0 \ , \end{aligned} \tag{6.39}$$

where now of course all indices are raised and lowered with the metric $g_{\mu\nu}$,

$$F^{\mu\nu} = g^{\mu\lambda}g^{\nu\rho}F_{\lambda\rho} \ . \tag{6.40}$$

<u>Remarks:</u>

(a) Regarding the use of the covariant derivative in the second equation, the same *caveat* as above applies.

(b) In particular, using the results derived in section 5.5, we can rewrite these two equations as

$$\begin{aligned} \text{GR:} \qquad & \partial_\mu(\sqrt{g}F^{\mu\nu}) = -\sqrt{g}J^\nu \\ & \partial_{[\mu}F_{\nu\lambda]} = 0 \ . \end{aligned} \tag{6.41}$$

(c) It is clear from the first of these equations that the Maxwell equations imply that the current is covariantly conserved: since

$$\partial_\nu\partial_\mu(\sqrt{g}F^{\mu\nu}) = 0 \tag{6.42}$$

by anti-symmetry of $F^{\mu\nu}$, it follows that

$$\partial_\nu(\sqrt{g}J^\nu) = 0 \quad \Leftrightarrow \quad \nabla_\nu J^\nu = 0 \ . \tag{6.43}$$

From the covariant version $\nabla_\mu F^{\mu\nu} = -J^\nu$ this follows in the seemingly more roundabout way from the identity (8.58) for the commutator of covariant derivatives that we will establish later, in the context of our discussion of the Riemann curvature tensor in section 8.

(d) In Special Relativity, the inhomogeneous Maxwell equations can be decoupled by imposing the Loren(t)z (see footnote 65 in section 23.5) gauge condition $\partial_a A^a = 0$,

$$\partial_a A^a = 0 \quad \Rightarrow \quad \partial_a F^{ab} = -J^b \rightarrow \Box A_b = -J_b \ . \tag{6.44}$$

This gauge condition has the virtue of preserving Lorentz invariance. Similarly, its covariantised version

$$\nabla_\mu A^\mu = 0 \tag{6.45}$$

has the virtue of preserving general covariance, because $\nabla_\mu A^\mu$ is a scalar. However, the inhomogeneous Maxwell equations in this gauge do not take the form $\Box A_\nu = -J_\nu$ one might perhaps have anticipated on the basis of minimal coupling. Rather, using the covariant Lorenz gauge condition (6.45), the covariant divergence of the Maxwell field strength tensor can be written as

$$\nabla_\mu A^\mu = 0 \quad \Rightarrow \quad \nabla_\mu F^{\mu\nu} = \nabla_\mu(\nabla^\mu A^\nu - \nabla^\nu A^\mu) = \Box A^\nu - [\nabla_\mu, \nabla_\nu] A^\mu \;\; , \tag{6.46}$$

where $\Box A_\mu = \nabla^\nu \nabla_\nu A_\mu$ is the "naive" Laplacian on scalars. The second term would of course be zero in Minkowski space, but here it is not. Indeed, as we will see in section 8, the quintessence of a non-trivial geometry is that covariant derivatives do not commute on tensors other than scalars. In particular, here one finds that as a consequence of (8.51) the Maxwell equations in the covariant Lorenz gauge can be written as

$$\Box A^\nu - R^\nu_{\;\mu} A^\mu = -J^\nu \;\; , \tag{6.47}$$

where $R_{\mu\nu}$ is the *Ricci tensor*, a particular contraction of the Riemann curvature tensor, constructed from the second derivatives of the metric.

(e) Thus, these equations appear to display a *non-minimal coupling to the gravitational field*, even though we started off with the minimally coupled equations which we can also derive, see below, from the minimally coupled action. We will return to a discussion of this issue in section 8.10.

3. The electromagnetic force acting on a particle of charge e is given in Special Relativity by the Lorentz force

$$\text{SR:} \quad f^a = eF^a_{\;b}\dot{\xi}^b \;\; . \tag{6.48}$$

Thus in General Relativity it becomes

$$\text{GR:} \quad f^\mu = eg^{\mu\lambda} F_{\lambda\nu}\dot{x}^\nu \;\; . \tag{6.49}$$

4. The Lorentz-invariant action of (vacuum) Maxwell theory is

$$\text{SR:} \quad S[A_a] = -\tfrac{1}{4}\int d^4\xi \; F_{ab}F^{ab} \tag{6.50}$$

in Special Relativity, and thus becomes

$$\text{GR:} \quad S[A_\alpha, g_{\alpha\beta}] = -\tfrac{1}{4}\int \sqrt{g} d^4x \; F_{\mu\nu}F^{\mu\nu} \equiv -\tfrac{1}{4}\int \sqrt{g} d^4x \; g^{\mu\lambda}g^{\nu\rho}F_{\mu\nu}F_{\lambda\rho} \tag{6.51}$$

in General Relativity.

As for the scalar field, depending on whether one writes the field strength tensor as $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ or as $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$, by varying this action with respect to the A_μ one obtains the vacuum Maxwell equations $\nabla_\mu F^{\mu\nu} = 0$ in either of the 2 forms

$$\frac{\delta}{\delta A_\nu} S[A_\alpha, g_{\alpha\beta}] = 0 \quad \Rightarrow \quad \begin{cases} \partial_\mu(\sqrt{g} F^{\mu\nu}) = 0 \\ \nabla_\mu F^{\mu\nu} = 0 \end{cases} \tag{6.52}$$

<u>Remarks:</u>

(a) Writing out explicitly the Lagrangian in terms of its components (with respect to some coordinate system $x^\alpha = (t, x^k)$) one finds

$$\begin{aligned} -\tfrac{1}{4} F_{\alpha\beta} F^{\alpha\beta} = &- \tfrac{1}{2} F_{0i} F_{0k} (g^{00} g^{ik} - g^{0i} g^{0k}) \\ &- \tfrac{1}{4} F_{ij} F_{kl} g^{ik} g^{jl} \\ &- \tfrac{1}{2} F_{0i} F_{kl} (g^{0k} g^{il} - g^{0l} g^{ik}) \end{aligned} \tag{6.53}$$

While the 1st and 2nd lines look just like "gravitationally dressed" standard terms $\sim \vec{E}^2$ and $\sim \vec{B}^2$, the last line appears to suggest a gravitationally induced coupling between the electric and magnetic fields. This, however, is misleading and simply not a meaningful way of expressing things. After all, even in Minkowski space the decomposition of the electro-magnetic field into electric and magnetic fields depends on the choice of inertial reference system.

(b) In order to add sources, one can add $\int \sqrt{g} d^4x \; A_\mu J^\mu$ to the Maxwell action, thus coupling the matter current to the Maxwell gauge field. Instead of just adding such a (phenomenological) source-term by hand, a more coherent microscopic approach (which also provides the sources with their own dynamics) is to consider a matter action (minimally) coupled to the Maxwell field,

$$S_M[\phi] \rightarrow S_M[\phi, A_\alpha] \quad . \tag{6.54}$$

The combined Maxwell + matter action will then give rise to the Maxwell equations with a source provided that one defines the current J^α as the variation of the matter action with respect to the gauge field,

$$J^\alpha \sim \frac{\delta S_M[\phi, A_\alpha]}{\delta A_\alpha} \quad . \tag{6.55}$$

As in the case of scalar fields, we will postpone a discussion of the energy-momentum tensor and how to properly define it (something that is already an issue in Minkowski space because for Maxwell theory the Noether energy-momentum tensor turns out to be neither symmetric nor gauge-invariant!) to section 7.

In anticipation of this I just want to point out that, by the same rationale as that leading to (6.55), perhaps we should define the source term for the gravitational field by the variation of the gravitationally minimally coupled matter action with respect to the metric. If we now call this source term the energy-momentum tensor, then we have a candidate definition of the energy-momentum tensor which is natural and appropriate from the gravitational point of view. We will pursue this point of view in section 7.6.

6.7 Minimal Coupling and (quasi-)Topological Couplings

In all the cases considered so far, the minimal coupling prescription resulted in a minimally coupled matter action that depends explicitly on the metric - this is as it should be and is not a surprise. What would be more of a surprise would be to find minimally coupled and hence generally covariant contributions to an action that do not depend on the metric, but such examples do indeed exist (and play an important role in many branches of physics and even mathematics, ranging from the strong-CP problem in QCD to high-T_c superconductors to topology). Such terms in the action are usually referred to as "topological terms" in the physics literature but as they need not be (and usually are not) purely topological in the mathematics sense, for lack of a better name I refer to them as "quasi-topological".

Here are 2 prototypical examples illustrating this phenomenon:

1. Axionic Coupling in (3+1) Dimensions

   The first toy-model we will consider consists of Maxwell-theory coupled to a neutral scalar field through what is known as an axionic coupling only (with analogous considerations for the more interesting case of a non-Abelian Yang-Mills field),

   $$S[\phi, A] = S_s[\phi] + S_m[A] + S_a[\phi, A] \ , \tag{6.56}$$

   with

   $$S_s[\phi] = \int d^4x \; L_s(\phi, \partial_\alpha \phi) \tag{6.57}$$

   some arbitrary standard scalar field action (of the type already discussed), $S_m[A]$ the usual Maxwell action,

   $$S_m[A] = \int d^4x \; L_m(\partial_\alpha A_\beta) = -\tfrac{1}{4}\int d^4x \; F^{\alpha\beta}F_{\alpha\beta} \tag{6.58}$$

   and the axionic coupling term is

   $$S_a[\phi, A] = -\tfrac{1}{4}\int d^4x f(\phi)\tilde{F}^{\alpha\beta}F_{\alpha\beta} \tag{6.59}$$

   where

   $$\tilde{F}^{\alpha\beta} = \tfrac{1}{2} \in^{\alpha\beta\gamma\delta} F_{\gamma\delta} \tag{6.60}$$

with $\in^{\alpha\beta\gamma\delta} = 0, \pm 1$ the Levi-Civita symbol, a tensor density (cf. remark 2 in section 4.5) and $f(\phi)$ some function of the scalar field ϕ. Note that for $f(\phi) = 1$ (or in the absence of a scalar field) the axionic term would be (locally) a total derivative and would hence not contribute to the equations of motion. For non-trivial $f(\phi)$, on the other hand, the axionic term is itself non-trivial.

Minimal coupling for the first two (standard) terms proceeds as already discussed above. For the third, axionic term, we make the usual replacement $d^4x \to \sqrt{g}d^4x$ and recall from (4.83) that

$$\epsilon^{\alpha\beta\gamma\delta} \equiv \frac{1}{\sqrt{g}} \in^{\alpha\beta\gamma\delta} \tag{6.61}$$

is a (4,0) tensor, so that the generally covariant generalisation of the axionic action is

$$\begin{aligned} S_a[\phi, A, g_{\alpha\beta}] &= -\tfrac{1}{8} \int \sqrt{g}d^4x f(\phi)\epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta}F_{\alpha\beta} \\ &= -\tfrac{1}{8} \int d^4x f(\phi) \in^{\alpha\beta\gamma\delta} F_{\gamma\delta}F_{\alpha\beta} = S_a[\phi, A] \end{aligned} \tag{6.62}$$

We see that, as announced, the metric dependence drops out of the minimally coupled generally covariant action. The reason for this is that the axionic Lagrangian is already all by itself a scalar density of weight $w = +1$, and that therefore its integral (4.76) is well-defined and generally covariant without having to take recourse to a metric to construct an auxiliary weight-one object like $\sqrt{g}$.

2. Maxwell - Chern-Simons Theory in (2+1) Dimensions

   The second prominent example involves the addition of what is known as an Abelian Chern-Simons term to the Maxwell action in (2+1) dimensions (with analogous considerations for the more interesting case of a non-Abelian Yang-Mills field). The Minkowski space Lagrangian of this model is

$$L = L_m + k\; L_{cs} = -\tfrac{1}{4}F^{\alpha\beta}F_{\alpha\beta} + \tfrac{1}{2}k \;\in^{\alpha\beta\gamma} A_\alpha F_{\beta\gamma} \;\; . \tag{6.63}$$

   Minimal coupling for the first term is standard and for the 2nd term one finds, as above, that the generally covariant minimally coupled Chern-Simons action is actually metric independent (since the Chern-Simons Lagrangian is a density of weight $w = 1$),

$$\begin{aligned} S_{cs}[A, g_{\alpha\beta}] &= \tfrac{1}{2}k \int \sqrt{g}d^3x\; \epsilon^{\alpha\beta\gamma} A_\alpha F_{\beta\gamma} \\ &= \tfrac{1}{2}k \int d^3x \;\in^{\alpha\beta\gamma} A_\alpha F_{\beta\gamma} = S_{cs}[A] \end{aligned} \tag{6.64}$$

   As an aside note that the above theory is also known as topologically massive Maxwell theory, since the CS term provides a gauge-invariant mass term for the photon. One quick way to see this is to note that the equations of motion are

$$\partial_\alpha F^{\alpha\beta} + k \;\in^{\beta\gamma\delta} F_{\gamma\delta} = 0 \;\; . \tag{6.65}$$

and that in terms of the dual field strength

$$G^{\beta} = \tfrac{1}{2} \in^{\beta\gamma\delta} F_{\gamma\delta} \tag{6.66}$$

the equations of motion and the Bianchi identity take the form

$$\partial_\alpha G_\beta - \partial_\beta G_\alpha = 2k \ \in_{\alpha\beta\gamma} G^\gamma \quad , \quad \partial_\beta G^\beta = 0 \tag{6.67}$$

respectively. Acting with ∂^α on the equation of motion and using the Bianchi identity and again the equation of motion one finds

$$\begin{aligned} \Box G_\beta &= 2k \ \in_{\alpha\beta\gamma} \partial^\alpha G^\gamma = k \ \in_{\alpha\beta\gamma} (\partial^\alpha G^\gamma - \partial^\gamma G^\alpha) \\ &= 2k^2 \ \in_{\alpha\beta\gamma} \in^{\alpha\gamma\delta} \ G_\delta = 4k^2 G_\beta \end{aligned} \tag{6.68}$$

so that the theory describes excitations of mass $m^2 = 4k^2$.

These quasi-topological terms modify the equations of motion. Moreover, since they depend on the derivatives of the fields, they will contribute to the canonical Noether energy-momentum tensor. On the other hand, since they do not depend on the metric, they do not contribute to the covariant energy-momentum tensor, defined in section 7 in terms of the variation of the matter action with respect to the metric (and as such playing the role of the source term for the Einstein gravitational field equations).

How it nevertheless conspires that this tensor is conserved on-shell (meaning: for a solution to the matter equations of motion) even though the equations of motion have been modified and how the improved canonical energy-momentum tensor nevertheless ends up agreeing with the covariant energy-momentum tensor on-shell will be explored and explained in section 22.5.

6.8 Conserved Charges from Covariantly Conserved Currents

In Special Relativity a conserved current J^a is characterised by the vanishing of its divergence, i.e. by $\partial_a J^a = 0$. It leads to a conserved charge Q by integrating J^a over a spacelike hypersurface, say the one described by $t = t_0$,

$$Q = \int_{t=t_0} d^3x \ J^0 \quad . \tag{6.69}$$

That Q is conserved, i.e. independent of t_0, is a consequence of the fact that by virtue of the Gauss theorem

$$Q(t_1) - Q(t_0) = \int_V d^4\xi \ \partial_a J^a = 0 \ , \tag{6.70}$$

where V is the four-volume $\mathbb{R}^3 \times [t_0, t_1]$. This holds provided that J vanishes at spatial infinity.

Now in General Relativity, the conservation law will be replaced by the covariant conservation law $\nabla_\mu J^\mu = 0$, and one may wonder if this also leads to some conserved charges in the ordinary sense. The answer is yes because, recalling the formula for the covariant divergence of a vector,

$$\nabla_\mu J^\mu = g^{-1/2}\partial_\mu(g^{1/2}J^\mu) \quad , \tag{6.71}$$

we see that

$$\nabla_\mu J^\mu = 0 \Leftrightarrow \partial_\mu(g^{1/2}J^\mu) = 0 \quad , \tag{6.72}$$

so that $g^{1/2}J^\mu$ is a conserved current in the ordinary sense. We then obtain conserved quantities in the ordinary sense by integrating J^μ over a spacelike hypersurface Σ. We will develop a more precise formula for this, an appropriate version of the Gauss theorem for hypersurfaces in curved space-times, in section 16.3.

The factor $g^{1/2}$ apearing in the current conservation law can be understood physically. To see what it means, split J^μ into its space-time direction u^μ, with $u^\mu u_\mu = -1$, and its magnitude ρ as

$$J^\mu = \rho u^\mu \quad . \tag{6.73}$$

This defines the average four-velocity of the conserved quantity represented by J^μ and its density ρ measured by an observer moving at that average velocity (rest mass density, charge density, number density, ...). Since u^μ is a vector, in order for J^μ to be a vector, ρ has to be a scalar. Therefore this density is defined as per unit proper volume. The factor of $g^{1/2}$ transforms this into density per coordinate volume and this quantity is conserved (in a comoving coordinate system where $J^0 = \rho, J^i = 0$).

We will come back to this in the context of cosmology later on in this course, but for now just think of the following picture (Figure 44 in section 34): take a balloon, draw lots of dots on it at random, representing particles or galaxies. Next choose some coordinate system on the balloon and draw the coordinate grid on it. Now inflate or deflate the balloon. This represents a time dependent metric, roughly of the form $ds^2 = r^2(t)(d\theta^2 + \sin^2\theta d\phi^2)$. You see that the number of dots per coordinate volume element (area element in this case) does not change, whereas the number of dots per unit proper volume (area) will.

7 Energy-Momentum Tensor I: Basics

7.1 Introduction

Newton's gravitational field equation for the gravitational potential ϕ is the Poisson equation $\Delta\phi = 4\pi G_N \mu$, with μ the mass density. Thus in Newton's theory, mass is the source of gravity. We can also more usefully, and thinking relativistically, write this in terms of the energy density $\rho = \mu c^2$ as

$$\Delta\phi = \frac{4\pi G_N}{c^2}\rho \overset{(c=1)}{=} 4\pi G_N\, \rho \quad . \tag{7.1}$$

Now we already noted in section 1.1 that in Special Relativity ρ is not a scalar but rather just one component of a tensor, the energy-momentum tensor

$$T_{ab}: \quad T_{00} = \rho \quad , \tag{7.2}$$

with the components T_{ab} transforming into each other under Lorentz transformations according to the transformation rules for Lorentz tensors.

It is therefore entirely plausible that in a relativistic theory of gravity the source of gravity should be the entire energy-momentum tensor. In particular, also the other components of T_{ab}, T_{0k} ($\sim$ energy flux), T_{k0} ($\sim$ momentum density) and T_{ik} ($\sim$ stresses or pressure) are a source of gravity. Clearly, therefore, the notion of energy-momentum tensor will play a crucial role in the following. This then immediately raises the question how to find or define an energy-momentum tensor.

Within the framework of special relativity and relativistic field theories there are (at least) 2 common approaches to constructing or defining an energy-momentum tensor, namely

1. a Macroscopic Phenomenological Description
2. a Microscopic Lagrangian Prescription

and we will now briefly discuss these in turn.

7.2 Perfect Fluid Energy-Momentum Tensor in Special Relativity

A macroscopic phenomenological description is useful when one does not know (or does not care about) the microscopic description of the matter one is dealing with but rather tries to characterise its properties in terms of the specification of some macroscopic (thermodynamic, hydrodynamic) parameters such as energy (density), pressure, viscosity etc. For many purposes this is the appropriate language for describing e.g. gases or fluids.

In this case, one constructs the energy-momentum tensor in such a way that it encodes the physics one is trying to describe (primarily conservation laws and dynamics). As a simple example of this (not by coincidence the one which is of most relevance for gravitational physics and thus also later on in these notes), we consider a *perfect fluid.*

By definition, a perfect fluid is one in which a comoving observer (i.e. an observer in a local rest-frame of the fluid) sees the fluid around him as isotropic (rotation-invariant). This means that in this reference system the components of the energy-momentum tensor have the form (any non-zero T_{0k} would break rotation invariance, and δ_{ik} is the unique rotation-invariant symmetric $(0,2)$-tensor)

$$T_{00} = \rho \quad , \quad T_{0k} = 0 \quad , \quad T_{ik} = p\delta_{ik} \quad . \tag{7.3}$$

Here ρ and p are any functions of the coordinates, interpreted as the energy density and the pressure of the fluid .

To specify the kind of fluid one is working with, one should supplement this by an *equation of state* which provides a relation between ρ and p. Typically this amounts to specifying p as a function of ρ,

$$\text{Equation of State:} \quad p = p(\rho) \tag{7.4}$$

(and possibly other parameters).

In terms of the 4-velocity u^a of the fluid, which in the local rest frame has the components

$$u^a = (1,0,0,0) \quad , \tag{7.5}$$

one can combine the components of the energy-momentum tensor into the expression,

$$T_{ab} = (\rho + p)u_a u_b + p\eta_{ab} \quad , \tag{7.6}$$

(note that energy density and pressure = force per unit area have the same dimensions). As this is now a tensorial equation it is now valid in any inertial system. It defines the energy-momentum tensor of a perfect fluid. The conditions

$$\partial^a T_{ab} = 0 \tag{7.7}$$

imply a continuity equation and (as we will see below) a relativistic generalisation of the Euler equations for a perfect fluid. These are usually supplemented by a further continuity equation for the fluid density current

$$j^a = nu^a \tag{7.8}$$

with n e.g. the number density or particle density, say, namely

$$\partial_a j^a = 0 \quad . \tag{7.9}$$

Now let us look at the consequences of these equations. Since

$$u^a u_a = -1 \quad \Rightarrow \quad (\partial_a u_b)u^b = \partial_a(u_b u^b)/2 = 0 \ \ , \tag{7.10}$$

the u-component of (7.7) can be written as

$$(\partial^a T_{ab})u^b = 0 \quad \Leftrightarrow \quad u^a\partial_a\rho + (\rho + p)\partial_a u^a = 0 \ \ . \tag{7.11}$$

With the help of the current conservation equation, this equation can be recast into the form

$$\begin{aligned}
0 &= u^a\partial_a\rho + (\rho + p)\partial_a(j^a/n) \\
&= u^a\partial_a\rho + (\rho + p)j^a\partial_a(1/n) \\
&= u^a[\partial_a\rho + (\rho + p)n\partial_a(1/n)] \\
&= nu^a[p\partial_a(1/n) + \partial_a(\rho/n)] \ \ .
\end{aligned} \tag{7.12}$$

The point of rewriting the equation in this way is that (assuming a situation of thermodynamic equilibrium) the 2nd law of thermodynamics says that pressure p, energy density ρ and the volume per particle $(1/n)$ are related by

$$Tds = pd(1/n) + d(\rho/n) \ \ . \tag{7.13}$$

where T is the temperature and s the specific entropy, i.e. the entropy per particle.[17] Thus the above equation says that the specific entropy s is constant along the flow,

$$u^a\partial_a s = 0 \ \ . \tag{7.14}$$

The significance of the spatial (transverse to u) components of (7.7) is easier to decipher if one writes the equations non-covariantly by setting

$$u^0 = \gamma(v) \quad , \quad u^i = \gamma(v)v^i = v^i u^0 \ \ , \tag{7.15}$$

so that

$$u^a\partial_a = \gamma(v)(\partial_t + \vec{v}.\vec{\nabla}) \tag{7.16}$$

is ($\gamma(v)$ times) the usual *convective derivative* or comoving time-derivative, and the above equation for the conservation of the specific entropy can be written as

$$(\partial_t + \vec{v}.\vec{\nabla})s = 0 \ \ . \tag{7.17}$$

Moreover, the continuity equation for the current j^a with components

$$j^0 = \gamma(v)n \quad , \quad j^i = \gamma(v)nv^i \tag{7.18}$$

becomes

$$\partial_t(\gamma(v)n) + \vec{\nabla}.(\gamma(v)n\vec{v}) = 0 \ \ , \tag{7.19}$$

[17] See e.g. J. van Holten, *Relativistic Fluid Dynamics*, `http://www.nikhef.nl/~t32/relhyd.pdf` for a derivation of this and further discussion.

and the time-component of (7.7) can be written as

$$\partial_t(p - \gamma(v)^2(\rho + p)) - \vec{\nabla}.[\gamma(v)^2(\rho + p)\vec{v}] = 0 \ . \tag{7.20}$$

Using this equation the spacelike components of (7.7) can then be written as

$$\gamma(v)^2(\rho + p)(\partial_t\vec{v} + \vec{v}.\vec{\nabla}\vec{v}) + \vec{v}\partial_t p + \vec{\nabla}p = 0 \ . \tag{7.21}$$

In a suitable non-relativistic limit ($v \ll 1, p \ll \rho$), this latter equation reduces to

$$\rho(\partial_t\vec{v} + \vec{v}.\vec{\nabla}\vec{v}) + \vec{\nabla}p = 0 \ , \tag{7.22}$$

which is the non-relativistic Euler equation for a perfect fluid.

As we will see, it is straightforward to promote such a perfect fluid energy-momentum tensor by minimal coupling to the energy-momentum tensor describing a perfect fluid in a gravitational field, and we will come back to this below.

For a covariant rendition and elementary covariant derivation of the ensuing equations of motion in a general gravitational field from the conservation of the energy-momentum tensor, see e.g. the derivation of (35.81) and (35.82) in sections 35.4 and 35.5.

7.3 Noether Energy-Momentum Tensor in Special Relativity (Review)

A microscopic Lagrangian description is the method of choice when one has a Poincaré-invariant Lagrangian field theory description of the matter one is trying to describe. In particular, this applies to the scalar and Maxwell field theories we have already discussed and, more generally, to the modern microscopic and action-based description of the fundamental interactions of particle physics.

In this case, there is a canonical procedure for constructing an energy-momentum tensor, namely from Noether's theorem applied to translations, resulting in what is then appropriately known as the *Noether energy-momentum tensor* or the *canonical energy-momentum tensor* Θ_{ab}.

For a Lagrangian $L = L(\phi, \partial_a\phi)$ depending on some fields ϕ and their 1st derivatives (these could be scalar, vector, ... fields), this tensor is defined by

$$\Theta^a_{\ b} = -\frac{\partial L}{\partial(\partial_a\phi)}\partial_b\phi + \delta^a_{\ b}L \tag{7.23}$$

(sign conventions are such that Θ_{00} rather than $\Theta^0_{\ 0}$ is the energy density). It is built from the 4 Noether currents

$$\Theta^a_{\ b} \equiv J^a_{(b)} \tag{7.24}$$

associated to translation invariance in the x^b-direction, $\delta_{(b)}\phi = \partial_b\phi$. By calculating its divergence, one finds

$$\partial_a\Theta^a_{\ b} = \frac{\delta L}{\delta\phi}\partial_b\phi \ , \tag{7.25}$$

where $\delta L/\delta\phi$ is the Euler-Lagrange variational derivative,

$$\frac{\delta L}{\delta \phi} = \frac{\partial L}{\partial \phi} - \partial_a \frac{\partial L}{\partial(\partial_a \phi)} \quad . \tag{7.26}$$

Thus Θ_{ab} is on-shell (meaning: for a solution to the matter equations of motion) conserved,

$$\partial_a \Theta^a_{\ b} = 0 \quad \text{on-shell} \quad , \tag{7.27}$$

and leads to the conserved energy-momentum 4-vector

$$P_b = \int d^3x \; J^0_{(b)} = \int d^3x \; \Theta^0_{\ b} \quad . \tag{7.28}$$

This procedure and prescription is perfectly adequate and sufficient for scalar (spin 0) fields, but it turns out to be far from satisfactory and far from the end of the story for other fields (e.g. for Maxwell theory, for which Θ_{ab} turns out to be neither symmetric nor gauge invariant). In this more general situation one is then required to "improve" this prescription in order to obtain an energy-momentum tensor T_{ab} with the desired properties.

As a first example where everything works out nicely, consider the energy-momentum tensor of a Klein-Gordon scalar field in Minkowski space. In this case,

$$\Theta_{ab} = \partial_a\phi\partial_b\phi + \eta_{ab}L = \partial_a\phi\partial_b\phi - \tfrac{1}{2}\eta_{ab}\left(\eta^{cd}\partial_c\phi\partial_d\phi + m^2\phi^2\right) \tag{7.29}$$

with

$$\Theta_{00} = \tfrac{1}{2}(\dot{\phi}^2 + (\vec{\nabla}\phi)^2 + m^2\phi^2) \quad . \tag{7.30}$$

This energy-momentum tensor is conserved for ϕ a solution to the equations of motion,

$$\left(\Box_\eta - m^2\right)\phi = 0 \quad \Rightarrow \quad \partial^a\Theta_{ab} = 0 \quad . \tag{7.31}$$

This energy-momentum tensor is also manifestly symmetric (off-shell, i.e. without using the equations of motion),

$$\Theta_{ba} = \Theta_{ab} \quad . \tag{7.32}$$

In particular, this implies that the angular momentum current associated to an infinitesimal Lorentz transformation (1.28) with parameters $\omega_{bc} = -\omega_{cb}$, namely

$$L^a = \tfrac{1}{2}\omega_{bc}L^{abc} \tag{7.33}$$

with

$$L^{abc} = x^b\Theta^{ac} - x^c\Theta^{ab} \quad , \tag{7.34}$$

is on-shell conserved,

$$\partial_a L^{abc} = \Theta^{bc} - \Theta^{cb} = 0 \quad . \tag{7.35}$$

Since Θ_{ab} is symmetric (and gauge invariance is not an issue), in this example there is no need to "improve" the Noether energy-momentum tensor, and we thus denote it by T_{ab},

$$T_{ab} = \Theta_{ab} = \partial_a \phi \partial_b \phi + \eta_{ab} L \tag{7.36}$$

As we will see below, it is also straightforward to promote this tensor by minimal coupling to a (covariantly conserved) energy-momentum tensor of a scalar field in a gravitational field,

Now let us take a look at Maxwell theory in Minkowski space. In this case the canonical Noether energy-momentum tensor is

$$\Theta_{ab} = -\frac{\partial L}{\partial(\partial^a A_c)} \partial_b A_c + \eta_{ab} L = F_a{}^c \partial_b A_c - \tfrac{1}{4} \eta_{ab} F_{cd} F^{cd} \ . \tag{7.37}$$

It is of course on-shell conserved by construction,

$$\partial^a \Theta_{ab} = 0 \quad \text{on-shell} \tag{7.38}$$

(note that both sets of Maxwell equations are required to derive this), but it is neither symmetric nor gauge-invariant. In particular, therefore, the angular momentum current (7.34) is not conserved (even though Maxwell theory is Lorentz invariant), and the expression for the energy-density is not gauge-invariant and does not agree with the standard expression

$$\Theta_{00} \neq \tfrac{1}{2}(\vec{E}^2 + \vec{B}^2) \ . \tag{7.39}$$

This can be rectified by manipulating Θ_{ab} as

$$\begin{aligned} \Theta_{ab} &= F_a{}^c(\partial_b A_c - \partial_c A_b) - \tfrac{1}{4}\eta_{ab} F_{cd} F^{cd} + F_{ac} \partial^c A_b \\ &= F_a{}^c F_{bc} - \tfrac{1}{4}\eta_{ab} F_{cd} F^{cd} + F_{ac}\partial^c A_b \ , \end{aligned} \tag{7.40}$$

and noting that the last term can be written as a sum of two terms,

$$F_{ac}\partial^c A_b = \partial^c(F_{ac} A_b) - (\partial^c F_{ac}) A_b \tag{7.41}$$

the first of which is identically conserved because of $F_{ac} = -F_{ca}$,

$$\partial^a \partial^c (F_{ac} A_b) = 0 \ , \tag{7.42}$$

and the second of which vanishes on-shell,

$$(\partial^c F_{ac}) A_b = 0 \quad \text{on-shell.} \tag{7.43}$$

Therefore one can redefine the energy-momentum tensor in a first step to

$$\hat{\Theta}_{ab} = \Theta_{ab} - \partial^c (F_{ac} A_b) \tag{7.44}$$

and notes that this energy-momentum tensor is still conserved on-shell,

$$\partial^a \hat{\Theta}_{ab} = 0 \quad \text{on-shell} \ , \tag{7.45}$$

as well as on-shell gauge invariant,

$$\begin{aligned}\hat{\Theta}_{ab} &= F_{ac}F_b{}^c - \tfrac{1}{4}\eta_{ab}F_{cd}F^{cd} - (\partial^c F_{ac})A_b \\ &= F_{ac}F_b{}^c - \tfrac{1}{4}\eta_{ab}F_{cd}F^{cd} \quad \text{on-shell} \ .\end{aligned} \tag{7.46}$$

Therefore one can define the "improved" energy-momentum tensor

$$T_{ab} = F_{ac}F_b{}^c - \tfrac{1}{4}\eta_{ab}F_{cd}F^{cd} \ , \tag{7.47}$$

in such a way that

- T_{ab} is still on-shell conserved,

$$\partial^a T_{ab} = 0 \quad \text{on-shell} \tag{7.48}$$

  (again both sets of Maxwell equations are required to establish this; with an external source,

$$\partial_{[a}F_{bc]} = 0 \quad , \quad \partial_a F^{ab} = -J^b \tag{7.49}$$

  one has the non-conservation law

$$\partial^a T_{ab} = J^a F_{ab} \tag{7.50}$$

  instead, which becomes a conservation law when one adds to T_{ab} the energy-momentum tensor of the source fields + interaction terms);

- T_{ab} is off-shell symmetric,

$$T_{ab} = T_{ba} \tag{7.51}$$

- T_{ab} is gauge-invariant and correctly gives the gauge-invariant and positive-definite energy-density as

$$T_{00} = \tfrac{1}{2}(\vec{E}^2 + \vec{B}^2) \ . \tag{7.52}$$

Moreover, the components of T_{0k} are the components of the Poynting vector and the spatial components T_{ik} are the components of the Maxwell stress tensor. Thus T_{ab} is *the* correct energy-momentum tensor of Maxwell theory.

This procedure to obtain T_{ab} from Θ_{ab} can be understood in a more general and systematic way, via the so-called Belinfante improvement (or symmetrisation) procedure. A brief synopsis of this construction will be provided in section 7.4 below.

One of the many useful properties of a symmetric, conserved energy-momentum tensor, and one that is occasionally used in general relativity, e.g. in the discussion of the energy and energy flux of gravitational waves, is the *Laue Theorem* (or *tensor virial theorem*). It states that for such an energy-momentum tensor and a localised source (so that one can integrate by parts with impunity) one has the relation

$$\left.\begin{array}{c}\partial^a T_{ab} = 0 \quad , \quad T_{ab} = T_{ba} \\ \text{localised source}\end{array}\right\} \quad \Rightarrow \quad \int d^3x \, T^{ik} = \tfrac{1}{2}(\partial_0)^2 \int d^3x \, T_{00}x^i x^k \tag{7.53}$$

between the integrated spatial components T_{ik} and the "quadrupole moments"

$$Q^{ik}(t) = \int d^3x\; T_{00} x^i x^k \tag{7.54}$$

of the energy density T_{00},

$$\int d^3x\; T^{ik} = \tfrac{1}{2}\ddot{Q}^{ik} \quad . \tag{7.55}$$

The proof of this identity is a straightforward repeated application of the conservation law and integration by parts. Indeed, using the symmetry, the time and space components of the conservation law

$$\partial^a T_{a0} = 0 \quad \Leftrightarrow \quad \partial_0 T_{00} = \partial^i T_{i0} \quad , \quad \partial^a T_{ak} = 0 \quad \Leftrightarrow \quad \partial_0 T_{0i} = \partial^k T_{ki} \tag{7.56}$$

and discarding boundary terms, one calculates

$$\begin{aligned}
\tfrac{1}{2}(\partial_0)^2 \int d^3x\; T_{00} x^i x^k &= +\tfrac{1}{2}\partial_0 \int d^3x\; \partial_0 T_{00} x^i x^k \\
&= +\tfrac{1}{2}\partial_0 \int d^3x\; (\partial_j T^j_{\ 0}) x^i x^k \\
&= -\tfrac{1}{2}\partial_0 \int d^3x\; (T^i_{\ 0} x^k + T^k_{\ 0} x^i) \\
&= -\tfrac{1}{2} \int d^3x\; (\partial_0 T^i_{\ 0} x^k + \partial_0 T^k_{\ 0} x^i) \\
&= -\tfrac{1}{2} \int d^3x\; ((\partial_j T^{ij}) x^k + (\partial_j T^{kj}) x^i) \\
&= +\int d^3x\; T^{ik} \quad .
\end{aligned} \tag{7.57}$$

7.4 Synopsis of the Belinfante Improvement Procedure (Review)

The procedure to obtain a symmetric and conserved T_{ab} from the canonical Noether energy-momentum tensor Θ_{ab} of a Poincaré-invariant field theory, illustrated above in the case of Maxwell theory, can be understood in a more general and systematic, but also somewhat round-about way by appealing to the Lorentz-invariance of the action and taking into account the non-trivial transformation behaviour of the fields with spin $\neq 0$ under Lorentz transformations. This recipe is known as the *Belinfante improvement procedure.*[18] Here is, just for reference purposes, a brief description of the general features of this construction:

[18] This is explained in many places, with varying degree of comprehensibility or comprehension. For a detailed explanation, geared also towards applications to general relativity, see section 2 of T. Ortin, *Gravity and Strings*; for a succinct description, and an extension of the usual procedure to Lagrangians depending also on second derivatives of the fields, see section II of D. Bak, D. Cangemi, R. Jackiw, *Energy-Momentum Conservation in General Relativity*, `arXiv:hep-th/9310025`.

- In general (with the exception of spin zero scalar fields), $\Theta_{ab} = \eta_{ac}\Theta^c_{\ b}$ is not symmetric,

$$\Theta_{ab} \neq \Theta_{ba} \quad . \tag{7.58}$$

- As a consequence, the would-be angular momentum current (7.34) is now not on-shell conserved,

$$\partial_a L^{abc} = \Theta^{bc} - \Theta^{cb} \neq 0 \quad . \tag{7.59}$$

- By Lorentz invariance of the action and Noether's theorem, the total (orbital + spin) angular momentum should be conserved, and the above (purely orbital) angular momentum current fails to be conserved because it does not take into account the spin, i.e. the fact that the ϕ are possibly non-trivial Lorentz tensors (an irrelevant fact as far as the translational symmetries and hence the Noether energy-momentum tensor are concerned).

- This can be rectified by constructing the conserved total angular momentum current J^{abc} directly from Noether's theorem applied to Lorentz transformations $\delta\phi = \delta_L\phi$ of the fields and coordinates. This gives rise to an additional (spin) contribution to the current, schematically of the form

$$J^a = J^a_{\text{orbit}} + \frac{\partial L}{\partial(\partial_a\phi)}\delta_L\phi \quad , \quad J^{a[bc]}_{\text{orbit}} = L^{abc} \quad . \tag{7.60}$$

  From the conservation of this current one can then via some gymnastics deduce and extract a candidate energy-momentum tensor $\hat{\Theta}_{ab}$ which is such that the total angular momentum current J^a takes the form

$$J^{abc} = x^b\hat{\Theta}^{ac} - x^c\hat{\Theta}^{ab} \quad . \tag{7.61}$$

  Note that the spin-contribution to the total angular momentum has in this way been transformed into an orbital contribution with respect to the new energy-momentum tensor $\hat{\Theta}_{ab}$.

- This tensor $\hat{\Theta}_{ab}$ turns out to differ from the canonical energy-momentum tensor Θ_{ab} by an identically conserved term,

$$\hat{\Theta}_{ab} = \Theta_{ab} + \partial_c\Psi^{cab} \quad , \tag{7.62}$$

  with

$$\Psi^{cab} = -\Psi^{acb} \quad \Rightarrow \quad \partial_a\partial_c\Psi^{cab} \equiv 0 \tag{7.63}$$

  so that

$$\partial_a\Theta^a_{\ b} = 0 \quad \text{on-shell} \quad \Rightarrow \quad \partial_a\hat{\Theta}^a_{\ b} = 0 \quad \text{on-shell} \quad . \tag{7.64}$$

- Addition of such a term to the energy-momentum tensor is always possible as it does not violate the conservation law. While this changes the definition of the local energy and momentum densities, with suitable fall-off conditions on the Ψ^{abc} this has no effect on the total energy-momentum P_b (7.28),

$$\begin{aligned} P^b \to P^b + \int d^3x\,\partial_c\Psi^{c0b} &= P^b + \int d^3x\,\partial_k\Psi^{k0b} \quad . \\ &= P^b + \oint dS_k\Psi^{k0b} \quad . \end{aligned} \tag{7.65}$$

- Angular momentum conservation together with (7.64) now implies

$$\partial_a J^{abc} = 0 \quad \text{on-shell} \quad \Rightarrow \quad \hat{\Theta}_{ab} = \hat{\Theta}_{ba} \quad \text{on-shell} \quad . \tag{7.66}$$

- Thus on-shell $\hat{\Theta}_{ab}$ agrees with a tensor T_{ab}, which can be chosen to be symmetric (off-shell) and on-shell conserved,

$$\begin{aligned} \hat{\Theta}_{ab} \to T_{ab}: \quad T_{ab} &= T_{ba} \quad \text{off-shell} \\ \partial_a T^a_{\ b} &= 0 \quad \text{on-shell} \quad . \end{aligned} \tag{7.67}$$

- This tensor T_{ab} (or occasionally just $\hat{\Theta}_{ab}$) is known as the *Belinfante improvement* of the energy-momentum tensor, and $\partial_c\Psi^{cab}$ as the (identically conserved) *improvement term*,

$$T_{ab} = \{\text{Improvement of } \Theta\}_{ab} \quad . \tag{7.68}$$

T_{ab} is then generally considered to be the "correct" choice of energy-momentum tensor for the Lagrangian field thory at hand, but it should be clear from the above discussion that this somewhat round-about procedure for finding and obtaining it leaves something to be desired (to put it mildly), already in the framework of Special Relativity.

7.5 Energy-Momentum Tensor from Minimal Coupling?

Given the success of the minimal coupling prescription, it is natural to try to define the matter energy-momentum tensor in a gravitational field in the same way. While this is certainly possible to a certain extent (as the examples will show), this procedure also leaves something to be desired (as the examples will also show).

Let us start by considering the "phenomenological" perfect fluid energy-momentum tensor (7.6),

$$T_{ab} = (\rho + p)u_a u_b + p\eta_{ab} \quad . \tag{7.69}$$

Following the minimal coupling rules, we promote this to the energy-momentum tensor

$$T_{\alpha\beta} = (\rho + p)u_\alpha u_\beta + pg_{\alpha\beta} \quad , \tag{7.70}$$

where u^α denotes the proper-time normalised velocity field of the fluid, $g_{\alpha\beta}u^\alpha u^\beta = -1$.

The covariantisation of the conservation law (7.7) evidently reads

$$\partial^a T_{ab} = 0 \quad \to \quad \nabla^\alpha T_{\alpha\beta} = 0 \ . \tag{7.71}$$

This generalises the continuity equation and the relativistic Euler equations to a fluid moving in a gravitational field and reduces to the special relativistic laws at the origin of a freely falling coordinate system, as it should.

There are neither conceptual nor technical complications in this example, and we will adopt this perfect fluid energy-momentum tensor, supplemented by an appropriate equation of state, to model the interior of a star (section 24.7) and the matter content of the universe (in our discussion of cosmology). In both of these examples, such a phenomenological description is quite appropriate and sufficient (although for more detailed investigations one may need to go beyond the perfect fluid approximation). For a detailed analysis of the conservation equations in the context of cosmology, see sections 35.4 and 35.6.

Let us now turn to energy-momentum tensors for Lagrangian field theories, starting with the example of the Klein-Gordon scalar field. As we saw above, in Minkowski space its (Noether = improved) energy-momentum tensor is given by

$$T_{ab} = \partial_a\phi\partial_b\phi + \eta_{ab}L = \partial_a\phi\partial_b\phi - \tfrac{1}{2}\eta_{ab}\left(\eta^{cd}\partial_c\phi\partial_d\phi + m^2\phi^2\right) \ . \tag{7.72}$$

The corresponding minimally coupled energy-momentum tensor in a gravitational field is then evidently

$$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi + g_{\alpha\beta}L = \partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}g_{\alpha\beta}\left(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + m^2\phi^2\right) \ , \tag{7.73}$$

and it is easy to check that it is covariantly conserved for ϕ a solution to the equations of motion in a gravitational background,

$$\left(\Box_g - m^2\right)\phi = 0 \quad \Rightarrow \quad \nabla^\alpha T_{\alpha\beta} = 0 \ . \tag{7.74}$$

For the action (6.14) with a potential $V(\phi)$, the energy-momentum tensor of course also has the form (7.73) with $m^2\phi^2/2$ unsurprisingly replaced by $V(\phi)$,

$$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}g_{\alpha\beta}\, g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - g_{\alpha\beta}V(\phi) \ , \tag{7.75}$$

with

$$\Box_g\phi = V'(\phi) \quad \Rightarrow \quad \nabla^\alpha T_{\alpha\beta} = 0 \ . \tag{7.76}$$

So far so good. However, the significance of this energy-momentum tensor outside the realm of special relativity is not clear. In special relativity, it encodes the conserved quantities associated to translation invariance, but in a general gravitational field there is no translation invariance (or other symmetry). In particular, in a general gravitational field

- one cannot even derive the energy-momentum tensor (7.73) from Noether's theorem applied to translations

- and, related to this is the fact that one does not obtain an ordinary conservation law but the covariant conservation law $\nabla_\alpha T^{\alpha\beta} = 0$.

Regarding the first point, it is fair to wonder if $T_{\alpha\beta}$ possesses an intrinsic gravitational significance beyond being merely the non-conserved minimally coupled counterpart of something that happens to have a significance in the absence of gravity. We will see below that, yes indeed, it is precisely the *source of gravity* arising from scalar fields.

Regarding the second point, we will see in sections 7.9 and 10.1 below that to any continous symmetry of a gravitational field (metric) and the covariantly conserved energy-momentum tensor one *can* associate a covariantly conserved current and thus also (as discussed in section 6.8) a conserved charge.

Now let us turn to Maxwell theory. Here the situation is a priori a bit murkier, because in principle we have both the canonical Noether energy-momentum tensor Θ_{ab} (7.37),

$$\Theta_{ab} = F_a^{\ c} \partial_b A_c - \tfrac{1}{4}\eta_{ab} F_{cd} F^{cd} \quad , \tag{7.77}$$

and its Belinfante-improved symmetric gauge-invariant sibling T_{ab} (7.47),

$$T_{ab} = F_{ac} F_b^{\ c} - \tfrac{1}{4}\eta_{ab} F_{cd} F^{cd} \quad , \tag{7.78}$$

at our disposal. Let us start with the latter, not only because it is the nicer object but also because it turns out to give the "correct" result. Applying the rules of minimal coupling, one finds the tensor

$$T_{\alpha\beta} = F_{\alpha\gamma} F_\beta^{\ \gamma} - \tfrac{1}{4} g_{\alpha\beta} F_{\gamma\delta} F^{\gamma\delta} \quad , \tag{7.79}$$

where indices of the (metric independent) field strength tensor $F_{\alpha\beta}$ are of course raised with the aid of the inverse metric $g^{\alpha\beta}$. This object turns out to have all the right properties to qualify as a candidate energy-momentum tensor of Maxwell theory in a gravitational field. In particular, it is off-shell symmetric and moreover on-shell covariantly conserved,

$$\nabla_\alpha T^{\alpha\beta} = 0 \quad \text{on-shell} \quad , \tag{7.80}$$

where "on-shell" of course refers to the equations of motion (6.39) in a gravitational field. Again both sets of vacuum Maxwell equations are required to verify this. In the presence of an external current, this is modified to

$$\nabla_\alpha T^{\alpha\beta} = J_\alpha F^{\alpha\beta} \quad \text{on-shell} \quad . \tag{7.81}$$

While one may have anticipated these last two equations on the basis of the minimal coupling recipe, it is important (and a useful exercise) to verify by direct calculation that they indeed hold. The point of this verification is to make sure that no commutators of covariant derivatives, i.e. "curvature terms", arise in and mess up this equation, as they will in the calculation below involving the Noether energy-momentum tensor.

So let us take a brief look at the covariantised or minimally coupled Noether energy-momentum tensor, namely

$$\Theta_{\alpha\beta} = F_{\alpha}^{\ \gamma}\nabla_{\beta}A_{\gamma} - \tfrac{1}{4}g_{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta} \quad . \tag{7.82}$$

While the canonical energy-momentum tensor in Minkowski space had some undesirable properties, its one redeeming feature was that it was on-shell conserved. In contrast to this, $\Theta_{\alpha\beta}$ is neither on-shell conserved nor on-shell covariantly conserved. In order to establish $\partial^a\Theta_{ab} = 0$ in Minkowski space, one uses the fact that partial derivatives commute. Thus, analogously, in calculating $\nabla^{\alpha}\Theta_{\alpha\beta}$ one encounters the *commutator of covariant derivatives*. Explicitly on-shell one finds

$$\nabla_{\alpha}\Theta^{\alpha}_{\ \beta} = \tfrac{1}{2}F^{\alpha\gamma}[\nabla_{\alpha},\nabla_{\gamma}]A_{\beta} \quad , \tag{7.83}$$

However, as we will discuss at length in section 8, the characteristic and defining feature of a non-trivial curved space-time is that these covariant derivatives do not commute when acting on tensors other than scalars (their commutator defining the curvature tensor of the space-time).

Likewise the covariant version of the improvement term in

$$\hat{\Theta}_{ab} = \Theta_{ab} - \partial^c(F_{ac}A_b) \quad , \tag{7.84}$$

namely $\nabla^{\gamma}(F_{\alpha\gamma}A_{\beta})$ is not identically conserved anymore, rather one has

$$\nabla^{\alpha}\nabla^{\gamma}(F_{\alpha\gamma}A_{\beta}) = \tfrac{1}{2}F^{\alpha\gamma}[\nabla_{\alpha},\nabla_{\gamma}]A_{\beta} \quad , \tag{7.85}$$

so that it would not qualify as an "improvement term" in the standard sense. Nevertheless, subtracting this term from the (non-conserved) Noether energy-momentum tensor, one finds that this indeed cancels the commutator term arising form (7.83), thus giving rise to an on-shell covariantly conserved $\hat{\Theta}_{\alpha\beta}$ or $T_{\alpha\beta}$. From the present perspective, however, this must be considered to be somewhat of a miracle or fluke. For some more comments on this, see section 22.2.

Thus we adopt (7.79),

$$T_{\alpha\beta} = F_{\alpha\gamma}F_{\beta}^{\ \gamma} - \tfrac{1}{4}g_{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta} \quad , \tag{7.86}$$

as our (preliminary) definition of the energy-momentum tensor of Maxwell theory in a gravitational field, but we now face the same issue as in the case of scalar fields, namely the question what, if any, is the intrinsic gravitational significance of this energy-momentum tensor.

As we have seen, there are some irritating conceptual and technical issues associated with the "Noether + minimal coupling" procedure in general. These irritants turn out to be a good thing, though, because they motivate us to rethink this issue from scratch, and this will now lead us to a much more compelling and both conceptually and technically perfectly satisfactory general definition of the energy-momentum tensor of any Lagrangian field theory in a gravitational field.

Thus let us think about this issue from a Lagrangian, action-based, perspective. So far we have discussed what is the appropriate form of the action for matter fields in a gravitational field, namely a generally covariant action

$$S_{matter} = S_M[\phi; g_{\alpha\beta}] = \int \sqrt{g} d^4x \; L_M(\phi, \partial_\alpha \phi, \ldots, g_{\alpha\beta}, \ldots) \tag{7.87}$$

for the matter fields ϕ in a gravitational background $g_{\alpha\beta}$, obtained e.g. by the minimal coupling description and thus describing the dynamics of the fields in a gravitational background and encoding the coupling of the matter fields to gravity. Ultimately, this action should then be one part of the total gravitational + matter action describing the dynamics of the matter fields *and* of the gravitational field,

$$S_{total} = S_{gravity} + S_{matter} \quad . \tag{7.88}$$

Since the gravitational field is described by the (now dynamical) variables $g_{\alpha\beta}(x)$, we can write this marginally more explicitly as

$$S[g_{\alpha\beta}, \phi] = S_g[g_{\alpha\beta}] + S_M[\phi; g_{\alpha\beta}] \quad . \tag{7.89}$$

The precise form of the gravitational action S_g will not be relevant here - this is something that we will discuss at length in section 20. All we need to keep in mind is that this action is to provide us with the gravitational part of the gravitational field equations, i.e. with the appropriate tensorial generalisation of the left-hand side $\Delta\phi$ of the Newtonian field equation $\Delta\phi = 4\pi G_N \rho$.

Variation of this total action with respect to the matter fields ϕ is equivalent to the variation of the matter action S_M alone with respect to the matter fields,

$$\frac{\delta S[g_{\alpha\beta}, \phi]}{\delta \phi} = 0 \quad \Leftrightarrow \quad \frac{\delta S_M[\phi; g_{\alpha\beta}]}{\delta \phi} = 0 \quad , \tag{7.90}$$

and will thus simply give rise to the equations of motion of the matter fields in a gravitational field, as required.

Now let us consider the variation of the total action with respect to the gravitational dynamical variables $g_{\alpha\beta}$,

$$\frac{\delta S[g_{\alpha\beta}, \phi]}{\delta g_{\alpha\beta}} = \frac{\delta S_g[g_{\alpha\beta}]}{\delta g_{\alpha\beta}} + \frac{\delta S_M[\phi; g_{\alpha\beta}]}{\delta g_{\alpha\beta}} \stackrel{!}{=} 0 \tag{7.91}$$

Variation of the gravitational action with respect to the gravitational field $g_{\alpha\beta}$ will give us the gravitational part of the field equations. Thus variation of the matter action with respect to the gravitational field will give us the source term for the gravitational field equations provided by the matter fields,

$$\frac{\delta S_M[\phi; g_{\alpha\beta}]}{\delta g_{\alpha\beta}} = \text{Source of Gravity} \quad . \tag{7.92}$$

On the other hand, as recalled in the introduction to this section (section 7.1), we expect the energy-momentum tensor to act as the source of gravity. Therefore we should simply define the energy-momentum tensor by this relation,

$$\begin{aligned} T^{\alpha\beta} &:= \text{Source of Gravity} \\ \Rightarrow \quad T^{\alpha\beta} &\sim \frac{\delta S_M[\phi; g_{\alpha\beta}]}{\delta g_{\alpha\beta}} \quad . \end{aligned} \tag{7.93}$$

We will fix the proportionality factor momentarily.

Note that this is precisely analogous to the way a source term for the Maxwell equations, a current J^α, arises from the variation of the coupled matter-Maxwell action with respect to the gauge field A_α (6.55),

$$J^\alpha \sim \frac{\delta S_M[\phi, A_\alpha]}{\delta A_\alpha} \quad . \tag{7.94}$$

In order to test this suggestion, let us take a look at our two standard examples, a scalar field and Maxwell theory. For a scalar field, the minimally coupled action is (6.14)

$$S[\phi, g_{\alpha\beta}] = \int \sqrt{g} d^4x \left(-\tfrac{1}{2} g^{\alpha\beta} \partial_\alpha\phi \partial_\beta\phi - V(\phi)\right) \quad . \tag{7.95}$$

Since the action depends explicitly on the inverse metric, it is more convenient to determine the variation of the action under variations

$$g^{\alpha\beta} \to g^{\alpha\beta} + \delta g^{\alpha\beta} \tag{7.96}$$

of the inverse metric. Under such a variation, the volume factor $\sqrt{g}$ varies as (5.88)

$$\delta\sqrt{g} = -\tfrac{1}{2}\sqrt{g} g_{\alpha\beta} \delta g^{\alpha\beta} \quad . \tag{7.97}$$

Thus the metric-variation of the scalar field action is

$$\delta S_M[\phi, g_{\alpha\beta}] = -\tfrac{1}{2} \int \sqrt{g} d^4x \left(\partial_\alpha\phi\partial_\beta\phi + g_{\alpha\beta}(-\tfrac{1}{2} g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi))\right) \delta g^{\alpha\beta} \quad , \tag{7.98}$$

Comparison with the minimally coupled energy-momentum tensor (7.75) of a scalar field,

$$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2} g_{\alpha\beta}\, g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - g_{\alpha\beta} V(\phi) \quad , \tag{7.99}$$

shows that this is precisely what we have obtained from the metric-variation of the matter action,

$$\delta S_M[\phi, g_{\alpha\beta}] = -\tfrac{1}{2} \int \sqrt{g} d^4x\, T_{\alpha\beta} \delta g^{\alpha\beta} \quad . \tag{7.100}$$

Now let us look at Maxwell theory, our litmus test. In this case, the action is (6.51)

$$S[A_\alpha, g_{\alpha\beta}] = -\tfrac{1}{4}\int \sqrt{g}d^4x\; g^{\mu\lambda}g^{\nu\rho}F_{\mu\nu}F_{\lambda\rho} \quad . \tag{7.101}$$

The variation of $\sqrt{g}$ is as before, and as regards the variation of the inverse metric, there is now an additional relative factor of two compared with the calculation for the scalar fields because the action depends quadratically on the inverse metric. Thus one has

$$\delta S[A_\alpha, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g}d^4x\; \left(g^{\nu\rho}F_{\alpha\nu}F_{\beta\rho} - \tfrac{1}{4}g_{\alpha\beta}F_{\mu\nu}F^{\mu\nu}\right)\delta g^{\alpha\beta} \quad . \tag{7.102}$$

Comparing with (7.79),

$$T_{\alpha\beta} = F_{\alpha\gamma}F_{\beta}^{\;\gamma} - \tfrac{1}{4}g_{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta} \quad , \tag{7.103}$$

we see that we once again have

$$\delta S_M[\phi, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g}d^4x\; T_{\alpha\beta}\delta g^{\alpha\beta} \quad . \tag{7.104}$$

Thus the metric variation of the matter action has given us on the nose the symmetric, gauge invariant, on-shell conserved energy-momentum tensor of Maxwell theory, without any need to appeal to any improvement procedures!

Thus, when it comes to defining the energy-momentum tensor for Maxwell theory, the above approach based on the variation of the matter action with respect to the metric wins hands down over the painful canonical definition based on Noether's theorem for translations and the Belinfante improvement procedure combined with minimal coupling.

Encouraged by this, we now define the energy-momentum tensor $T_{\alpha\beta}$ in general by

$$\delta_{\text{metric}}S_M[\phi, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g}d^4x\; T_{\alpha\beta}\delta g^{\alpha\beta} \quad , \tag{7.105}$$

or, equivalently,

$$T_{\alpha\beta} := -\frac{2}{\sqrt{g}}\frac{\delta}{\delta g^{\alpha\beta}}S_M[\phi, g_{\alpha\beta}] \quad . \tag{7.106}$$

Even though, as we have seen, there are other definitions of the energy-momentum tensor, this is the modern, and by far the most useful, definition of the energy-momentum tensor, namely as the response of the matter action to a variation of the metric (equivalently, as the source of gravity).

Moreover, crucially for the present context, whatever the virtues of other definitions may be, from the variational principle for general relativity it is this energy-momentum tensor that plays the role of the source term for the Einstein equations.

<u>Remarks:</u>

1. The energy-momentum tensor as defined by (7.105) or (7.106) is frequently called the *metric* energy-momentum (or stress-energy) tensor, or also the *Hilbert* or

Rosenfeld energy-momentum tensor. It is sometimes also referred to as the gravitational energy-momentum tensor, but that is confusing as it does not describe the energy-momentum of the gravitational field itself, a more mysterious and elusive quantity we will briefly look at and for in section 22.6.

I prefer the attribute *covariant*, to distinguish it from what is usually called the *canonical* Noether energy-momentum tensor. Thus, even though this terminology is not standard, I will henceforth refer to $T_{\alpha\beta}$ as defined by (7.105) or (7.106), as the *Covariant Energy-Momentum Tensor*.

2. One of the many advantages of this definition is that it automatically and in general gives a symmetric and gauge invariant tensor (no improvement terms or similar gymnastics required). This is obvious from the definition.

3. This energy-momentum tensor turns out to also automatically be covariantly conserved (on-shell, i.e. for matter fields satisfying their Euler-Lagrange equations of motion). We will establish this latter fact in section 20.6 below where we will see that this is simply a consequence of the general covariance of the matter action S_M,
$$\text{general covariance of } S_M \quad \Rightarrow \quad \nabla^\alpha T_{\alpha\beta} = 0 \quad \text{on-shell} \quad . \tag{7.107}$$

4. When the minimally coupled matter Lagrangian depends only on the metric and not on the first derivatives of the metric (i.e. not on the Christoffel symbols),
$$L_M(x) = L_M(\phi(x), \partial_\mu\phi(x), g_{\mu\nu}(x)) \quad , \tag{7.108}$$
as in the case of scalar or Maxwell gauge fields, then more explicitly the covariant energy-momentum tensor can be written as (and calculated from)
$$T_{\mu\nu}(x) = -\frac{2}{\sqrt{g}}\frac{\partial(\sqrt{g}L_M(x))}{\partial g^{\mu\nu}(x)} = -2\frac{\partial L_M(x)}{\partial g^{\mu\nu}(x)} + g_{\mu\nu}(x)L_M(x) \tag{7.109}$$
or
$$T^{\mu\nu}(x) = 2\frac{\partial L_M(x)}{\partial g_{\mu\nu}(x)} + g^{\mu\nu}(x)L_M(x) \quad . \tag{7.110}$$
Here the sign change is due to the fact that $\delta g^{\mu\nu}$ denotes the variation of the inverse metric, not the contravariant components of $\delta g_{\mu\nu}$. Thus it is not the same as $g^{\mu\lambda}g^{\nu\rho}\delta g_{\lambda\rho}$, but rather minus this expression,
$$\delta g^{\mu\nu} = -g^{\mu\lambda}g^{\nu\rho}\delta g_{\lambda\rho} \quad , \tag{7.111}$$
as can be seen by varying $g^{\mu\nu}g_{\nu\lambda} = \delta^\mu_\lambda$,
$$0 = \delta(g^{\mu\nu}g_{\nu\lambda}) = (\delta g^{\mu\nu})g_{\nu\lambda} + g^{\mu\rho}\delta g_{\rho\lambda} \quad \Leftrightarrow \quad \delta g^{\mu\nu} = -g^{\mu\lambda}g^{\nu\rho}\delta g_{\lambda\rho} \quad . \tag{7.112}$$

5. The definition (7.105) or the explicit expression (7.109) also provides an efficient strategy to determine the energy-momentum tensor even if one is just interested in Poincaré-invariant field theories in Minkowski space:

   In order to determine a symmetric, gauge invariant, and on-shell conserved energy-momentum tensor T_{ab} for such a theory, one

   - temporarily minimally couples the theory to a metric $g_{\alpha\beta}(x)$,
   - uses (7.105) or (7.106) or (7.109) to determine $T_{\alpha\beta}$,
   - and then replaces $g_{\alpha\beta} \to \eta_{ab}$ etc. again at the end.

   In equations, one defines T_{ab} by

   $$T_{ab} := (T_{\alpha\beta})|_{x^\alpha \to \xi^a, g_{\alpha\beta} \to \eta_{ab}} \quad . \tag{7.113}$$

   It can be shown that for fields of any spin this energy-momentum tensor agrees on-shell with what one could have also obtained by invoking the Belinfante improvement procedure of the Noether energy-momentum tensor,

   $$\hat{\Theta}_{ab} = T_{ab} \quad \text{on-shell} \tag{7.114}$$

   (see the discussion and references in section 7.4).

6. When the minimally coupled matter action depends also on the first derivatives of the metric, through the covariant derivative $\nabla_\mu \psi$ of some (non-scalar) field ψ, say, by the usual rules of variational calculus there will be additional contributions to the energy-momentum tensor, arising from an integration by parts of

   $$\int \sqrt{g} d^4x \, \left(-2 \frac{\partial L_M(x)}{\partial \nabla_\mu \psi(x)} (\delta \nabla_\mu \psi)(x) \right)$$

   where $\delta(\nabla_\mu \psi) = (\delta \nabla_\mu)\psi$ denotes the variation of the covariant derivative induced by the metric-variation (e.g. via the corresponding variation (20.14) of the Christoffel symbols). The precise form of the resulting contribution to the energy-momentum tensor depends on the tensorial type of ψ, is rarely needed, and it is unedifying to attempt to write down a general formula for this.

7.7 On the Energy-Momentum Tensor for Weyl-invariant Actions

Another general feature of the energy-momentum tensor that is readily understood by adopting the definition

$$T_{\alpha\beta} = -\frac{2}{\sqrt{g}} \frac{\delta S_{\text{matter}}}{\delta g^{\alpha\beta}} \quad . \tag{7.115}$$

is the relation between Weyl invariance and the trace of the energy-momentum tensor.

We consider the situation where the minimally coupled matter action happens to be invariant under *Weyl rescalings*, i.e. under rescalings of the metric

$$g_{\alpha\beta}(x) \to \mathrm{e}^{2\omega(x)} g_{\alpha\beta}(x) \tag{7.116}$$

by a positive definite function, or infinitesimally

$$\delta_\omega g_{\alpha\beta}(x) = 2\omega(x) g_{\alpha\beta}(x) \ . \tag{7.117}$$

In particular, thus, we consider the (admittedly very special) situation where one has such a symmetry without any accompanying transformation of the matter fields. The discussion can be extended to the case where also a transformation of the matter fields is required, but for present purposes this special case is good enough (see the end of this section for a comment on the general case).

Examples of such actions are e.g. the action of a massless scalar field (6.11) in $D = 2$ (space-time) dimensions

$$S[\phi, g_{\alpha\beta}] = -\tfrac{1}{2} \int d^2x \ \sqrt{g} g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi \tag{7.118}$$

and that of Maxwell theory (6.51) in $D = 4$ dimensions,

$$S[A_\alpha, g_{\alpha\beta}] = -\tfrac{1}{4} \int d^4x \ \sqrt{g} g^{\alpha\beta} g^{\gamma\delta} F_{\alpha\gamma} F_{\beta\delta} \ . \tag{7.119}$$

Indeed, in that case the metric dependence of the action is precisely such that the combination of of the determinant $\sqrt{g}$ and the inverse metric that appears is invariant under Weyl rescalings,

$$g_{\alpha\beta} \to \mathrm{e}^{2\omega} g_{\alpha\beta} \quad \Rightarrow \quad \begin{cases} D = 2 & \sqrt{g} g^{\alpha\beta} \to \sqrt{g} g^{\alpha\beta} \\ D = 4 & \sqrt{g} g^{\alpha\beta} g^{\gamma\delta} \to \sqrt{g} g^{\alpha\beta} g^{\gamma\delta} \end{cases} \tag{7.120}$$

This is reflected in the fact that the corresponding energy-momentum tensor is traceless precisely in these dimensions: from (7.73) and (7.79) one finds

$$\begin{aligned} T_{\alpha\beta} = \partial_\alpha \phi \partial_\beta \phi - \tfrac{1}{2} g_{\alpha\beta} \left(g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \right) \quad &\Rightarrow \quad T^\alpha_{\ \alpha} = -\tfrac{1}{2}(D-2) g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \\ T_{\alpha\beta} = F_{\alpha\lambda} F_\beta^{\ \lambda} - \tfrac{1}{4} g_{\alpha\beta} F_{\lambda\sigma} F^{\lambda\sigma} \quad &\Rightarrow \quad T^\alpha_{\ \alpha} = -\tfrac{1}{4}(D-4) F_{\mu\nu} F^{\mu\nu} \ . \end{aligned} \tag{7.121}$$

The relation between these two observations / assertions is provided by noting that if the matter action is invariant under Weyl rescalings one has

$$\begin{aligned} 0 = \delta_\omega S_{\text{matter}} &= -\tfrac{1}{2} \int \sqrt{g} d^D x \ T_{\alpha\beta}(x) \delta_\omega g^{\alpha\beta}(x) \\ &= \int \sqrt{g} d^D x \ T_{\alpha\beta}(x) g^{\alpha\beta}(x) \omega(x) = \int \sqrt{g} d^D x \ T^\alpha_{\ \alpha}(x) \omega(x) \ . \end{aligned} \tag{7.122}$$

Since this is to be zero for all functions $\omega(x)$, this proves

$$\text{invariance under Weyl rescalings of the metric} \quad \Rightarrow \quad T^\alpha_{\ \alpha} = 0 \ . \tag{7.123}$$

In the special case that we have considered here (invariance under scalings of the metric alone, without transforming the matter fields), this is true off-shell, i.e. without using the equations of motion for the matter fields. In the more general case of an invariance under joint Weyl rescalings of the metric and accompanying scalings of the matter fields, in the above chain of arguments one would need to also vary the matter action with respect to the matter fields to establish the invariance of the action. The term arising from the variation of the matter fields is evidently proportional to the Euler-Lagrange equations of the matter fields, and therefore in that case one could only conclude that $T^\alpha_\alpha = 0$ on-shell,

$$\left.\begin{array}{c}\text{invariance under joint Weyl rescalings}\\ \text{of the metric and the matter fields}\end{array}\right\} \quad \Rightarrow \quad T^\alpha_\alpha = 0 \;\; \text{on-shell.} \tag{7.124}$$

An example of this is provided by the so-called *conformally coupled scalar field*. This conformal coupling involves a space-time dependent mass term that represents a non-minimal coupling of the scalar field to the scalar curvature (a contraction of the Riemann curvature tensor to be introduced in section 8), and understanding the Weyl invariance of this model requires a formula for the variation of the scalar curvature with respect to the metric which we will derive in section 20.2. Therefore we will need to postpone a discussion of this model to section 22.3.

7.8 Klein-Gordon Scalar Field in (1+1) Minkowski and Rindler Space

As an aside, but as a concrete, and the simplest non-trivial, example, and an illustration of the above remarks regarding Weyl invariance, let us consider a massless scalar field in (1+1)-dimensions, in either the usual Minkowski coordinates, or in the Rindler coordinates discussed in sections 1.3 and 3.4 (we will in particular make use of the results in section 3.4).

In inertial coordinates in (1+1)-dimensional Minkowski space-time, $ds^2 = -dt^2 + dx^2$, the action and equation of motion of a massless scalar field are

$$S_M[\phi] = -\tfrac{1}{2}\int dtdx\; \eta^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi \quad \Rightarrow \quad \Box\phi = (-\partial_t^2 + \partial_x^2)\phi = 0 \;\; . \tag{7.125}$$

Thus a natural basis of solutions to this equation is provided by the plane waves $f_k \sim \exp(-i\omega t + ikx)$, with $k^2 = \omega^2$, i.e. $k = \pm\omega, \omega > 0$, and their complex conjugates. For a given ω there are thus two linearly-independent positive frequency solutions,

$$\begin{aligned} f_\omega(t,x) &= \frac{1}{(4\pi\omega)^{1/2}}\mathrm{e}^{-i\omega(t-x)} \\ g_\omega(t,x) &= \frac{1}{(4\pi\omega)^{1/2}}\mathrm{e}^{-i\omega(t+x)} \end{aligned} \tag{7.126}$$

(the normalisation factors are inserted for QFT-pedantry reasons only and are irrelevant for the following). Thus the basis of solutions splits into *right-movers* or *right-moving* *modes* f_ω and *left-movers* g_ω. It is thus convenient to introduce the corresponding null coordinates $u_M = t - x, v_M = t + x$ as in (3.94), in terms of which the solutions can be written as

$$\begin{aligned} f_\omega = f_\omega(u_M) &= \frac{1}{(4\pi\omega)^{1/2}} \mathrm{e}^{-i\omega u_M} \\ g_\omega = g_\omega(v_M) &= \frac{1}{(4\pi\omega)^{1/2}} \mathrm{e}^{-i\omega v_M} \quad . \end{aligned} \tag{7.127}$$

That the solutions split in this way could have also been deduced from the form of the wave operator in these lightcone (null) coordinates, namely $\Box = -4\partial_{u_M}\partial_{v_M}$, and the ensuing solutions to the equation of motion,

$$\Box\phi = 0 \quad \Rightarrow \quad \phi = f(u_M) + g(v_M) \quad . \tag{7.128}$$

Here f and g can now be arbitrary wave packets constructed from the solutions f_ω and g_ω respectively.

The energy-density $\rho_M = T_{tt}$ of the scalar field with respect to Minkowski time is

$$\rho_M = \tfrac{1}{2}((\partial_t\phi)^2 + (\partial_x\phi)^2) \tag{7.129}$$

and in terms of lightcone coordinates this splits into a sum of left-moving and right-moving contributions,

$$\rho_M = (\partial_{u_M}\phi)^2 + (\partial_{v_M}\phi)^2 \quad , \tag{7.130}$$

with $f(u_M)$ evidently only contributing to the former and $g(v_M)$ to the latter.

Now let us consider the same issue in Rindler coordinates. In terms of the coordinates (η, ξ) (3.89), the metric takes the form (3.87)

$$ds^2 = \mathrm{e}^{2a\xi}(-d\eta^2 + d\xi^2) \quad . \tag{7.131}$$

Note that, as mentioned in section 3.4, the metric in these coordinates is conformally flat. Thus, by the reasoning above, in section 7.7, in particular the discussion around equation (7.120), we know that the action and equation of motion for a scalar field in Rindler coordinates will look just like those in Minkowski coordinates, with the replacement $(t, x) \to (\eta, \xi)$,

$$S_R[\phi] = -\tfrac{1}{2}\int \sqrt{g}d\eta d\xi \; g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi = -\tfrac{1}{2}\int d\eta d\xi \; \eta^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi \quad , \tag{7.132}$$

and

$$\Box_g\phi = 0 \quad \Leftrightarrow \quad (-\partial_\eta^2 + \partial_\xi^2)\phi = 0 \quad . \tag{7.133}$$

Thus by the same reasoning as above, the solutions can be split into left- and right-movers and are conveniently written in terms of the Rindler lightcone coordinates (3.95)

$$(u_R, v_R) = \eta \mp \xi \quad , \tag{7.134}$$

i.e. one has

$$\Box_g \phi = 0 \quad \Leftrightarrow \quad \partial_{u_R}\partial_{v_R}\phi = 0 \quad \Leftrightarrow \quad \phi = f(u_R) + g(v_R) \ . \tag{7.135}$$

The energy-density $\rho_R = T_{\eta\eta}$ of the scalar field with respect to Rindler time is

$$\rho_R = \tfrac{1}{2}((\partial_\eta \phi)^2 + (\partial_\xi \phi)^2) \tag{7.136}$$

and in terms of lightcone coordinates this splits into a sum of left-moving and right-moving contributions,

$$\rho_R = (\partial_{u_R}\phi)^2 + (\partial_{v_R}\phi)^2 \ , \tag{7.137}$$

with $f(u_R)$ evidently only contributing to the former and $g(v_R)$ to the latter.

The interest in these (fairly trivial) considerations lies in the fact that the exponential relation (3.97) between the Minkowski and Rindler null coordinates

$$u_M = -a^{-1}\mathrm{e}^{-au_R} \quad , \quad v_M = +a^{-1}\mathrm{e}^{+av_R} \ , \tag{7.138}$$

reflecting the exponential redshift of a Rindler relative to an inertial observer (and vice-versa) has a number of non-trivial and remarkable implications. I will just mention 2 of them here:

1. The exponential redshift expressed by (7.138) implies that the right-moving energy densities in Minkowski and Rindler coordinates are related by

$$\frac{\partial u_M}{\partial u_R} = -au_M \quad \Rightarrow \quad (\partial_{u_M}\phi)^2 = \frac{1}{a^2 u_M^2}(\partial_{u_R}\phi)^2 \tag{7.139}$$

(and likewise for the left-movers). Thus essentially any classical solution that is regarded as regular by the Rindler observer (finite and non-zero ρ_R) corresponds to a divergent Minkowski energy-density as $u_M \to 0$, i.e. on the future boundary (horizon) $t = x$ of the Rindler wedge.

2. The exponential redshift expressed by (7.138) also implies that the notions of *positive frequency* with respect to Minkowski and Rindler time are inequivalent, e.g. in the sense that $f_\omega(u_M)$, restricted to the right Rindler-wedge $u_M < 0$, say, cannot be written as a superposition of Rindler right-moving positive frequency waves alone,

$$f_\omega(u_M) \neq \int_0^\infty d\omega' \alpha(\omega,\omega') f_{\omega'}(u_R) \ . \tag{7.140}$$

Of course, the $f_\omega(u_R)$ and their complex conjugates $f^*_\omega(u_R)$ provide a basis of solutions for the right-moving modes (in the right Rindler-wedge), so that one can certainly expand the Minkowski plane waves as

$$f_\omega(u_M) = \int_0^\infty d\omega' \ \left(\alpha(\omega,\omega') f_{\omega'}(u_R) + \beta(\omega,\omega') f^*_{\omega'}(u_R)\right) \ , \tag{7.141}$$

but necessarily with some of the $\beta(\omega, \omega') \neq 0$.

If you know a little bit of quantum field theory, you will be able to anticipate that this means that the notions of creation and annihilation operators are inequivalent, and that therefore what is the vacuum, say, for an inertial observer, will not be seen as the vacuum by the accelerating observer (and vice-versa).

Combining the two facts, one also arrives at the conclusion that the "Rindler vacuum" is singular both at the future horizon (from right-movers) and at the past horizon (from left-movers).

In the spirit of the equivalence principle ("before studying gravity, let us study accelerations in flat space"), this *Unruh Effect* is a fascinating and rewarding first step towards understanding (or appreciating the difficulties encountered by) quantum field theory in curved space-times, i.e. in non-trivial gravitational fields. For more on this see the references given in section 27.7.

As further examples of scalar fields in particular gravitational backgrounds, in section 26.8 we will consider scalar fields in the Schwarzschild space-time, and in section 34.10 we will look at the equations of motion of scalar fields in a cosmological gravitational background.

7.9 Conserved Currents from the Energy-Momentum Tensor?

In section 6.8 we had discussed how to obtain conserved charges from covariantly conserved currents. Now in special relativity one can construct conserved currents (corresponding to the generators of Poincaré transformations) from the conserved energy-momentum tensor, and hence from there the corresponding conserved charges like energy, momentum and angular momentum. In this section we will take a first look at the question if or to which extent we can also obtain such conserved currents from the covariantly conserved energy-momentum tensor in a gravitational field.

To set the stage, recall that in Special Relativity, if T^{ab} is the energy-momentum tensor of a physical system, it generally satisfies an equation of the form

$$\partial_a T^{ab} = G^b \quad , \tag{7.142}$$

where G^b represents the density of the external forces acting on the system. In particular, if there are no external forces, the divergence of the energy-momentum tensor is zero. For example, in the case of Maxwell theory and a current corresponding to a charged particle we have

$$G^b = J_a F^{ab} = -F^a_{\ b} J^b \sim -F^a_{\ b} \dot{\xi}^b \quad , \tag{7.143}$$

which is indeed the relevant external (Lorentz) force density (in writing this I have suppressed the δ-function that localises the current to the worldline $\xi^a = \xi^a(\tau)$ of the particle).

When there are no external forces, i.e. when one has taken into account the complete matter action, the total energy-momentum tensor is conserved. In that case, $T^{ab} = J^{(b)a}$ defines four conserved currents, more or less (modulo Belinfante improvement terms, see e.g. the discussion in sections 7.4 and 22.2 and the references given there) the currents associated to translation invariance of the action via Noether's theorem. One is thus in the setting of conserved currents of section 6.8, and one can define conserved quantities like total energy and momentum, P^a, and angular momentum J^{ab}, by integrals of T^{0a} or $\xi^a T^{0b} - \xi^b T^{0a}$ (the latter being conserved if T_{ab} is symmetric) over spacelike hypersurfaces.

The situation in general relativity is somewhat different (exactly how different it is perceived to be is partly a matter of personal preconceptions or desires). In particular, in general relativity, and assuming that $T_{\mu\nu}$ is the complete matter energy-momentum tensor (otherwise we certainly cannot expect to derive any conservation law), we will have a "conservation law" of the form

$$\nabla_\mu T^{\mu\nu} = g^{-1/2}\partial_\mu(g^{1/2}T^{\mu\nu}) + \Gamma^\nu{}_{\mu\lambda}T^{\mu\lambda} = 0 \quad . \tag{7.144}$$

We see that, due to the second term, this does not define four conserved currents in the ordinary or covariant sense (and we will return to the interpretation of this equation, and the related issue of energy and energy density of the gravitational field, in section 22.6).

Nevertheless, in analogy with special relativity, one might like to attempt to define conserved quantities like total energy and momentum, P^μ, and angular momentum $J^{\mu\nu}$, by integrals of $T^{0\mu}$ or $x^\mu T^{0\nu} - x^\nu T^{0\mu}$ over spacelike hypersurfaces. However, these quantities are rather obviously not covariant, and nor are they conserved.

This should perhaps not be too surprising because, after all, for a Poincaré-invariant field theory in Minkowski space these quantities are preserved as a consequence of Poincaré invariance, i.e. because of the symmetries (isometries) of the Minkowski metric (as well as of the action).

A generic metric has no isometries whatsoever (the explicit examples of metrics in these notes not withstanding, all of which exhibit at least some symmetries). As it has no symmetries, we have no reason to expect to find associated conserved quantities in general.

However, if there are symmetries then one should indeed be able to define conserved quantities (think of Noether's theorem again), one for each symmetry generator. In order to implement this we need to understand how to define and detect isometries of the metric. For this we need the concepts of Lie derivatives and Killing vectors. These already made occasional brief appearances in previous sections and will be discussed more systematically in section 9, the corresponding conserved charges then being the subject of section 10.

Alternatively, one might try to just go ahead optimistically and attempt to construct a covariant current-like object (with a corresponding conservation law and the ensuing possibility to define conserved charges) by contracting the energy-momentum tensor not with the coordinates but with a vector field V^λ, along the lines of

$$J_V^\mu = T^\mu_{\ \lambda} V^\lambda \quad . \tag{7.145}$$

At least this now has the merit of clearly being a vector field, but is it conserved? Calculating its covariant divergence, and using the fact that $T^{\mu\nu}$ is symmetric and conserved, one finds

$$\nabla_\mu J_V^\mu = \tfrac{1}{2} T^{\mu\nu}(\nabla_\mu V_\nu + \nabla_\nu V_\mu) \quad . \tag{7.146}$$

Thus we would have a conserved current (and associated conserved charge by the discussion in section 6.8) for any conserved energy-momentum tensor if the vector field V^λ were such that it satisfies

$$\nabla_\mu V_\nu + \nabla_\nu V_\mu = 0 \quad \Rightarrow \quad \nabla_\mu (T^\mu_{\ \lambda} V^\lambda) = 0 \quad . \tag{7.147}$$

The link between this observation and the one in the preceding paragraph regarding symmetries is that this is precisely the condition characterising (infinitesimal) symmetries of metric:

- First of all, this is the condition we already found and encountered in (3.35), as reformulated in (5.68), for the infinitesimal coordinate transformation $\delta x^\mu = \epsilon V^\mu$ to generate a symmetry of the metric, thus leading to a conserved charge for geodesics.

- More generally, as we will discuss in detail in section 9 below, vector fields satisfying the equation $\nabla_\mu V_\nu + \nabla_\nu V_\mu = 0$ are indeed in one-to-one correspondence with infinitesimal generators of continuous symmetries of a metric (isometries).

Thus this gives a satisfactory and coherent overall picture of symmetries and conservation laws in a gravitational field.

8 Curvature I: The Riemann Curvature Tensor

8.1 Curvature: Preliminary Remarks

We now come to one of the most important concepts of General Relativity and Riemannian Geometry, that of curvature and how to describe it in tensorial terms. Among other things, this will finally allow us to decide unambiguously if a given metric is just the (flat) Minkowski metric in disguise or the metric of a genuinely curved space (but a proof of this statement is postponed to section 11). More importantly (for present purposes) it will allow us to construct tensors that depend on the 2nd derivatives of the metric and will thus allow us to construct tensorial (generally covariant) differential equations for the metric. In particular, this will then lead us fairly directly to the Einstein equations (section 19), i.e. to the field equations for the gravitational field.

Recall that the equations that describe the behaviour of particles and fields in a gravitational field involve the metric and the Christoffel symbols determined by the metric. Thus the equations for the gravitational field should be generally covariant (tensorial) differential equations for the metric.

At first, here we seem to face a dilemma. How can we write down covariant differential equations for the metric when the covariant derivative of the metric is identically zero? Having come to this point, Einstein himself reached an impasse and required the help of his mathematician friend Marcel Grossmann ("Grossmann, you have to help me, or else I'll go crazy!") whom he had asked to investigate if there were any tensors that could be built from the second derivatives of the metric.

Grossmann soon found that this problem had indeed been addressed and solved in the mathematics literature, in particular by Riemann (generalising work of Gauss on curved surfaces), Ricci-Curbastro and Levi-Civita. It was shown by them that there are indeed non-trivial tensors that can be constructed from (ordinary) derivatives of the metric. These can then be used to write down covariant differential equations for the metric.[19]

The most important among these are the Riemann curvature tensor and its various contractions. In fact, it is known that these are the only tensors that can be constructed from the metric and its first and second derivatives, and they will therefore play a central role in all that follows.

Technically the most straightforward way of introducing the Riemann curvature tensor is via the commutator of covariant derivatives. In this section we will adopt this pragmatic (and relatively streamlined) approach, as it is sufficient to

[19] Of course, the story is not as simple and straightforward as that. For an account of Marcel Grossmann's (often overlooked) contributions to tensor calculus and the development of general relativity, see T. Sauer, *Marcel Grossmann and his contribution to the general theory of relativity*, `arXiv:1312.4068 [physics.hist-ph]`.

- determine the most important algebraic and differential properties of the curvature tensor (symmetries and Bianchi identities)
- assess its physical significance (gravitational tidal forces) via the influence of the curvature tensor on the motion of (families of) freely falling particles
- and to thus provide us with all the information and ingredients we need to then discuss the Einstein equations (section 19) and their formulation in terms of an action principle (section 20).

However, this is not geometrically the most intuitive way to introduce the concept of curvature, and it downplays the extent to which the curvature tensor reflects and encodes the geometric properties of space-time and, more generally, does not do justice to the fundamental differential geometric notion and significance of curvature. Some of these aspects are discussed in Part B of these notes, in particular in sections 11, 12, 13 and 14.

8.2 Riemann Tensor from the Commutator of Covariant Derivatives

As mentioned before, second covariant derivatives do not commute on (p,q)-tensors unless $p=q=0$. However, the fact that they *do* commute on scalars has the pleasant consequence that e.g. the commutator of covariant derivatives acting on a vector field V^μ does not involve any derivatives of V^μ. In fact, I will first show, without actually calculating the commutator, that

$$[\nabla_\mu,\nabla_\nu](\phi V^\lambda)=\phi[\nabla_\mu,\nabla_\nu]V^\lambda \tag{8.1}$$

for any scalar field ϕ. This implies that $[\nabla_\mu,\nabla_\nu]V^\lambda$ cannot depend on derivatives of V because if it did it would also have to depend on derivatives of ϕ.

Hence, the commutator can be expressed purely algebraically in terms of V. As the dependence on V is clearly linear, the commutator of covariant derivatives must then act like a linear transformation. There must therefore be an object $R^\lambda_{\ \sigma\mu\nu}$ such that

$$[\nabla_\mu,\nabla_\nu]V^\lambda=R^\lambda_{\ \sigma\mu\nu}V^\sigma \quad . \tag{8.2}$$

This can of course also be verified by a direct calculation, and we will come back to this below. For now let us just note that, since the left hand side of this equation is clearly a tensor for any V, the quotient theorem implies that the quantities $R^\lambda_{\ \sigma\mu\nu}$ are the components of a tensor.

Let us first verify (8.1). We have

$$\nabla_\mu\nabla_\nu(\phi V^\lambda)=(\nabla_\mu\nabla_\nu\phi)V^\lambda+(\nabla_\nu\phi)(\nabla_\mu V^\lambda)+(\nabla_\mu\phi)(\nabla_\nu V^\lambda)+\phi\nabla_\mu\nabla_\nu V^\lambda \quad . \tag{8.3}$$

Thus, upon taking the commutator the 2nd and 3rd terms drop out (because the 3rd is the symmetrisation of the 2nd), and we are left with

$$
\begin{aligned}
[\nabla_\mu, \nabla_\nu](\phi V^\lambda) &= ([\nabla_\mu, \nabla_\nu]\phi)V^\lambda + \phi[\nabla_\mu, \nabla_\nu]V^\lambda \\
&= \phi[\nabla_\mu, \nabla_\nu]V^\lambda \quad , \qquad (8.4)
\end{aligned}
$$

where the last line follows from the fact that 2nd covariant derivatives do commute on scalars. Thus we have established (8.1).

By explicitly calculating the commutator, one can confirm the structure displayed in (8.2). This explicit calculation shows that the *Riemann-Christoffel Curvature Tensor* (or Riemann tensor for short) is given by

$$
\boxed{R^\lambda_{\ \sigma\mu\nu} = \partial_\mu \Gamma^\lambda_{\ \sigma\nu} - \partial_\nu \Gamma^\lambda_{\ \sigma\mu} + \Gamma^\lambda_{\ \mu\rho}\Gamma^\rho_{\ \nu\sigma} - \Gamma^\lambda_{\ \nu\rho}\Gamma^\rho_{\ \mu\sigma}} \qquad (8.5)
$$

Remarks:

1. Note how useful the quotient theorem is in this case. It would be quite unpleasant to have to verify the tensorial nature of this expression by explicitly checking its behaviour under coordinate transformations.

2. Note also that this tensor is clearly zero for the Minkowski metric written in Cartesian coordinates. Hence it is also zero for the Minkowski metric written in any other coordinate system.

   In particular, if for some metric written in some coordinate system one finds that any component of the Riemann tensor is non-zero, this metric cannot possibly be the Minkowski metric written in some coordinates!

   We will prove the converse, that vanishing of the Riemann curvature tensor implies that the metric is (locally) equivalent to the Minkowski metric, in section 11.2.

3. In the above we have defined the Riemann tensor by the relation (8.2) and then deduced the explicit expression (8.5). While this is, pragmatically speaking, a useful way of proceeding, it may be more logical to initially define the Riemann tensor in a different way, e.g. directly by (8.5) (for instance because by painful calculations one has discovered that this particular combination of non-tensorial objects miraculously happens to transform as a tensor). In that case, (8.2) is a result rather than a definition, known as the *Ricci identity*.

It is straightforward to extend the above to an action of the commutator $[\nabla_\mu, \nabla_\nu]$ on arbitrary tensors. For covectors we have, since we can raise and lower the indices with the metric with impunity,

$$\begin{aligned} [\nabla_\mu, \nabla_\nu] V_\rho &= g_{\rho\lambda} [\nabla_\mu, \nabla_\nu] V^\lambda \\ &= g_{\rho\lambda} R^\lambda_{\ \sigma\mu\nu} V^\sigma \\ &= R_{\rho\sigma\mu\nu} V^\sigma \\ &= R_\rho^{\ \sigma}{}_{\mu\nu} V_\sigma \quad . \end{aligned} \tag{8.6}$$

We will see later that the Riemann tensor is anti-symmetric in its first two indices. Hence we can also write

$$[\nabla_\mu, \nabla_\nu] V_\rho = -R^\sigma_{\ \rho\mu\nu} V_\sigma \quad . \tag{8.7}$$

The extension to arbitrary (p,q)-tensors now follows the usual pattern, with one Riemann curvature tensor, contracted as for vectors, appearing for each of the p upper indices, and one Riemann curvature tensor, contracted as for covectors, for each of the q lower indices. Thus, e.g. for a $(2,0)$-tensor $T^{\alpha\beta}$ one has

$$[\nabla_\mu, \nabla_\nu] T^{\alpha\beta} = R^\alpha_{\ \gamma\mu\nu} T^{\gamma\beta} + R^\beta_{\ \gamma\mu\nu} T^{\alpha\gamma} \tag{8.8}$$

and for a (1,1)-tensor $A^\lambda_{\ \rho}$ one has

$$[\nabla_\mu, \nabla_\nu] A^\lambda_{\ \rho} = R^\lambda_{\ \sigma\mu\nu} A^\sigma_{\ \rho} - R^\sigma_{\ \rho\mu\nu} A^\lambda_{\ \sigma} \quad . \tag{8.9}$$

I will give two other versions of the fundamental formula (8.2) which are occasionally useful and used.

1. Instead of looking at the commutator $[\nabla_\mu, \nabla_\nu]$ of two derivatives in the coordinate directions x^μ and x^ν, we can look at the commutator $[\nabla_X, \nabla_Y]$ of two directional covariant derivatives aong the vector fields X, Y. Evidently, from

$$\nabla_X \nabla_Y = X^\mu \nabla_\mu Y^\nu \nabla_\nu = X^\mu Y^\nu \nabla_\mu \nabla_\nu + (X^\mu \nabla_\mu Y^\nu) \nabla_\nu \tag{8.10}$$

one finds that, in addition to the term defining the Riemann curvature tensor, the commutator $[\nabla_X, \nabla_Y]$ contains terms involving the covariant derivative along

$$\nabla_X Y^\mu - \nabla_Y X^\mu = X^\nu \nabla_\nu Y^\mu - Y^\nu \nabla_\nu X^\mu = X^\nu \partial_\nu Y^\mu - Y^\nu \partial_\nu X^\mu \tag{8.11}$$

(the terms involving involving the Christoffel symbols cancel by symmetry of the latter). Let us denote this vector field by $[X, Y]^\mu$, i.e.

$$[X, Y]^\mu = X^\nu \partial_\nu Y^\mu - Y^\nu \partial_\nu X^\mu \quad . \tag{8.12}$$

The rationale for this notation is that this is precisely the commutator of the vector fields X and Y regarded as 1st order differential operators $X = X^\mu \partial_\mu, Y = Y^\nu \partial_\nu$, as in (4.90). This will be explored in more detail in section 9.3. In any case, with this notation, and subtracting this term, we can now write the formula for the curvature tensor as

$$([\nabla_X, \nabla_Y] - \nabla_{[X,Y]}) V^\lambda = R^\lambda_{\ \sigma\mu\nu} X^\mu Y^\nu V^\sigma \quad . \tag{8.13}$$

2. Secondly, one can consider a net of curves $x^\mu(s_1, s_2)$ parametrising, say, a two-dimensional surface, and look at the commutators of the covariant derivatives along the s_1- and s_2-curves. Let us denote by

$$Z^\mu_{(1)} = \frac{\partial x^\mu}{\partial s_1} \quad , \quad Z^\mu_{(2)} = \frac{\partial x^\mu}{\partial s_2} \tag{8.14}$$

the tangent vector fields along the curves parametrised by s_1 and s_2 respectively, and by D_{s_k} the corresponding covariant derivative along the curves parametrised by s_k (section 5.7),

$$D_{s_k} = Z^\mu_{(k)} \nabla_\mu \quad . \tag{8.15}$$

The formula one obtains in this case is

$$(D_{s_1} D_{s_2} - D_{s_2} D_{s_1})\, V^\lambda = R^\lambda{}_{\sigma\mu\nu} Z^\mu_{(1)} Z^\nu_{(2)} V^\sigma \quad . \tag{8.16}$$

It can be obtained from (8.13) by noting that X and Y commute in this case. Indeed,

$$[Z_{(1)}, Z_{(2)}] = [\partial_{s_1}, \partial_{s_2}] = 0 \quad , \tag{8.17}$$

or, in more detail,

$$[Z_{(1)}, Z_{(2)}]^\mu = Z^\nu_{(1)} \partial_\nu Z^\mu_{(2)} - Z^\nu_{(2)} \partial_\nu Z^\mu_{(1)} = \frac{\partial^2 x^\mu}{\partial s_1 \partial s_2} - \frac{\partial^2 x^\mu}{\partial s_2 \partial s_1} = 0 \quad . \tag{8.18}$$

This can also be written as the statement that in terms of covariant derivatives one has

$$\frac{\partial^2 x^\mu}{\partial s_1 \partial s_2} = \frac{\partial^2 x^\mu}{\partial s_2 \partial s_1} \quad \Leftrightarrow \quad D_{s_1} Z^\mu_{(2)} = D_{s_2} Z^\mu_{(1)} \quad . \tag{8.19}$$

We will make use of the identity (8.16) in our discussion of tidal forces (the geodesic deviation equation) in section 8.4 below.

8.3 Symmetries and Algebraic Properties of the Riemann Tensor

A priori, the Riemann tensor has $256 = 4^4$ components in 4 dimensions. However, because of a large number of symmetries, the actual number of independent components is much smaller.

In general, to read off all the symmetries from the formula (8.5) is difficult. One way to simplify things is to look at the Riemann curvature tensor at the origin x_0 of a Riemann normal coordinate system (or some other inertial coordinate system). In that case, all the first derivatives of the metric disappear and only the first two terms of (8.5) contribute. One finds

$$\begin{aligned} R_{\alpha\beta\gamma\delta}(x_0) &= g_{\alpha\lambda}(\partial_\gamma \Gamma^\lambda{}_{\beta\delta} - \partial_\delta \Gamma^\lambda{}_{\beta\gamma})(x_0) \\ &= (\partial_\gamma \Gamma_{\alpha\beta\delta} - \partial_\delta \Gamma_{\alpha\beta\gamma})(x_0) \\ &= \tfrac{1}{2}(g_{\alpha\delta},_{\beta\gamma} + g_{\beta\gamma},_{\alpha\delta} - g_{\alpha\gamma},_{\beta\delta} - g_{\beta\delta},_{\alpha\gamma})(x_0) \quad . \end{aligned} \tag{8.20}$$

In principle, this expression is sufficiently simple to allow one to read off all the symmetries of the Riemann tensor. However, it is more insightful to derive these symmetries in a different way, one which will also make clear *why* the Riemann tensor has these symmetries.

1. Anti-symmetry in the second pair of indices:

$$R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma} \tag{8.21}$$

This is obviously true from the definition or by construction.

2. Anti-symmetry in the first pair of indices:

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} \tag{8.22}$$

This is a consequence of the fact that the metric is covariantly constant. In fact, we can calculate

$$\begin{aligned} 0 &= [\nabla_\gamma, \nabla_\delta] g_{\alpha\beta} \\ &= R_{\alpha\ \gamma\delta}^{\ \lambda} g_{\lambda\beta} + R_{\beta\ \gamma\delta}^{\ \lambda} g_{\alpha\lambda} \\ &= (R_{\alpha\beta\gamma\delta} + R_{\beta\alpha\gamma\delta}) \ . \end{aligned} \tag{8.23}$$

As mentioned before, this implies that we can write the commutator of covariant derivatives on a covector as

$$[\nabla_\mu, \nabla_\nu] V_\rho = R_{\rho\ \mu\nu}^{\ \sigma} V_\sigma = -R^\sigma_{\ \rho\mu\nu} V_\sigma \ . \tag{8.24}$$

3. Cyclic permutation symmetry (or first Bianchi identity)

$$R_{\alpha[\beta\gamma\delta]} = 0 \quad \Leftrightarrow \quad R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta} = 0 \tag{8.25}$$

This Bianchi identity is a consequence of the fact that there is no torsion. In fact, applying $[\nabla_\gamma, \nabla_\delta]$ to the covector $\nabla_\beta \phi$, ϕ a scalar, one has

$$\nabla_{[\gamma} \nabla_\delta \nabla_{\beta]} \phi = 0 \quad \Rightarrow \quad R^\lambda_{\ [\beta\gamma\delta]} \nabla_\lambda \phi = 0 \ . \tag{8.26}$$

As this has to be true for all scalars ϕ, this implies $R_{\alpha[\beta\gamma\delta]} = 0$ (to see this you could e.g. choose the (locally defined) coordinate functions $\phi^{(\mu)}(x) = x^\mu$ with $\nabla_\lambda \phi^{(\mu)} = \delta^\mu_\lambda$).

In turn this identity now implies that for any covector (not necessarily a gradient) one has the identity (also called Bianchi identity)

$$\nabla_{[\gamma} \nabla_\delta V_{\beta]} = -R^\alpha_{\ [\beta\gamma\delta]} V_\alpha = 0 \ . \tag{8.27}$$

4. Symmetry under exchange of the two pairs of indices

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta} \tag{8.28}$$

This identity, stating that the Riemann tensor is symmetric in its two pairs of indices, is not an independent symmetry but can be deduced from the three other symmetries by some not particularly interesting algebraic manipulations. One (quite possibly not optimal or minimal) possibility is

$$\begin{aligned}
R_{\gamma\delta\alpha\beta} &\overset{(3)}{=} -R_{\gamma\alpha\beta\delta} - R_{\gamma\beta\delta\alpha} \\
&\overset{(2)}{=} R_{\alpha\gamma\beta\delta} + R_{\beta\gamma\delta\alpha} \\
&\overset{(3)}{=} -R_{\alpha\delta\gamma\beta} - R_{\alpha\beta\delta\gamma} - R_{\beta\alpha\gamma\delta} - R_{\beta\delta\alpha\gamma} \\
&\overset{(1,2)}{=} 2R_{\alpha\beta\gamma\delta} + R_{\delta\alpha\gamma\beta} + R_{\delta\beta\alpha\gamma} \\
&\overset{(3)}{=} 2R_{\alpha\beta\gamma\delta} - R_{\delta\gamma\beta\alpha} \\
&\overset{(1,2)}{=} 2R_{\alpha\beta\gamma\delta} - R_{\gamma\delta\alpha\beta} \quad ,
\end{aligned} \tag{8.29}$$

from which the claim follows.

Slightly more elegant (but equally obtuse) is the following argument.[20] Consider the matrix

$$\mathcal{R} = \begin{pmatrix} R_{\alpha\beta\gamma\delta} & R_{\alpha\delta\beta\gamma} & R_{\alpha\gamma\delta\beta} \\ R_{\delta\alpha\beta\gamma} & R_{\delta\gamma\alpha\beta} & R_{\delta\beta\gamma\alpha} \\ \vdots & \ldots & \ldots \end{pmatrix} \tag{8.30}$$

where the first column consists of the 4 cyclic permutations of all 4 indices of the Riemann tensor, while each row consists of the 3 cyclic permutations of the last 3 indices. Thus the sum σ_k of the entries of the k'th row is zero for all k (by symmetry (3)),

$$\sigma_k = \sum_{l=1}^{3} \mathcal{R}_{kl} \overset{(3)}{=} 0 \quad , \tag{8.31}$$

while

$$\sigma_1 + \sigma_2 - \sigma_3 - \sigma_4 \overset{(1,2)}{=} 2R_{\alpha\beta\gamma\delta} - 2R_{\gamma\delta\alpha\beta} \quad . \tag{8.32}$$

We can now count how many independent components the Riemann tensor really has. (1) implies that the second pair of indices can only take $N = (4 \times 3)/2 = 6$ independent values. (2) implies the same for the first pair of indices. (4) thus says that the Riemann curvature tensor behaves like a symmetric (6×6) matrix and therefore has $(6 \times 7)/2 = 21$ components. We now come to the remaining condition (3): if two of the indices in (3) are equal, (3) is equivalent to (4) and (4) we have already taken into account. With

[20] See e.g. D. Bleecker, *Gauge Theory and Variational Principles.*

all indices unequal, (3) then provides one and only one more additional constraint. We conclude that the total number of independent components is 20.

<u>Remarks:</u>

1. Note that this agrees precisely with our previous counting in section 3.6 of how many of the second derivatives of the metric cannot be set to zero by a coordinate transformation: the second derivative of the metric has 100 independent components, to be compared with the $4 \times (4 \times 5 \times 6)/(2 \times 3) = 80$ components of the third derivatives of the coordinates. This also leaves 20 components. We thus see very explicitly that the Riemann curvature tensor contains all the coordinate independent information about the geometry up to second derivatives of the metric. In fact, it can be shown that in a Riemann normal coordinate system one has

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + 0 + \tfrac{1}{3}R_{\mu\lambda\sigma\nu}(x_0)(x-x_0)^\lambda(x-x_0)^\sigma + \mathcal{O}((x-x_0)^3) \quad . \tag{8.33}$$

2. Just for the record, I note here that in general dimension $D = d+1$ the Riemann tensor has $D^2(D^2-1)/12$ independent components. This number arises as

$$\begin{aligned} \frac{D^2(D^2-1)}{12} &= \frac{N(N+1)}{2} - \binom{D}{4} \\ N &= \frac{D(D-1)}{2} \end{aligned} \tag{8.34}$$

and describes (as above) the number of independent components of a symmetric $(N \times N)$-matrix, now subject to $\binom{D}{4}$ conditions which arise from all the possibilities of choosing 4 out of D possible distinct values for the indices in (3). Just as for $D = 4$, this number of components of the Riemann tensor coincides with the number of second derivatives of the metric minus the number of independent components of the third derivatives of the coordinates determined in (3.179),

$$\frac{D(D+1)}{2} \times \frac{D(D+1)}{2} - D \times \frac{D(D+1)(D+2)}{2 \times 3} = \frac{D^2(D^2-1)}{12} \quad . \tag{8.35}$$

For $D = 2$ this formula predicts one independent component, and this is as it should be. Rather obviously the only independent non-vanishing component of the Riemann tensor in this case is R_{1212}. We will discuss curvature in 2 dimensions in more detail in sections 8.6 and 11.3 below.

Finally, a word of warning: there are a large number of sign conventions involved in the definition of the Riemann tensor (and its contractions we will discuss below), so whenever reading a book or article, in particular when you want to use results or equations presented there, make sure what conventions are being used and either adopt those or translate the results into some other convention. As a check: the conventions used here are such that $R_{\phi\theta\phi\theta}$ as well as the curvature scalar (to be introduced below) are *positive* for the standard metric on the two-sphere.

8.4 Tidal Forces: Influence of Curvature on Particle Trajectories

In a certain sense the main effect of curvature (or gravity) is that initially parallel trajectories of freely falling non-interacting particles (dust, pebbles,...) do not remain parallel, i.e. that gravity is an attractive force that has the tendency to focus matter. This statement find its mathematically precise formulation in equations describing the influence of space-time curvature on the behaviour of (families of) geodesics.

Let us, as we will need this later anyway, recall first the situation in the Newtonian theory. One particle moving under the influence of a gravitational field is governed by the equation

$$\tfrac{d^2}{dt^2}x^i = -\partial^i\phi(x) \quad , \tag{8.36}$$

where ϕ is the potential. Now consider a family of particles, or just two nearby particles, one at $x^i(t)$ and the other at $x^i(t)+\delta x^i(t)$. The other particle will of course obey the equation

$$\tfrac{d^2}{dt^2}(x^i+\delta x^i) = -\partial^i\phi(x+\delta x) \quad . \tag{8.37}$$

From these two equations one can deduce an equation for δx itself, namely

$$\tfrac{d^2}{dt^2}\delta x^i = -\partial^i\partial_j\phi(x)\delta x^j \quad . \tag{8.38}$$

It describes the effect of *gravitational tidal forces* (the gradient of the gravitational force) on a family of particles moving in a gravitational field.

In particular, when there is no gravitational force, and the trajectories are straight lines, one has

$$\tfrac{d^2}{dt^2}\delta x^i = 0 \quad \Rightarrow \quad \delta x^i = (\delta x^i)_0 + (\delta v^i)t \quad . \tag{8.39}$$

Thus one recovers Euclid's parallel axiom, that two straight lines intersect at most once (for suitable choices of $\delta v^i \neq 0$) and that they never intersect when they are initially parallel ($\delta v^i = 0$). Any departure from this equation or its Minkowskian counterpart

$$\tfrac{d^2}{d\tau^2}\delta\xi^a = 0 \quad , \tag{8.40}$$

will therefore indicate a departure from Euclidean geometry!

It is the counterpart of (8.38) that we will be seeking in the context of General Relativity. One derivation of this can be modelled on the Newtonian derivation above. It is elementary but looks non-covariant (and therefore somewhat messy) at intermediate stages of the calculation. I will then give you a very quick manifestly covariant derivation of the same result using the formula (8.16). Other aspects and generalisations of the geodesic deviation equation are discussed in section 12.1.

The starting point for the 1st derivation is of course, by analogy with the above Newtonian discussion, the geodesic equation for x^μ and for its nearby partner $x^\mu + \delta x^\mu$,

$$\tfrac{d^2}{d\tau^2}x^\mu + \Gamma^\mu_{\ \nu\lambda}(x)\tfrac{d}{d\tau}x^\nu\tfrac{d}{d\tau}x^\lambda = 0 \quad , \tag{8.41}$$

and

$$\tfrac{d^2}{d\tau^2}(x^\mu + \delta x^\mu) + \Gamma^\mu_{\ \nu\lambda}(x+\delta x)\tfrac{d}{d\tau}(x^\nu + \delta x^\nu)\tfrac{d}{d\tau}(x^\lambda + \delta x^\lambda) = 0 \ \ . \tag{8.42}$$

As above, from these one can deduce an equation for δx, namely

$$\tfrac{d^2}{d\tau^2}\delta x^\mu + 2\Gamma^\mu_{\ \nu\lambda}(x)\tfrac{d}{d\tau}x^\nu\tfrac{d}{d\tau}\delta x^\lambda + \partial_\rho\Gamma^\mu_{\ \nu\lambda}(x)\delta x^\rho\tfrac{d}{d\tau}x^\nu\tfrac{d}{d\tau}x^\lambda = 0 \ \ . \tag{8.43}$$

Now this does not look particularly covariant. Thus instead of in terms of $d/d\tau$ we would like to rewrite this in terms of the covariant operator D_τ, with

$$D_\tau\delta x^\mu = \frac{d}{d\tau}\delta x^\mu + \Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\delta x^\lambda \ \ . \tag{8.44}$$

Calculating $(D_\tau)^2\delta x^\mu$, replacing $\ddot{x}^\mu$ appearing in that expression by $-\Gamma^\mu_{\ \nu\lambda}\dot{x}^\nu\dot{x}^\lambda$ (because x^μ satisfies the geodesic equation) and using (8.43), one eventually finds the nice covariant *geodesic deviation equation*

$$(D_\tau)^2\delta x^\mu = R^\mu_{\ \nu\lambda\rho}\dot{x}^\nu\dot{x}^\lambda\delta x^\rho \tag{8.45}$$

<u>Remarks:</u>

1. This shows very clearly that curvature, as captured by the Riemann curvature tensor, leads to non-Euclidean geometry in which e.g. the parallel axiom is not necessarily satisifed.

2. In general, solutions to the geodesic deviation equation are called *Jacobi fields*. They describe the difference between the given geodesic and a (hypothetical) infinitely close neighbouring geodesic.

Clearly the present derivation of this result leaves something to be desired. It is also possible to give a manifestly covariant, and thus perhaps slightly more satisfactory, derivation of the above geodesic deviation equation,

Thus we think of a family of geodesics as given by $x^\alpha = x^\alpha(\tau, s)$, where the parameter s labels the individual geodesisc, parametrised by proper time (affine parameter) τ. We introduce the tangent vectors

$$u^\alpha = \frac{\partial x^\alpha}{\partial\tau} \quad , \quad \delta x^\alpha \equiv \xi^\alpha = \frac{\partial x^\alpha}{\partial s} \ \ , \tag{8.46}$$

and the corresponding covariant derivatives D_τ, D_s along the curves. Now all we need to use are the identity (8.16) for the commutator $[D_\tau, D_s]$, the fact that the u^α are geodesics,

$$D_\tau u^\alpha = 0 \ \ , \tag{8.47}$$

and the fact that by (8.19) the vector fields u^α and ξ^α satisfy

$$\frac{\partial^2 x^\alpha}{\partial s\partial\tau} = \frac{\partial^2 x^\alpha}{\partial\tau\partial s} \quad \Leftrightarrow \quad D_s u^\alpha = D_\tau\xi^\alpha \ \ . \tag{8.48}$$

Then one can simply calculate the desired quantity $(D_\tau)^2\xi^\alpha$ by

$$(D_\tau)^2\xi^\alpha = D_\tau D_s u^\alpha = [D_\tau, D_s]u^\alpha = R^\alpha_{\beta\gamma\delta}u^\gamma\xi^\delta u^\beta \tag{8.49}$$

which is identical to (8.45). I hope you agree that this derivation is much more satisfactory than the 1st derivation! For some more details and generalisations see also also the discussion in section 12, in particular section 12.1.

8.5 Contractions of the Riemann Tensor: Ricci Tensor and Ricci Scalar

The Riemann tensor, as we have seen, is a four-index tensor. For many purposes this is not the most useful object, but we can create new tensors by contractions of the Riemann tensor. Due to the symmetries of the Riemann tensor, there is essentially only one possibility, namely the *Ricci tensor*

$$R_{\mu\nu} := R^\lambda{}_{\mu\lambda\nu} = g^{\lambda\sigma}R_{\sigma\mu\lambda\nu} \quad . \tag{8.50}$$

It arises naturally from the definition (8.2) of the Riemann tensor in terms of commutators of covariant derivatives, when one considers a contracted commutator,

$$[\nabla_\mu, \nabla_\nu]V^\lambda = R^\lambda{}_{\sigma\mu\nu}V^\sigma \quad \Rightarrow \quad [\nabla_\mu, \nabla_\nu]V^\mu = R^\mu{}_{\sigma\mu\nu}V^\sigma \equiv R_{\sigma\nu}V^\sigma \quad . \tag{8.51}$$

In particular, this identity explains why the Maxwell equations in the covariant Lorenz gauge (6.46) take the non-minimally coupled form (6.47).

It follows from the symmetries of the Riemann tensor that $R_{\mu\nu}$ is symmetric. Indeed

$$R_{\nu\mu} = g^{\lambda\sigma}R_{\sigma\nu\lambda\mu} = g^{\lambda\sigma}R_{\lambda\mu\sigma\nu} = R^\sigma{}_{\mu\sigma\nu} = R_{\mu\nu} \quad . \tag{8.52}$$

Thus, for $D = 4$, the Ricci tensor has 10 independent components, for $D = 3$ it has 6, while for $D = 2$ there is only 1 because there is only one independent component of the Riemann curvature tensor to start off with.

There is one more contraction of the Riemann tensor we can perform, namely on the Ricci tensor itself, to obtain what is called the *Ricci scalar* or *curvature scalar*

$$R := g^{\mu\nu}R_{\mu\nu} \quad . \tag{8.53}$$

Remarks:

1. One might have thought that at least in four dimensions there is another way of constructing a (pseudo-)scalar, by contracting the Riemann tensor with the Levi-Civita tensor, but

$$\epsilon^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} = 0 \tag{8.54}$$

because of the Bianchi identity (cyclic symmetry of the Riemann tensor).

2. Note that for $D = 2$ the Riemann curvature tensor has as many independent components as the Ricci scalar, namely one, and that for $D = 3$ the Ricci tensor has as many components as the Riemann tensor, namely 6. Thus in $D = 2$ one can express the entire Riemann tensor in terms of the Ricci scalar (and the metric) alone, and one has

$$D = 2: \quad R_{\alpha\beta\gamma\delta} = \tfrac{1}{2}(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma})R \tag{8.55}$$

(we will establish this relation in section 11.3, see (11.29)), while in $D = 3$ one has

$$\begin{aligned} D = 3: \quad R_{\alpha\beta\gamma\delta} &= (g_{\alpha\gamma}R_{\beta\delta} + R_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}R_{\beta\gamma} - R_{\alpha\delta}g_{\beta\gamma}) \\ &\quad + \tfrac{1}{2}(g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta})R \end{aligned} \tag{8.56}$$

(and we will prove this in section 11.4).

3. It is thus only in four (and more) dimensions that there are strictly less components of the Ricci tensor than of the Riemann tensor. This has profound implications for the dynamics of gravity in these dimensions. In fact, we will see that it is only in dimensions $D > 3$ that gravity becomes truly dynamical, where empty space can be curved, where gravitational waves can exist etc.

4. Contracting (8.8), one consequence of the symmetry of the Ricci tensor is the useful general result

$$[\nabla_\mu, \nabla_\nu]T^{\mu\nu} = R_{\mu\nu}(T^{\mu\nu} - T^{\nu\mu}) = 0 \tag{8.57}$$

for any tensor $T^{\mu\nu}$. If $T^{\mu\nu} = F^{\mu\nu}$ is anti-symmetric, $F^{\mu\nu} = -F^{\nu\mu}$, it is not necessary to take the commutator, so one also has

$$F^{\mu\nu} = -F^{\nu\mu} \quad \Rightarrow \quad \nabla_\mu(\nabla_\nu F^{\mu\nu}) = \tfrac{1}{2}[\nabla_\mu, \nabla_\nu]F^{\mu\nu} = 0 \ . \tag{8.58}$$

Note that this can also be deduced (without knowing anything about curvature in general or the Ricci tensor in particular) from the general expression (5.66) for the divergence of an anti-symmetric tensor,

$$\nabla_\mu F^{\mu\nu} = g^{-1/2}\partial_\mu(g^{1/2}F^{\mu\nu}) \quad \Rightarrow \quad \nabla_\nu\nabla_\mu F^{\mu\nu} = g^{-1/2}\partial_\nu\partial_\mu(g^{1/2}F^{\mu\nu}) = 0 \ . \tag{8.59}$$

This is how we had shown in section 6.6 that the Maxwell equations imply covariant current conservation,

$$\nabla_\mu F^{\mu\nu} = -J^\nu \quad \Rightarrow \quad \nabla_\nu J^\nu = 0 \ . \tag{8.60}$$

Now we see that we can alternatively directly use the identity (8.58) to arrive at this result.

5. There are other scalars that can be built from the curvature tensor, but these are necessarily of higher order in the curvature tensor, such as (trivially) R^2 or (somewhat less trivially) $R_{\mu\nu}R^{\mu\nu}$ or the square of the Riemann tensor, the so-called *Kretschmann scalar*

$$K = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} \quad . \tag{8.61}$$

   Analogously, scalars can be built from higher powers of the Riemann tensor and or from powers of covariant derivatives of the Riemann tensor ($\Box R$ being the simplest example).

6. Such scalars are useful in analysing a given metric because, since they are scalars they are invariant under coordinate transformations. Thus they directly provide coordinate-invariant information about a metric. For instance if K is singular at some point in some coordinate system then it will be singular at that point in all coordinate systems, and thus such a singularity is not an artefact of a bad choice of coordinate system but a property of the space(-time) itself described by that metric. A prominent example is the singularity at the origin $r = 0$ of the Schwarzschild metric, unambiguously unveiled by the singularity of its Kretschmann scalar (27.163).

7. Contracting (8.51) with V^ν, one finds

$$V^\nu\nabla_\mu\nabla_\nu V^\mu - V^\nu\nabla_\nu\nabla_\mu V^\mu = R_{\mu\nu}V^\mu V^\nu \quad . \tag{8.62}$$

   Rewriting the first term as

$$V^\nu\nabla_\mu\nabla_\nu V^\mu = \nabla_\mu(V^\nu\nabla_\nu V^\mu) - (\nabla_\mu V^\nu)(\nabla_\nu V^\mu) \tag{8.63}$$

   this identity can be written as

$$V^\nu\nabla_\nu(\nabla_\mu V^\mu) + (\nabla_\mu V_\nu)(\nabla^\nu V^\mu) - \nabla_\mu(V^\nu\nabla_\nu V^\mu) + R_{\mu\nu}V^\mu V^\nu = 0 \quad . \tag{8.64}$$

   This is a very useful and versatile "master equation" which provides valuable information about the relation between vector fields and curvature when specialised e.g. to geodesic vector fields, $V^\nu\nabla_\nu V^\mu = 0$, or Killing vector fields, $\nabla_\mu V_\nu = -\nabla_\nu V_\mu$ and $\nabla_\mu V^\mu = 0$. Various specialisations of this equation will therefore appear later on in these notes, and even though we will then usually rederive them from scratch in the case at hand, it is good to keep in mind that e.g. (12.22) (our starting point for the discussion of the Raychaudhuri equation in section 12.2) and (13.12) (a useful identity relating Killing vectors and curvature) are special cases of (8.64).

8. As an a(far)side, and as an illustration of what one can do with (8.64), assume that V is such that its curl $\nabla_\mu V_\nu - \nabla_\nu V_\mu = 0$ and its divergence $\nabla_\mu V^\mu = 0$ are zero. Locally, the first condition has the solution $V_\mu = \partial_\mu f$, and then the second condition says that $\Box f = 0$, i.e. that f is *harmonic*. Therefore let us call a

V^μ *harmonic* if it satisfies the above two conditions (to the mathematically more sophisticated: yes, I know that this is backwards, but we will specialise to the compact Riemannian case below ...).

For V harmonic in this sense, (8.64) reduces to

$$(\nabla_\mu V_\nu)(\nabla^\mu V^\nu) + R_{\mu\nu}V^\mu V^\nu = \nabla_\mu(V^\nu \nabla_\nu V^\mu) \quad . \tag{8.65}$$

The simplest (albeit perhaps not of most direct relevance for physics) situation where one can deduce something of substance from this equation is when one has a Riemannian (i.e. positive-definite) metric and the space one is dealing with is compact, without boundary. Then (a) the first term is non-negative, and (b) upon integration over the space the total derivative term on the right-hand side gives zero upon use of the Gauss theorem (5.63) (discussed in some more detail in section 16.3).

This implies that for a harmonic V to exist on such a space, the integral of $R_{\mu\nu}V^\mu V^\nu$ must be non-positive. In particular,

- if the Ricci tensor is positive (as a quadratic form), there are no harmonic vector fields at all,
- and if $R_{\mu\nu}V^\mu V^\nu$ =0, then a harmonic vector field is necessarily covariantly constant, $\nabla_\mu V_\nu = 0$.

In more mathematical terms this means that the first Betti number of a compact manifold admitting a metric with positive Ricci curvature is equal to zero. A variant of this kind of argument for Killing vectors will be given in section 13.3.[21]

8.6 Example: Curvature Tensor of the 2-Sphere

To see how calculations of the curvature tensor can be done in practice, let us work out the example of the two-sphere of unit radius, i.e. with line element

$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2 \equiv g_{ab}dx^a dx^b \quad . \tag{8.66}$$

We already know that the non-zero Christoffel symbols necessarily have two ϕ-indices and one θ-index (from $g_{\phi\phi} = \sin^2\theta$), and are given by

$$\Gamma^\phi{}_{\phi\theta} = \cot\theta \ , \qquad \Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta \quad . \tag{8.67}$$

We also know that the Riemann curvature tensor has only one independent component. Let us therefore work out $R^\theta{}_{\phi\theta\phi}$. From the definition we find

$$R^\theta{}_{\phi\theta\phi} = \partial_\theta \Gamma^\theta{}_{\phi\phi} - \partial_\phi \Gamma^\theta{}_{\theta\phi} + \Gamma^\theta{}_{\theta c}\Gamma^c{}_{\phi\phi} - \Gamma^\theta{}_{\phi c}\Gamma^c{}_{\theta\phi} \quad . \tag{8.68}$$

[21] Theorems of this kind, and proved with the help of these kinds of techniques, were pioneered by S. Bochner and K. Yano in the 1930s - 1950s. See e.g. S. Bochner, K. Yano, *Curvature and Betti Numbers* or W. Poor, *Differential Geometric Structures* for more details, rigour and applications.

The second and third terms are manifestly zero, and we are left with

$$R^{\theta}{}_{\phi\theta\phi} = \partial_\theta(-\sin\theta\cos\theta) + \sin\theta\cos\theta\cot\theta = \sin^2\theta \quad . \tag{8.69}$$

Thus we have

$$\begin{aligned} R^{\theta}{}_{\phi\theta\phi} &= R_{\theta\phi\theta\phi} = \sin^2\theta \\ R^{\phi}{}_{\theta\phi\theta} &= 1 \quad . \end{aligned} \tag{8.70}$$

Therefore the Ricci tensor R_{ab} has the components

$$\begin{aligned} R_{\theta\theta} &= 1 \\ R_{\theta\phi} &= 0 \\ R_{\phi\phi} &= \sin^2\theta \quad . \end{aligned} \tag{8.71}$$

These equations can succinctly be written as

$$R_{ab} = g_{ab} \quad , \tag{8.72}$$

showing that the standard metric on the two-sphere is what we will later call an *Einstein metric*. The Ricci scalar R is

$$R = g^{\theta\theta}R_{\theta\theta} + g^{\phi\phi}R_{\phi\phi} = 1 + \frac{1}{\sin^2\theta}\sin^2\theta = 2 \quad . \tag{8.73}$$

In particular, we have here our first concrete example of a space with non-trivial, in fact positive, curvature.

The result for the Riemann tensor can be written succinctly as

$$R^{a}{}_{bcd} = \delta^{a}{}_{c}g_{bd} - \delta^{a}{}_{d}g_{bc} \quad , \tag{8.74}$$

which also immediately implies (8.72),

$$R_{bd} = R^{a}{}_{bad} = g_{bd} \quad . \tag{8.75}$$

We will see later on, in section 14, that this form of the curvature tensor, or its equivalent,

$$R_{abcd} = g_{ac}g_{bd} - g_{ad}g_{bc} \quad , \tag{8.76}$$

is characteristic of the curvature tensor of the sphere in any dimension.

8.7 More Examples: Curvature Tensor and Polar/Spherical Coordinates

We now turn to some variations of the above theme (and some other generalisations are discussed in section 11.3 below).

1. First of all, let us address the question what is the curvature (scalar) of a sphere of radius L, i.e. of the space with line element

$$ds^2 = L^2(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{8.77}$$

   There are at least 3 ways to answer this question:

- The first is to simply and blindly redo the above calculations in this case and to see what one gets.
- Alternatively, and somewhat more insightfully, rather than redoing the calculation in that case one can argue as follows. Let us observe first of all that the Christoffel symbols are invariant under constant rescalings of the metric because they are schematically of the form $g^{-1}\partial g$. Therefore the Riemann curvature tensor, which only involves derivatives and products of Christoffel symbols, is also invariant. Hence the Ricci tensor, which is just a contraction of the Riemann tensor, is also invariant:

  $$g_{ab} \to L^2 g_{ab} \quad \Rightarrow \quad \Gamma^a_{bc} \to \Gamma^a_{bc} \quad \Rightarrow \quad R^a_{bcd} \to R^a_{bcd} \quad \Rightarrow \quad R_{bd} \to R_{bd} \ . \tag{8.78}$$

  However, to construct the Ricci scalar, one needs the inverse metric. This introduces an explicit L-dependence and the result is that the curvature scalar of a sphere of radius L is $R = 2/L^2$,

  $$R(L^2 g_{ab}) = L^{-2} R(g_{ab}) = 2/L^2 \ . \tag{8.79}$$

  In particular, the curvature scalar of a large sphere is smaller than that of a small sphere, something which makes intuitve sense, a very large sphere locally "looking flatter" than a small sphere. However, one should use this intuition with care since, as we have seen, e.g. the Ricci tensor is independent of the size of the sphere.
- Finally, this result could also have been obtained on purely dimensional grounds. The curvature scalar is constructed from second derivatives of the metric. Hence it has length-dimension (-2). Therefore for a sphere of radius L, R has to be proportional to $1/L^2$. Comparing with the known result for $L = 1$ determines $R = 2/L^2$, as before.

2. Now let us consider, instead of the unit 2-sphere, the unit hyperboloid H^2 with metric (2.31)

   $$ds^2(H^2) = d\sigma^2 + \sinh^2\sigma \, d\phi^2 \ . \tag{8.80}$$

   It is clear that, apart from a few sign changes here and there, the calculation of the Riemann curvature tensor is identical to that for S^2. These sign changes ultimately lead to the conclusion that the curvature scalar of H^2 is (-2). While the sphere is the prototypical example of a space with positive curvature, the hyperboloid is the prototypical example of a space with negative curvature.

   Instead of just doing the calculation for this specific example, it is slightly more instructive to do it for the class of metrics

   $$ds^2 = dx^2 + f(x)^2 d\phi^2 \ , \tag{8.81}$$

for some (for the time being unspecified) function $f = f(x)$. Denoting the derivative with respect to x by a prime, $f'(x) = df/dx$, one finds (this is a simple but constructive exercise) that the Ricci tensor and Ricci scalar are

$$R_{ab} = -(f''/f)g_{ab} \quad , \quad R(x) = -2f''(x)/f(x) \ . \tag{8.82}$$

In particular, for the Euclidean metric and the standard metrics on the sphere and the hyperboloid one finds

$$f(x) = \begin{cases} x & (R^2) \\ \sin x & (S^2) \\ \sinh x & (H^2) \end{cases} \quad \Rightarrow \quad R = \begin{cases} 0 \\ +2 \\ -2 \end{cases} \tag{8.83}$$

In 2 dimensions, R is related to the *Gauss Curvature* K of a surface by $K = R/2$ so that $K = 0, \pm 1$ in these examples. See section 11.3 for some more information.

3. Now let us promote the constant radius L of S^2 to a new radial coordinate r and ask the question what is the curvature tensor of the 3-dimensional space with coordinates $(r, x^a) = (r, \theta, \phi)$ and line element

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{8.84}$$

On the one hand, because one seems to have just added a trivial r-direction to the 2-sphere, one might be tempted to suspect that also this 3-dimensional space has non-trivial curvature. On the other hand, we recognise the above metric as the Euclidean metric on $\mathbb{R}^3$, written in spherical coordinates, and as such we expect its curvature (in fact, all components of the Riemann tensor) to be zero.

The latter expectation is of course borne out, but it is instructive to see explicitly how this cancellation occurs. In fact, it will be even more instructive to consider an apparently harmless and innocuous modification of the above metric which consists in replacing dr^2 by some constant multiple of dr^2,

$$ds^2 = p\, dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{8.85}$$

Equivalently, up to a truly harmless overall constant factor, we can think of this as the Euclidean metric, but with the metric on the unit-sphere replaced by that of a sphere of radius $1/\sqrt{p} \neq 1$),

$$ds^2 = p\, \left(dr^2 + (r^2/p)(d\theta^2 + \sin^2\theta d\phi^2)\right) \ . \tag{8.86}$$

Proceeding in a pedestrian way, we thus have a metric $g_{\alpha\beta}$ with components

$$g_{rr} = p \quad , \quad g_{ar} = 0 \quad , \quad g_{ab} = r^2\gamma_{ab} \ , \tag{8.87}$$

with γ_{ab} in this example denoting the components of the metric on the unit sphere (and with γ^a_{bc} and r^a_{bcd} its associated Christoffel symbols and components of the

Riemann curvature tensor determined in the previous section). From these we can deduce that for $r > 0$ the non-trivial Christoffel symbols are

$$\Gamma^r{}_{ab} = -p^{-1} r \gamma_{ab} \quad , \quad \Gamma^a_{\ br} = r^{-1}\delta^a_b \quad , \quad \Gamma^a_{\ bc} = \gamma^a_{\ bc} \ . \tag{8.88}$$

From this, in turn, one finds that all the components of the Riemann tensor involving at least one r-index are zero, whereas for the purely angular components one finds

$$R^a_{\ bcd} = r^a_{\ bcd} + \Gamma^a_{\ cr}\Gamma^r{}_{bd} - \Gamma^a_{\ dr}\Gamma^r{}_{bc} \ . \tag{8.89}$$

Using (8.74) and (8.88), one sees that

$$R^a_{\ bcd} = (1 - p^{-1}) r^a_{\ bcd} = (1 - p^{-1})(\delta^a_{\ c}\gamma_{bd} - \delta^a_{\ d}\gamma_{bc}) \ . \tag{8.90}$$

Therefore precisely for $p = 1$ the two contributions to the curvature tensor indeed cancel and the curvature tensor is identically zero, as expected.

Equally interesting is the fact that for $p \neq 1$ the curvature is non-zero even away from $r = 0$ (in addition, there is a conical deficit angle singularity at $r = 0$, as in the next example below, but this shall not be our concern here). In particular it follows from the above result that the only non-vanishing components of the Ricci tensor of this 3-dimensional space are

$$R^a_{\ bcd} = (1 - p^{-1}) r^a_{\ bcd} \quad \Rightarrow \quad R_{bd} = (1 - p^{-1}) r_{bd} = (1 - p^{-1})\gamma_{bd} \ . \tag{8.91}$$

Therefore also its Ricci scalar is non-zero,

$$R = g^{\alpha\beta} R_{\alpha\beta} = g^{ab} R_{ab} = 2(1 - p^{-1}) r^{-2} \ . \tag{8.92}$$

We also see from this that this space actually has a curvature singularity as $r \to 0$. Since the Ricci scalar is a scalar (under coordinate transformations), this divergence cannot be an artefact of a bad choice of coordinates, and indicates that there is a genuine geometric singularity for $r \to 0$.

Extended to a four-dimensional space-time metric via

$$ds^2 = -dt^2 + p\, dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ , \tag{8.93}$$

this describes the gravitational field outside a "monopole".[22]

4. As a final variation of this theme, we consider the above example in one dimension less, i.e. we look at the metric one obtains if one replaces the Euclidean metric on $\mathbb{R}^2$ written in polar coordinates by

$$dr^2 + r^2 d\phi^2 \to p\, dr^2 + r^2 d\phi^2 \ , \tag{8.94}$$

[22] M. Barriola, A. Vilenkin, *Gravitational Field of a Global Monopole*, Phys. Rev. Lett. 63 (1989) 341-343.

where the angle ϕ has period 2π.

In this case there is an interesting twist (pun intended) and the situation is somewhat different. Pulling out the factor of p, one sees that (up to this irrelevant overall constant factor) the metric can be written as

$$ds^2 = dr^2 + r^2 d(\phi/\sqrt{p})^2 \equiv dr^2 + C^2 r^2 d\phi^2 \ . \tag{8.95}$$

This would be the standard Euclidean metric on $\mathbb{R}^2$ either for $p = 1$ or if the angle ϕ had periodicity $2\pi\sqrt{p}$, but since ϕ has period 2π, this results in a misidentification of the points in a plane, like when one rolls up a flat piece of paper into a cone. Away from $r = 0$, this space is intrinsically flat (all the components of the Riemann curvature tensor are zero, as one can easily calculate - see section 11.1 for an explanation of this use of the word "intrinsic"). There is, however, a conical singularity at the tip of the cone $r = 0$, which can be thought of as a δ-function contribution to the curvature localised at $r = 0$. Extended to a four-dimensional space-time metric,

$$ds^2 = -dt^2 + dz^2 + dr^2 + C^2 r^2 d\phi^2 \ , \tag{8.96}$$

it can be interpreted as the space-time metric of an idealised cosmic string extended in the z-direction.[23]

8.8 Bianchi Identities and the Einstein Tensor

So far, we have discussed algebraic properties of the Riemann tensor. The Riemann tensor also satisfies some differential identities which, in particular in their contracted form, will be of fundamental importance in the following.

The first identity is easy to derive. As a (differential) operator the covariant derivative clearly satisfies the Jacobi identity

$$[\nabla_{[\mu}, [\nabla_\nu, \nabla_{\lambda]}]] = 0 \tag{8.97}$$

(total anti-symmetrisation over all 3 indices). Since the commutator $[\nabla_\nu, \nabla_\lambda]$ is already anti-symmetric in the indices ν, λ, this anti-symmetrisation is equivalent to cyclic permutation of the 3 indices,

$$[\nabla_{[\mu}, [\nabla_\nu, \nabla_{\lambda]}]] = 0 \quad \Leftrightarrow \quad [\nabla_\mu, [\nabla_\nu, \nabla_\lambda]] + \circlearrowleft (\mu, \nu, \lambda) = 0 \ . \tag{8.98}$$

[23]It is far from straightforward, however, to find a formalism which allows one to calculate and derive the distributional Riemann tensor of this space-time - see R. Geroch, J. Traschen, *Strings and other distributional sources in general relativity*, Phys. Rev. D36 (1987) 1017-1031 for a general analysis of the problem and issues arising in this and related contexts, C. Clarke, J. Vickers, J. Wilson, *Generalized functions and distributional curvature of cosmic strings*, Class. Quantum Grav. 13 (1996) 2485-2498 for one approach (based on the Colombeau algebra of distributions), and D. Garfinkle, *Metrics with distributional curvature*, `arXiv:gr-qc/9906053` for a different approach. We will (mostly) stay away from distributional curvatures in these notes.

If you do not believe this identity (valid for any 3 associative linear operators), you can just write out the twelve relevant terms explicitly to see that there is indeed a complete cancellation:

$$\begin{aligned}
[\nabla_{[\mu},[\nabla_\nu,\nabla_{\lambda]}]] \quad &\sim \quad \nabla_\mu\nabla_\nu\nabla_\lambda - \nabla_\mu\nabla_\lambda\nabla_\nu - \nabla_\nu\nabla_\lambda\nabla_\mu + \nabla_\lambda\nabla_\nu\nabla_\mu \\
&+ \quad \nabla_\lambda\nabla_\mu\nabla_\nu - \nabla_\lambda\nabla_\nu\nabla_\mu + \nabla_\nu\nabla_\mu\nabla_\lambda - \nabla_\mu\nabla_\nu\nabla_\lambda \\
&+ \quad \nabla_\nu\nabla_\lambda\nabla_\mu - \nabla_\nu\nabla_\mu\nabla_\lambda - \nabla_\lambda\nabla_\mu\nabla_\nu + \nabla_\mu\nabla_\lambda\nabla_\nu \\
&= \quad 0 \ .
\end{aligned} \tag{8.99}$$

To determine the implications of this identity for the Riemann tensor, we apply it to a vector field V, say. The first term in (8.98) is

$$\begin{aligned}
[\nabla_\mu,[\nabla_\nu,\nabla_\lambda]]V^\rho &= \nabla_\mu(R^\rho_{\ \sigma\nu\lambda}V^\sigma) - [\nabla_\nu,\nabla_\lambda](\nabla_\mu V^\rho) \\
&= (\nabla_\mu R^\rho_{\ \sigma\nu\lambda})V^\sigma + R^\rho_{\ \sigma\nu\lambda}\nabla_\mu V^\sigma - R^\rho_{\ \sigma\nu\lambda}\nabla_\mu V^\sigma + R^\sigma_{\ \mu\nu\lambda}\nabla_\sigma V^\rho \\
&= (\nabla_\mu R^\rho_{\ \sigma\nu\lambda})V^\sigma + R^\sigma_{\ \mu\nu\lambda}\nabla_\sigma V^\rho \ .
\end{aligned} \tag{8.100}$$

Upon taking the cyclic permutations, the sum of the 2nd terms vanishes by the cyclic symmetry of the Riemann tensor, and therefore one finds

$$(\nabla_\mu R^\rho_{\ \sigma\nu\lambda})V^\sigma + \circlearrowleft(\mu,\nu,\lambda) = 0 \ . \tag{8.101}$$

Since this holds for any V, one deduces the *Bianchi identity*

$$\nabla_\mu R^\rho_{\ \sigma\nu\lambda} + \circlearrowleft(\mu,\nu,\lambda) = 0 \quad \Leftrightarrow \quad \nabla_{[\mu}R^\rho_{\ |\sigma|\nu\lambda]} = 0 \tag{8.102}$$

(where $|\sigma|$ indicates that this index is to be excluded from the anti-symmetrisation). Using the symmetry (IV) of the Riemann tensor, this can equivalently be written as

$$\nabla_{[\mu}R_{\nu\lambda]\rho\sigma} = 0 \ . \tag{8.103}$$

We will mainly be interested in a (double) contraction of this identity. To that end we write out (8.102) explicitly as

$$\nabla_\lambda R_{\alpha\beta\mu\nu} + \nabla_\nu R_{\alpha\beta\lambda\mu} + \nabla_\mu R_{\alpha\beta\nu\lambda} = 0 \ . \tag{8.104}$$

By contracting this with $g^{\alpha\mu}$ we obtain

$$\nabla_\lambda R_{\beta\nu} - \nabla_\nu R_{\beta\lambda} + \nabla_\mu R^\mu_{\ \beta\nu\lambda} = 0 \ . \tag{8.105}$$

This is not yet particularly useful. To also turn the last term into a Ricci tensor we contract once more, with $g^{\beta\lambda}$ to obtain the contracted Bianchi identity

$$\nabla_\lambda R^\lambda_{\ \nu} - \nabla_\nu R + \nabla_\mu R^\mu_{\ \nu} = 0 \ , \tag{8.106}$$

or

$$\nabla^\mu(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R) = 0 \ . \tag{8.107}$$

The tensor appearing in this equation is the so-called *Einstein tensor* $G_{\mu\nu}$,

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R \quad . \tag{8.108}$$

It is the unique divergence-free tensor that can be built from the metric and its first and second derivatives (apart from $g_{\mu\nu}$ itself, of course),

$$\nabla^\mu G_{\mu\nu} = 0 \quad , \tag{8.109}$$

and this is why it will play *the* central role in the Einstein equations for the gravitational field.

A minor *caveat* regarding the above statement about the uniqueness of the Einstein tensor is that, as it stands, it is only true in $D = 4$ space-time dimensions. In $D > 4$, there are other tensors with this property, but they are non-linear in 2nd derivatives of the metric. The uniqueness statement continues to be true for $D > 4$ if one adds the requirement that the tensor is linear in 2nd derivatives of the metric. I will briefly come back to this in the discussion of the action principle for general relativity in section 20.1.

8.9 Riemann Normal Coordinates Revisited

In section 3.6 we had introduced Riemann normal coordinates as a special class of inertial coordinate systems, based on geodesics. The main idea was to introduce new coordinates $x^\alpha \to \xi^a$ in such a way that the coordinate lines of the new coordinates ξ^a are geodesics passing through the point p at which one wants to erect this coordinate system.

In particular,

1. we considered the Taylor expansion (3.170)

   $$x^\alpha(\tau) = x_0^\alpha + \tau u_0^\alpha - \tfrac{1}{2}\tau^2 \Gamma^\alpha_{\beta\gamma}(x_0)u_0^\beta u_0^\gamma + \ldots \quad , \tag{8.110}$$

   of a solution to the geodesic equation;

2. this led us to consider the coordinate transformation (3.172)

   $$x^\alpha(\xi) = x_0^\alpha + (\xi - \xi_0)^a e_a^\alpha - \tfrac{1}{2}(\xi - \xi_0)^b(\xi - \xi_0)^c \Gamma^\alpha_{\beta\gamma}(x_0)e_b^\beta e_c^\gamma + \ldots \tag{8.111}$$

   which has the property that the lines

   $$\xi^a(\tau) = \xi_0^a + \tau\lambda^a \tag{8.112}$$

   are geodesics for any constant λ^a;

3. for these geodesics the geodesic equation reduces to

$$\ddot{\xi}^a + \Gamma^a_{\ bc}\dot{\xi}^b\dot{\xi}^c = 0 \quad \Rightarrow \quad \Gamma^a_{\ bc}(\xi^a_0 + \tau\lambda^a)\lambda^b\lambda^c = 0 \ , \tag{8.113}$$

implying at $\tau = 0$

$$\Gamma^a_{\ bc}(\xi_0)\lambda^b\lambda^c = 0 \quad \forall\lambda^a \quad \Rightarrow \Gamma^a_{\ bc}(\xi_0) = 0 \tag{8.114}$$

(and we will look at the implications of the next term in the Taylor expansion of (8.113) below).

Therefore the Taylor expansion of the metric around $\xi = \xi_0$ has the form

$$g_{ab}(\xi) = \eta_{ab} + \tfrac{1}{2}(\xi - \xi_0)^c(\xi - \xi_0)^d g_{ab,\,cd}(\xi_0) + \dots \ , \tag{8.115}$$

and we will now determine the quadratic term in this expansion (and be able to express it in terms of the components of the Riemann tensor $R_{abcd}(\xi_0)$ at the point p in these coordinates). To that end we look at the next term in the Taylor expansion of (8.113). Thus we differentiate (8.113) along the geodesic, i.e. with respect to τ, and evaluate the results at $\tau = 0$ to deduce

$$\partial_d\Gamma^a_{\ bc}(\xi_0)\lambda^d\lambda^b\lambda^c = 0 \quad \forall \ \lambda^b \tag{8.116}$$

or, equivalently

$$\partial_{(d}\Gamma^a_{\ bc)}(\xi_0) = 0 \quad \Leftrightarrow \quad \partial_d\Gamma^a_{\ bc}(\xi_0) + \partial_b\Gamma^a_{\ cd}(\xi_0) + \partial_c\Gamma^a_{\ db}(\xi_0) = 0 \ . \tag{8.117}$$

This condition, and the analogous conditions

$$\partial_{(d\dots e}\Gamma^a_{\ bc)}(\xi_0) = 0 \tag{8.118}$$

arising from the higher-order terms in the Taylor expansion of (8.113) impose constraints on the Christoffel symbols and their derivatives that are satisfied in Riemann normal coordinates (but not in general inertial coordinate systems).

A useful way of reexpressing the condition (8.117) is the following (a certain amount of hindsight or trial-and-error is required for this): because $\Gamma^a_{\ bc}(\xi_0) = 0$, from the definition of the Riemann tensor we have

$$\begin{aligned} R^a_{\ bcd}(\xi_0) + R^a_{\ cbd}(\xi_0) &= \partial_c\Gamma^a_{\ bd}(\xi_0) - \partial_d\Gamma^a_{\ bc}(\xi_0) + \partial_b\Gamma^a_{\ cd}(\xi_0) - \partial_d\Gamma^a_{\ cb}(\xi_0) \\ &= \partial_c\Gamma^a_{\ bd}(\xi_0) + \partial_b\Gamma^a_{\ cd}(\xi_0) - 2\partial_d\Gamma^a_{\ bc}(\xi_0) \ , \end{aligned} \tag{8.119}$$

and using (8.117) this can be written as

$$\partial_d\Gamma^a_{\ bc}(\xi_0) = -\tfrac{1}{3}(R^a_{\ bcd}(\xi_0) + R^a_{\ cbd}(\xi_0)) \ . \tag{8.120}$$

On the other hand, from

$$g_{ab,\,c}(\xi) = \Gamma_{abc}(\xi) + \Gamma_{bac}(\xi) \tag{8.121}$$

we have

$$g_{ab,\,cd}(\xi) = \partial_d \Gamma_{abc}(\xi) + \partial_d \Gamma_{bac}(\xi) \tag{8.122}$$

and at ξ_0 we can use (8.120) and the symmetries of the Riemann tensor to deduce

$$\begin{aligned} g_{ab,\,cd}(\xi_0) &= -\tfrac{1}{3}(R_{abcd} + R_{acbd} + R_{bacd} + R_{bcad})(\xi_0) \\ &= -\tfrac{1}{3}(R_{acbd} + R_{bcad})(\xi_0) = -\tfrac{1}{3}(R_{acbd} + R_{adbc})(\xi_0) \enspace . \end{aligned} \tag{8.123}$$

We have thus found that, to quadratic order in a Taylor expansion of the metric around the origin of a Riemann normal coordinate system, the metric can be written as

$$\begin{aligned} g_{ab}(\xi) &= \eta_{ab} - \tfrac{1}{6}(R_{acbd}(\xi_0) + R_{adbc}(\xi_0))(\xi-\xi_0)^c(\xi-\xi_0)^d + \mathcal{O}(\xi^3) \\ &= \eta_{ab} - \tfrac{1}{3}R_{acbd}(\xi_0)(\xi-\xi_0)^c(\xi-\xi_0)^d + \mathcal{O}(\xi^3) \enspace . \end{aligned} \tag{8.124}$$

If required, higher order terms can be determined analogously with the help of the higher order terms in the Taylor expansion of (8.113), and (with a steady hand) can be expressed in terms of the covariant derivatives of the Riemann tensor at ξ_0.

8.10 Principle of Minimal Coupling Revisited

In sections 4.1 and 6.1 on the principles of general covariance and minimal coupling respectively, I mentioned that these do not necessarily fix the equations uniquely. In other words, there could be more than one generally covariant equation which reduces to a given equation in Minkowski space. Having the curvature tensor at our disposal now, we can construct examples of this kind.

Given some tensorial equation, obtained by the minimal coupling prescription, say, one can always contemplate the possibility to add additional terms to it involving the curvature tensor. Since such terms take the form of higher derivative corrections to the original equation, multiplied by appropriate dimensionful constants, one can usually get away with ignoring such terms when dealing with weak fields and other low-energy phenomena, and under such conditions the minimal coupling rule can usually be trusted. However, such terms are not negligible under extreme conditions involving e.g. very strong or strongly fluctuating gravitational fields.

An example which shows very clearly that the minimal coupling prescription, at least the way we have formulated it, is itself ambiguous is, as already briefly pointed out in section 6.6, provided by Maxwell theory. In that case, we saw that in the covariant Lorenz gauge one has (6.46)

$$\nabla_\mu A^\mu = 0 \quad \Rightarrow \quad \nabla_\mu F^{\mu\nu} = \nabla_\mu(\nabla^\mu A^\nu - \nabla^\nu A^\mu) = \Box A^\nu - [\nabla_\mu, \nabla^\nu]A^\mu \enspace , \tag{8.125}$$

where $\Box A^\nu = \nabla^\mu \nabla_\mu A^\nu$. It thus follows from (8.51) that the Maxwell equations in the covariant Lorenz gauge can be written as (6.47)

$$\nabla_\mu A^\mu = 0 \quad \Rightarrow \quad \nabla_\mu F^{\mu\nu} = -J^\nu \quad \rightarrow \quad \Box A^\nu - R^\nu_{\ \mu} A^\mu = -J^\nu \enspace . \tag{8.126}$$

What this shows is that “minimal coupling” all by itself is not a unique prescription, as we would have obtained (8.126) without the curvature terms by applying the minimal coupling prescription to the special relativity Maxwell equation in the Lorenz gauge, namely just $\Box A_\nu = -J_\nu$.

In the present situation, (6.47) is superior to the equation without the curvature term because

- it follows from a variational principle (involving the minimally coupled counterpart of the Maxwell action)
- and (related to this) because (8.126) implies that the current is covariantly conserved (as we had verified in section 6.6 in an arbitrary gauge), while for the equation without the curvature term covariant current conservation would then be violated by a curvature term (as can easily be verified).

Thus occasionally some such additional criteria can be used to eliminate (or reduce) the ambiguity in the minimal coupling prescription, but this need not always be the case.

As another example, consider the wave equation for a (massless, say) scalar field Φ. In Minkowski space, this is the Klein-Gordon equation which has the obvious curved space analogue (5.56)

$$\Box\Phi = 0 \tag{8.127}$$

obtained by the minimal coupling description. However, one could equally well postulate the equation

$$(\Box + \xi R)\Phi = 0 \ \ , \tag{8.128}$$

where ξ is a (dimensionless) constant and R is the scalar curvature. This equation is generally covariant, and reduces to the ordinary Klein-Gordon equation in Minkowski space, so this is an acceptable curved-space extension of the wave equation for a scalar field. This equation of motion arises (in D space-time dimensions) from the action

$$S_\xi[\phi, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g}d^Dx \ \left(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \xi R\phi^2\right) \ \ , \tag{8.129}$$

exhibiting the non-minimal (yet generally covariant) coupling of the scalar field to gravity via the term $\xi R\phi^2$ (which acts as an x-dependent mass term for the scalar field).

Moreover, ξ is dimensionless, so one cannot argue that on dimensional grounds this ambiguity is irrelevant for weak fields. Indeed, one frequently postulates a specific non-zero value for ξ which makes the wave equation *conformally invariant* (invariant under position-dependent Weyl rescalings of the metric) for massless fields, and this criterion can be imposed to select a particular non-zero value for ξ (e.g. for a 4-dimensional space-time this turns out to be the value $\xi = 1/6$). This will be discussed and explained in section 22.3.

Thus in general such ambiguities are present and are something one has to live with.

B: General Relativity and Geometry

In this second part of the lecture notes I have collected a number of different topics that develop the formalism of tensor calculus in one way or another. This does not mean, however, that one necessarily needs to digest all these topics before continuing with the physical applications of general relativity, and I do not even recommend this.

Stricly speaking none of these topics are essential for understanding some of the more elementary aspects of general relativity to be treated later on, e.g. the discussion of the Einstein equations, the field equations for gravity, in section 19, the discussion of gravitational waves in section 23, or the analysis of geodesics in the Schwarzschild geometry and the corresponding solar system tests of general relativity in section 25.

Some of the topics treated below will reappear frequently in subsequent sections, e.g. Killing vectors (section 9) and their associated conserved quantities (section 10), or the Gauss integral formula derived in section 16.3, and it will be useful to develop at least some nodding acquaintance with these things.

Other topics have been included for a variety of reasons:

- either to illustrate the relation between the Riemann curvature tensor, a central object of interest in general relativity and defined in a somewhat pragmatic and perhaps unintuitive fashion in section 8, and more intuitive and/or geometric concepts of curvature;
- or because they provide an improved understanding of the tensor calculus we have developed so far;
- or because they are required at a later stage to understand, or even formulate, certain somewhat more advanced aspects of general relativity;
- or simply because they are fun or beautiful (or both), and provide an invitation to the wonderful world of differential geometry;
- or (usually) a combination thereof.

9 Lie Derivative, Symmetries and Killing Vectors

9.1 Symmetries of a Metric (Isometries): Preliminary Remarks

Symmetries and their consequences play a fundamental role in physics. In the present context, these are symmetries of the gravitational field or of the space-time metric.

Before trying to figure out how to detect symmetries of a metric, or so-called *isometries*, let us decide what we mean by symmetries of a metric.

For example, we would say that the Minkowski metric has the Poincaré group as a group of symmetries, because the corresponding coordinate transformations leave the metric invariant.

Likewise, we would say that the standard metrics on the two- or three-sphere have rotational symmetries because they are invariant under rotations of the sphere. We can look at this in one of two ways: either as an *active* transformation, in which we rotate the sphere and note that nothing changes, or as a *passive* transformation, in which we do not move the sphere, all the points remain fixed, and we just rotate the coordinate system. So this is tantamount to a relabelling of the points. From the latter (passive) point of view, the symmetry is again understood as an invariance of the metric under a particular family of coordinate transformations.

Thus consider a metric $g_{\mu\nu}(x)$ in a coordinate system $\{x^\mu\}$ and a change of coordinates $x^\mu \to y^\mu(x^\nu)$ (for the purposes of this and the following section it will be convenient not to label the two coordinate systems by different sets of indices). Of course, under such a coordinate transformation we get a new metric $g'_{\mu\nu}$, with (since here we do not distinguish coordinate indices associated to different coordinate systems, we now momentarily put primes on the objects themselves in order to keep track of what we are talking about)

$$g'_{\mu\nu}(y(x)) = \frac{\partial x^\rho}{\partial y^\mu}\frac{\partial x^\lambda}{\partial y^\nu} g_{\rho\lambda}(x) \quad . \tag{9.1}$$

However, so far this by itself has nothing to do with possible symmetries of the metric.

Thinking actively, in order to detect symmetries, we should e.g. compare the geometry, given by the line-element $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$, at two different points x and y related by $y^\mu(x)$. Thus we are led to consider the difference

$$g_{\mu\nu}(y)dy^\mu dy^\nu - g_{\mu\nu}(x)dx^\mu dx^\nu \quad . \tag{9.2}$$

Using the invariance of the line-element under coordinate transformations, i.e. the usual tensorial transformation behaviour of the components of the metric, we see that we can also write this as the difference

$$(g_{\mu\nu}(y) - g'_{\mu\nu}(y))dy^\mu dy^\nu \quad . \tag{9.3}$$

Thus we deduce that what we mean by a symmetry, i.e. invariance of the metric under a coordinate transformation, is the statement

$$g'_{\mu\nu}(y) = g_{\mu\nu}(y) \quad . \tag{9.4}$$

From the passive point of view, in which a coordinate transformation represents a relabelling of the points of the space, this equation compares the new metric at a point P' (with coordinates y^μ) with the old metric at the point P which has the same values of the old coordinates as the point P' has in the new coordinate system, $y^\mu(P') = x^\mu(P)$.

The above equality then states that the new metric at the point P' has the same functional dependence on the new coordinates as the old metric on the old coordinates at the point P. Thus a neighbourhood of P' in the new coordinates looks identical to a neighbourhood of P in the old coordinates, and they can be mapped into each other *isometrically*, i.e. such that all the metric properties, like distances, are preserved. Thus either actively or passively one is led to the above condition.

Note that to detect a continuous symmetry in this way, we only need to consider infinitesimal coordinate transformations. In that case, the above amounts to the statement that metrically the space-time looks the same when one moves infinitesimally in the direction given by the coordinate transformation.

9.2 Lie Derivative for Scalars

We now want to translate the above discussion into a condition for an infinitesimal coordinate transformation

$$x^\mu \to y^\mu(x) = x^\mu + \epsilon V^\mu(x) \tag{9.5}$$

to generate a symmetry of the metric. Here you can and should think of V^μ as a vector field because, even though coordinates themselves of course do not transform like vectors, their infinitesimal variations δx^μ do,

$$z^{\mu'} = z^{\mu'}(x) \to \delta z^{\mu'} = \frac{\partial z^{\mu'}}{\partial x^\mu}\delta x^\mu \tag{9.6}$$

and we think of δx^μ as ϵV^μ.

In fact, we will do something slightly more general than just trying to detect symmetries of the metric. After all, we can also speak of functions or vector fields with symmetries, and this can be extended to arbitrary tensor fields (although that may be harder to visualise). So, for a general tensor field T we will want to compare $T'(y(x))$ with $T(y(x))$ - this is of course equivalent to, and only technically slightly more convenient in the following than, comparing $T'(x)$ with $T(x)$.

As usual, we start the discussion with scalars. In that case, we want to compare $\phi(y(x))$ with $\phi'(y(x)) = \phi(x)$. We find

$$\phi(y(x)) - \phi'(y(x)) = \phi(x + \epsilon V) - \phi(x) = \epsilon V^\mu \partial_\mu \phi + \mathcal{O}(\epsilon^2) \quad . \tag{9.7}$$

We now define the *Lie derivative* of ϕ along the vector field V^μ to be

$$L_V \phi := \lim_{\epsilon \to 0} \frac{\phi(y(x)) - \phi'(y(x))}{\epsilon} \quad . \tag{9.8}$$

Evaluating this, we find

$$L_V \phi = V^\mu \partial_\mu \phi \quad . \tag{9.9}$$

Thus for a scalar, the Lie derivative is just the ordinary directional derivative, and this is as it should be since saying that a function has a certain symmetry amounts to the assertion that its derivative in a particular direction vanishes.

9.3 Lie Derivative for Vector Fields

We now follow the same procedure for a vector field W^μ. We will need the matrix $(\partial y^\mu / \partial x^\nu)$ and its inverse for the above infinitesimal coordinate transformation. We have

$$\frac{\partial y^\mu}{\partial x^\nu} = \delta^\mu_\nu + \epsilon \partial_\nu V^\mu \quad , \tag{9.10}$$

and

$$\frac{\partial x^\mu}{\partial y^\nu} = \delta^\mu_\nu - \epsilon \partial_\nu V^\mu + \mathcal{O}(\epsilon^2) \quad . \tag{9.11}$$

Thus we have

$$\begin{aligned} W'^\mu(y(x)) &= \frac{\partial y^\mu}{\partial x^\nu} W^\nu(x) \\ &= W^\mu(x) + \epsilon W^\nu(x) \partial_\nu V^\mu(x) \quad , \end{aligned} \tag{9.12}$$

and

$$W^\mu(y(x)) = W^\mu(x) + \epsilon V^\nu \partial_\nu W^\mu(x) + \mathcal{O}(\epsilon^2) \quad . \tag{9.13}$$

Hence, defining the Lie derivative $L_V W$ of W by V by

$$L_V W^\mu := \lim_{\epsilon \to 0} \frac{W^\mu(y(x)) - W'^\mu(y(x))}{\epsilon} \quad , \tag{9.14}$$

we find

$$L_V W^\mu = V^\nu \partial_\nu W^\mu - W^\nu \partial_\nu V^\mu \quad . \tag{9.15}$$

There are several important things to note about this expression:

1. The result looks non-covariant, i.e. non-tensorial, but as a difference of two vectors at the same point (recall the limit $\epsilon \to 0$) the result should again be a vector. This is indeed the case. One way to verify this is to check that it indeed transforms as a vector under coordinate transformations. Indeed, by a straightforward calculation one finds that under a coordinate transformation $x^\mu \to y^\alpha$ with Jacobi matrix J^α_μ one has

$$[V,W]^\alpha = J^\alpha_\mu [V,W]^\mu + J^\alpha_{\mu\nu}(V^\mu W^\nu - V^\nu W^\mu) = J^\alpha_\mu [V,W]^\mu \quad , \tag{9.16}$$

because

$$J^\alpha_{\mu\nu} \equiv \partial_\mu J^\alpha_\nu = J^\alpha_{\nu\mu} \tag{9.17}$$

is symmetric.

2. Alternatively, to make the tensorial character of thet Lie derivative manifest, one can rewrite (9.15) in terms of covariant derivatives,

$$\begin{aligned} L_V W^\mu &= V^\nu \nabla_\nu W^\mu - W^\nu \nabla_\nu V^\mu \\ &= \nabla_V W^\mu - \nabla_W V^\mu \quad . \end{aligned} \tag{9.18}$$

This shows that $L_V W^\mu$ is again a vector field. Note, however, that the Lie derivative, in contrast to the covariant derivative, is defined without reference to any metric.

3. There is an alternative, and perhaps more intuitive, derivation of the above expression (9.15) for the Lie derivative of a vector field along a vector field, which makes both its tensorial character and its interpretation manifest (and which also generalises to other tensor fields; in fact we had already applied it to the metric in section 3.2 to deduce (3.34)).

   Namely, let us assume that we are initially in a coordinate system $\{y^{\mu'}\}$ adapted to V in the sense that $V = \partial/\partial y^a$ for some particular a, i.e. $V^{\mu'} = \delta^{\mu'}_a$ (so that we are locally choosing the flow-lines of V as one of the coordinate lines). In this coordinate system we would naturally define the change of a vector field $W^{\mu'}$ along V as the partial derivative of W along y^a,

$$L_V W^{\mu'} := \frac{\partial}{\partial y^a} W^{\mu'} \quad . \tag{9.19}$$

   We now consider an arbitrary coordinate transformation $x^\alpha = x^\alpha(y^{\mu'})$, and require that $L_V W$ transforms as a vector under coordinate transformations. This will then give us the expression for $L_V W$ in an arbitrary coordinate system:

$$\begin{aligned} \frac{\partial}{\partial y^a} W^{\mu'} &= \frac{\partial x^\alpha}{\partial y^a} \frac{\partial}{\partial x^\alpha}\left(\frac{\partial y^{\mu'}}{\partial x^\beta} W^\beta\right) \\ &\overset{!}{=} \frac{\partial y^{\mu'}}{\partial x^\alpha} (L_V W)^\alpha \quad . \end{aligned} \tag{9.20}$$

Disentangling this, using $V^\alpha = \partial x^\alpha / \partial y^a$ and

$$\frac{\partial x^\alpha}{\partial y^a}\frac{\partial^2 y^{\mu'}}{\partial x^\alpha \partial x^\beta} = \frac{\partial x^\alpha}{\partial y^a}\frac{\partial}{\partial x^\beta}\frac{\partial y^{\mu'}}{\partial x^\alpha} = -\frac{\partial V^\alpha}{\partial x^\beta}\frac{\partial y^{\mu'}}{\partial x^\alpha} \quad , \tag{9.21}$$

one recovers the definition (9.15).

4. Note that (9.15) is anti-symmetric in V and W. Hence it defines a commutator $[V, W]$ on the space of vector fields,

$$[V, W]^\mu := L_V W^\mu = -L_W V^\mu \quad . \tag{9.22}$$

This is actually a Lie bracket, i.e. it satisfies the Jacobi identity

$$[V, [W, X]]^\mu + [X, [V, W]]^\mu + [W, [X, V]]^\mu = 0 \quad . \tag{9.23}$$

This can also be rephrased as the statement that the Lie derivative is also a derivation of the Lie bracket, i.e. that one has

$$L_V [W, X]^\mu = [L_V W, X]^\mu + [W, L_V X]^\mu \quad . \tag{9.24}$$

5. I want to reiterate at this point that it is extremely useful to think of vector fields as first order linear differential operators, via $V^\mu \to V = V^\mu \partial_\mu$. In this case, the Lie bracket $[V, W]$ is simply the ordinary commutator of differential operators,

$$\begin{aligned} [V, W] &= [V^\mu \partial_\mu, W^\nu \partial_\nu] \\ &= V^\mu (\partial_\mu W^\nu)\partial_\nu + V^\mu W^\nu \partial_\mu \partial_\nu - W^\nu (\partial_\nu V^\mu)\partial_\mu - W^\nu V^\mu \partial_\nu \partial_\nu \\ &= (V^\nu \partial_\nu W^\mu - W^\nu \partial_\nu V^\mu)\partial_\mu \\ &= (L_V W)^\mu \partial_\mu = [V, W]^\mu \partial_\mu \quad . \end{aligned} \tag{9.25}$$

From this point of view, the Jacobi identity is obvious.

6. From the above it is evident that if one has two vector fields of the form $V_{(k)} = \partial_{y^k}$, they commute as differential operators, i.e. their Lie bracket is zero,

$$V_{(k)} = \partial_{y^k} \quad \Rightarrow \quad [V_{(1)}, V_{(2)}] = 0 \quad . \tag{9.26}$$

Conversely it is also true that locally this is a sufficient condition for the existence of such coordinates,

$$[V_{(1)}, V_{(2)}] = 0 \quad \Leftrightarrow \quad \exists \ \text{(locally)} \ y^k : \quad V_{(k)} = \partial_{y^k} \quad . \tag{9.27}$$

7. For example, if one has a 2-parameter surface $x^\mu = x^\mu(\tau, \sigma)$, which one can think of as a 1-parameter family of curves $x^\mu(\tau)$ labelled by σ, then the tangent vector field $\partial_\tau = \dot{x}^\mu \partial_\mu$ to the family of curves and the *connecting vector field* (or *deviation vector field*) $\partial_\sigma = x'^\nu \partial_\nu$ have vanishing Lie bracket.

Conversely this also provides a good visualisations of what it means for two vector fields to Lie commute, namely that locally they span a 2-dimensional surface and generate a coordinate grid on that surface. We will make use of this in section 12.1 when discussing the so-called geodesic deviation equation.

8. Having equipped the space of vector fields with a Lie algebra structure, in fact with the structure of an *infinite-dimensional Lie algebra*, it is fair to ask 'the Lie algebra of what group?'. Well, we have seen above that we can think of vector fields as infinitesimal generators of coordinate transformations. Hence, formally at least, the Lie algebra of vector fields is the Lie algebra of the group of coordinate transformations (passive point of view) or diffeomorphisms (active point of view).[24] We will briefly come back to this below, in remark 1 of section 9.4.

9. In section 8.2 we had obtained the formula (8.13),

$$([\nabla_X, \nabla_Y] - \nabla_{[X,Y]})V^\lambda = R^\lambda_{\ \sigma\mu\nu} X^\mu Y^\nu V^\sigma \quad . \tag{9.28}$$

   for the relation between the commutator of directional covariant derivatives and the Riemann curvature tensor. There we had used the abbreviation $[X, Y]$ for the vector field $\nabla_X Y^\mu - \nabla_Y X^\mu$. Comparing with (9.18), we see that this is indeed just the Lie bracket $[X, Y]^\mu$. Thus one way of interpreting the Riemann tensor is that the curvature measures the failure of the covariant derivative to provide a representation of the Lie algebra of vector fields.

9.4 Lie Derivative for other Tensor Fields

To extend the definition of the Lie derivative to other tensors, we can proceed in one of two ways. We can either extend the above procedure to other tensor fields by defining

$$L_V T^{\cdots}_{\cdots} := \lim_{\epsilon\to 0} \frac{T^{\cdots}_{\cdots}(y(x)) - T'^{\cdots}_{\cdots}(y(x))}{\epsilon} \quad . \tag{9.29}$$

Or we can extend it to other tensors by proceeding as in the case of the covariant derivative, i.e. by demanding the Leibniz rule. The Lie derivative on an arbitrary tensor is then uniquely determined by its action on scalars and vectors.

In either case, the result can be rewritten in manifestly tensorial form in terms of covariant derivatives. For example, for a covector one finds

$$L_V A_\mu = V^\nu \partial_\nu A_\mu + (\partial_\mu V^\nu) A_\nu = V^\nu \nabla_\nu A_\mu + (\nabla_\mu V^\nu) A_\nu \quad . \tag{9.30}$$

The general result is that the Lie derivative of a (p, q)-tensor T is, like the covariant derivative, the sum of three kinds of terms: the directional covariant derivative of T along V, p terms with a minus sign, involving the covariant derivative of V contracted with each of the upper indices, and q terms with a plus sign, involving the convariant derivative of V contracted with each of the lower indices (note that the plus and minus

[24] See e.g. H. Glöckner, *Fundamental problems in the theory of infinite-dimensional Lie groups*, `arXiv:math/0602078 [math.GR]` for an introduction and a survey of the problems that arise when dealing with or trying to define infinite-dimensional Lie groups.

signs are interchanged with respect to the covariant derivative). Thus, e.g., the Lie derivatives of a (0,2) and a (1,2)-tensor are

$$
\begin{aligned}
L_V T_{\nu\lambda} &= V^\rho \nabla_\rho T_{\nu\lambda} + T_{\rho\lambda}\nabla_\nu V^\rho + T_{\nu\rho}\nabla_\lambda V^\rho \\
L_V T^\mu_{\nu\lambda} &= V^\rho \nabla_\rho T^\mu_{\nu\lambda} - T^\rho_{\nu\lambda}\nabla_\rho V^\mu + T^\mu_{\rho\lambda}\nabla_\nu V^\rho + T^\mu_{\nu\rho}\nabla_\lambda V^\rho \quad .
\end{aligned}
\tag{9.31}
$$

Remarks:

1. While it is not obvious from the somewhat pedestrian definition of the Lie derivative that we have given here, the Lie derivative is an extremely natural operation on tensors. In differential geometry textbooks (and mathematically more sophisticated accounts of general relativity) it is defined as follows:

   (a) Given a vectorfield V, associate to it the 1-parameter family of diffeomorphisms Φ^t_V it generates (with $\Phi^{t=0}_V$ the idenitity), i.e. the flow along the integral curves of this vector field.

   (b) This diffeomorphism induces an action on tensor fields (by pull-back), denoted by $(\Phi^t_V)^*$,
   $$T \mapsto (\Phi^t_V)^* T \quad . \tag{9.32}$$

   (c) Define the Lie derivative to be the infinitesimal generator of this action,
   $$L_V T := \frac{d}{dt}(\Phi^t_V)^* T|_{t=0} \quad . \tag{9.33}$$

   While this definition can be shown to be equivalent to the definition of the Lie derivative given above in terms of coordinates, Taylor expansions etc., this definition is evidently more compact, more illuminating and somewhat more to the point. In particular, it makes the tensorial nature of the Lie derivative manifest. However, in order to arrive at explicit expressions for the Lie derivative of the components of a tensor, one then still needs to perform a calculation equivalent to (9.29).

2. The fact that the Lie derivative provides a representation of the Lie algebra of vector fields by first-order differential operators on the space of (p,q)-tensors is expressed by the identity
   $$[L_V, L_W] = L_{[V,W]} \quad . \tag{9.34}$$
   While it is a bit painful to verify this explicitly on arbitrary tensors, in view of the fact that by the Leibniz rule the Lie derivative of an arbitrary tensor is determined by its action on scalars and vectors, it is actually sufficient to verify (9.34) on scalars and vectors. This is trivial because it is just the statement that the Lie bracket is the commutator of first order differential operators (9.25),
   $$[L_V, L_W]f = [V^\alpha\partial_\alpha, W^\beta\partial_\beta]f = [V,W]^\alpha\partial_\alpha f = L_{[V,W]}f \quad , \tag{9.35}$$

and that this commutator satisfies the Jacobi identity (9.24),

$$\begin{aligned}[L_V, L_W]Z &= L_V[W,Z] - L_W[V,Z] = [V,[W,Z]] - [W,[V,Z]] \\ &= [[V,W],Z] = L_{[V,W]}Z \quad .\end{aligned} \tag{9.36}$$

9.5 Lie Derivative of the Metric and Killing Vectors

The above general formula (9.31) for the Lie derivative of a tensor becomes particularly simple for the metric tensor $g_{\mu\nu}$. The first term is not there (because the metric is covariantly constant), so the Lie derivative is the sum of two terms (with plus signs) involving the covariant derivative of V,

$$L_V g_{\mu\nu} = g_{\lambda\nu}\nabla_\mu V^\lambda + g_{\mu\lambda}\nabla_\nu V^\lambda \quad . \tag{9.37}$$

Lowering the index of V with the metric, this can be written more succinctly as

$$L_V g_{\mu\nu} = \nabla_\mu V_\nu + \nabla_\nu V_\mu \quad . \tag{9.38}$$

The not manifestly covariant avatar of this equation (recall that fundamentally the Lie derivative requires no notion of a covariant differentiation) is

$$L_V g_{\mu\nu} = V^\lambda \partial_\lambda g_{\mu\nu} + \partial_\mu V^\lambda g_{\lambda\nu} + \partial_\nu V^\lambda g_{\mu\lambda} \quad . \tag{9.39}$$

A quick alternative way to arrive at this result is to look directly at the infinitesimal version of the difference

$$g_{\mu\nu}(y)dy^\mu dy^\nu - g_{\mu\nu}(x)dx^\mu dx^\nu \tag{9.40}$$

which was the starting point of our discussion in section 9.1 above. Namely, we consider the infinitesimal coordinate transformation

$$\begin{aligned}\delta_V x^\mu = V^\mu \quad \Rightarrow \quad \delta_V dx^\mu &= dV^\mu = (\partial_\lambda V^\mu)dx^\lambda \\ \delta_V g_{\mu\nu}(x) &= V^\lambda \partial_\lambda g_{\mu\nu}(x) \quad ,\end{aligned} \tag{9.41}$$

and *define* the Lie derivative of the metric by the change this operation δ_V induces in the line element,

$$\delta_V(g_{\mu\nu}dx^\mu dx^\nu) \equiv (L_V g_{\mu\nu})dx^\mu dx^\nu \quad . \tag{9.42}$$

This leads directly to (9.39) and thus to (9.38).

We are now ready to return to our discussion of isometries (symmetries of the metric). Evidently, an infinitesimal coordinate transformation is a symmetry of the metric if $L_V g_{\mu\nu} = 0$. By (9.38) this can be written as (see also (5.69))

$$\begin{aligned}V \text{ generates an isometry} \quad &\Leftrightarrow \quad L_V\, g_{\mu\nu} = 0 \\ &\Leftrightarrow \quad \nabla_\mu V_\nu + \nabla_\nu V_\mu = 0 \quad .\end{aligned} \tag{9.43}$$

Vector fields V satisfying this equation are called *Killing vectors* - not because they kill the metric but after the 19th century mathematician W. Killing.

The alternative non-covariant way (9.39) of writing the Killing equation makes it manifest that only components and derivatives of the metric in the V-direction enter in this condition,

$$\nabla_\mu V_\nu + \nabla_\nu V_\mu = 0 \quad \Leftrightarrow \quad V^\lambda \partial_\lambda g_{\mu\nu} + \partial_\mu V^\lambda g_{\lambda\nu} + \partial_\nu V^\lambda g_{\mu\lambda} = 0 \ . \tag{9.44}$$

This is precisely the condition (3.35) we had encountered first in our discussion of first integrals of motion for the geodesic equation, and which we had already rewritten in terms of covariant derivatives, as in (9.38) above, in (5.68).

Since they are associated with symmetries of space-time, and since symmetries are always of fundamental importance in physics, Killing vectors will play an important role in the following. Our most immediate concern (in section 10, in particular section 10.1) will be with the conserved quantities associated with Killing vectors. Other aspects of Killing vectors and their interplay with the geometry of a space-time will be discussed in sections 13 and 14. For now we just note the following simple facts and examples:

1. Note that by virtue of (9.34) Killing vectors form a Lie algebra, i.e. if V and W are Killing vectors, then also $[V,W]$ is a Killing vector,

   $$L_V g_{\mu\nu} = L_W g_{\mu\nu} = 0 \Rightarrow L_{[V,W]} g_{\mu\nu} = 0 \ . \tag{9.45}$$

   Indeed one has

   $$L_{[V,W]} g_{\mu\nu} = L_V L_W g_{\mu\nu} - L_W L_V g_{\mu\nu} = 0 \ . \tag{9.46}$$

   An explicit proof of this fact will be given later on in section 13.2.

2. The resulting algebra of Killing vectors is the Lie algebra of the isometry group of the metric. For example, the collection of all Killing vectors of the Minkowski metric generates the Lie algebra of the Poincaré group. Indeed, for the Minkowski space-time in inertial (Cartesian) coordinates ξ^a, i.e. with the constant standard metric η_{ab}, the Killing condition simply becomes

   $$\partial_a V_b + \partial_b V_a = 0 \ , \tag{9.47}$$

   which is solved by

   $$V^a = \omega^a_{\ b}\, \xi^b + \epsilon^a \tag{9.48}$$

   where the ϵ^a are constant parameters and the constant matrices $\omega^a_{\ b}$ satisfy $\omega_{ab} = -\omega_{ba}$. These are precisely the infinitesimal Lorentz transformations and translations of the Poincaré algebra, as given e.g. in (1.30).

   Choosing as a basis for the Killing vectors of Minkowski space the vectors

   $$P_a = \partial_a \quad , \quad M_{ab} = \xi_a \partial_b - \xi_b \partial_a \ , \tag{9.49}$$

so that the general Killing vector V^a (9.48) can be expanded as

$$V = V^a \partial_a = \tfrac{1}{2}\omega^{ab} M_{ab} + \epsilon^a P_a \quad , \tag{9.50}$$

the Lie algebra (algebra of Lie brackets) is given by

$$\begin{aligned} [P_a, P_b] &= 0 \\ [M_{ab}, P_c] &= -\eta_{ac} P_b + \eta_{bc} P_a \\ [M_{ab}, M_{cd}] &= \eta_{ad} M_{bc} + \eta_{bc} M_{ad} - \eta_{ac} M_{bd} - \eta_{bd} M_{ac} \quad . \end{aligned} \tag{9.51}$$

This is of course the Lie algebra of the Poincaré group.

3. Another simple example is provided by the two-sphere: as mentioned before, in some obvious sense the standard metric on the two-sphere is rotationally invariant. In particular, with our new terminology we would expect the vector field ∂_ϕ, i.e. the vector field with components $V^\phi = 1, V^\theta = 0$ to be Killing. Let us check this. With the metric $d\theta^2 + \sin^2\theta d\phi^2$, the corresponding covector V_μ, obtained by lowering the indices of the vector field V^μ, are

$$V_\theta = 0 \quad , \qquad V_\phi = \sin^2\theta \quad . \tag{9.52}$$

The Killing condition breaks up into three equations, and we verify

$$\begin{aligned} \nabla_\theta V_\theta &= \partial_\theta V_\theta - \Gamma^\mu_{\ \theta\theta} V_\mu \\ &= -\Gamma^\phi_{\ \theta\theta} \sin^2\theta = 0 \\ \nabla_\theta V_\phi + \nabla_\phi V_\theta &= \partial_\theta V_\phi - \Gamma^\mu_{\ \theta\phi} V_\mu + \partial_\phi V_\theta - \Gamma^\mu_{\ \theta\phi} V_\mu \\ &= 2\sin\theta\cos\theta - 2\cot\theta\sin^2\theta = 0 \\ \nabla_\phi V_\phi &= \partial_\phi V_\phi - \Gamma^\mu_{\ \phi\phi} V_\mu = 0 \quad . \end{aligned} \tag{9.53}$$

Alternatively, using the non-covariant form (9.44) of the Killing equation, one finds, since $V^\phi = 1, V^\theta = 0$ are constant, that the Killing equation reduces to

$$\partial_\phi g_{\mu\nu} = 0 \quad , \tag{9.54}$$

which is obviously satisfied. This is clearly a simpler and more efficient argument.

By solving the Killing equations on S^2, in addition to $\partial_\phi \equiv V_{(3)}$ one finds two other linearly independent Killing vectors $V_{(1)}$ and $V_{(2)}$, namely

$$\begin{aligned} V_{(1)} &= \sin\phi\partial_\theta + \cot\theta\cos\phi\partial_\phi \\ V_{(2)} &= \cos\phi\partial_\theta - \cot\theta\sin\phi\partial_\phi \\ V_{(3)} &= \partial_\phi \quad . \end{aligned} \tag{9.55}$$

Note that $V_{(3)}$ evidently relates these two other Killing vectors by

$$[V_{(3)}, V_{(1)}] = V_{(2)} \quad , \quad [V_{(3)}, V_{(2)}] = -V_{(1)} \quad . \tag{9.56}$$

Since one also has

$$[V_{(1)}, V_{(2)}] = V_{(3)} \quad , \tag{9.57}$$

the $V_{(a)}$ form the Lie algebra

$$[V_{(a)}, V_{(b)}] = \epsilon_{abc} V_{(c)} \quad . \tag{9.58}$$

This is the Lie algebra of infinitesimal rotations, i.e. of the rotation group $SO(3)$, which is the isometry group of the standard metric on S^2.

4. In general, if the components of the metric are all independent of a particular coordinate, say y, then by the above argument $V = \partial_y$ is a Killing vector,

$$\partial_y g_{\mu\nu} = 0 \ \forall \ \mu, \nu \quad \Rightarrow \quad V = \partial_y \ \text{ is a Killing Vector} \tag{9.59}$$

Such a coordinate system, in which one of the coordinate lines agrees with the integral curves of the Killing vector, is said to be adapted to the Killing vector (or isometry) in question. For any given Killing vector V one can always introduce local coordinates such that V takes the form $V = \partial_y$. It suffices to choose as y the parameter along the integral curves of V, using the remaining coordinates to label the individual integral curves.

5. If one has two Killing vector fields $V_{(1)}$ and $V_{(2)}$, then the necessary and sufficient condition that one can introduce local coordinates $(y^1, y^2, \ldots)$ that are adapted to both of them, i.e. such that $V_{(k)} = \partial_{y^k}$ is that they commute as differential operators, i.e. that they have vanishing Lie bracket,

$$[V_{(1)}, V_{(2)}] = 0 \quad \Leftrightarrow \quad \exists \ \text{(locally)} \ y^k : \quad V_{(k)} = \partial_{y^k} \quad , \quad \partial_{y^k} g_{\mu\nu} = 0 \ . \tag{9.60}$$

6. As we did in section 3.2, one can also take the above equations (9.59) as the starting point for what one means by a symmetry of the metric (isometry) and then simply transform it to an arbitrary coordinate system by requiring that it transforms as a $(0, 2)$-tensor. Then one arrives at the Killing condition in the form (9.44).

7. Because by definition the geometry of a space-time does not change along the orbits of a Killing vector, it is intuitively obvious that in particular the norm of a Killing vector V should be constant along (the orbits of) V, and this is indeed easy to prove. Here are two simple proofs of this statement, one using covariant derivatives and the other using Lie derivatives:

   (a) Using covariant derivatives, one calculates

   $$V^\alpha \partial_\alpha (V^\beta V_\beta) = V^\alpha \nabla_\alpha (V^\beta V_\beta) = 2V^\alpha V^\beta \nabla_\alpha V_\beta = 0 \tag{9.61}$$

   by anti-symmetry of $\nabla_\alpha V_\beta$.

(b) Using Lie derivatives, one calculates

$$\begin{aligned} V^\alpha \partial_\alpha (V^\beta V_\beta) &= L_V (g_{\alpha\beta} V^\alpha V^\beta) \\ &= (L_V g_{\alpha\beta}) V^\alpha V^\beta + 2 g_{\alpha\beta} (L_V V^\alpha) V^\beta = 0 \end{aligned} \tag{9.62}$$

because $L_V g_{\alpha\beta} = 0$ (V is a Killing vector) and $L_V V^\alpha = [V, V]^\alpha = 0$ (which is true for any V).

8. An occasionally useful result that provides an interesting relation between geodesics and Killing vectors (different from the one to be discussed below in section 10.1) and that is straightforward to establish, is the fact that a Killing vector field is geodesic if and only if it is of constant length. This follows by contracting the Killing equation with V^μ and writing

$$0 = V^\mu (\nabla_\mu V_\nu + \nabla_\nu V_\mu) = V^\mu \nabla_\mu V_\nu + \tfrac{1}{2} \nabla_\nu (V^\mu V_\mu) \quad . \tag{9.63}$$

Since by definition V^μ is geodesic iff $V^\mu \nabla_\mu V_\nu = 0$ (5.100), the result follows. In particular, this implies that the integral curves of null Killing vector fields are always automatically (affinely parametrised) geodesics.

9. As an aside: a minimal variation of this proof establishes the same result for gradient vector fields $V_\mu = \partial_\mu S$ instead of Killing vector fields, namely that a gradient vector field is geodesic if and only if it is of constant length. Since a gradient vector field satisfies

$$V_\mu = \partial_\mu S \quad \Rightarrow \quad \nabla_\mu V_\nu - \nabla_\nu V_\mu = 0 \tag{9.64}$$

(instead of the Killing vector equation $\nabla_\mu V_\nu + \nabla_\nu V_\mu = 0$), it suffices to change one sign in (9.63),

$$0 = V^\mu (\nabla_\mu V_\nu - \nabla_\nu V_\mu) = V^\mu \nabla_\mu V_\nu - \tfrac{1}{2} \nabla_\nu (V^\mu V_\mu) \quad , \tag{9.65}$$

from which the claimed result follows.

9.6 Lie Derivative for Tensor Densities

It is straightforward to extend the Lie derivative to tensor densities. Given the fact expressed in (4.75) that any tensor density can be written as tensor times a suitble power of the determinant g of the metric, all we need to know is the Lie derivative acting on g. For this we can use the general variational formula (5.77) to deduce

$$L_V \, g = g \, g^{\alpha\beta} L_V g_{\alpha\beta} \quad . \tag{9.66}$$

With the aid of (9.38) this can be simplified to

$$L_V \, g = g \, g^{\alpha\beta} (\nabla_\alpha V_\beta + \nabla_\beta V_\alpha) = 2g \, \nabla_\alpha V^\alpha \quad , \tag{9.67}$$

and for the ubiquitous volume element $\sqrt{g}$ one finds

$$L_V \sqrt{g} = \sqrt{g}\, \nabla_\alpha V^\alpha \quad . \tag{9.68}$$

It follows for example that for a scalar density of weight 1 $\sqrt{g}F$, F a scalar, one has

$$L_V(\sqrt{g}\, F) = \sqrt{g}(V^\alpha \nabla_\alpha F + F \nabla_\alpha V^\alpha) = \sqrt{g}\, \nabla_\alpha (V^\alpha F) \quad . \tag{9.69}$$

Using (5.50), this can also be written as a total derivative

$$L_V(\sqrt{g}\, F) = \partial_\alpha (\sqrt{g}\, V^\alpha F) \quad . \tag{9.70}$$

This identity lies at the heart of the general covariance of actions built from scalars or scalar densities, and we will discuss this aspect in more detail in sections 20.6 and 22.2.

Analogously, the Lie derivative can be extended to tensor densities of any rank and weight.

10 Killing Vectors, Symmetries and Conserved Charges

10.1 Killing Vectors and Conserved Charges

We are used to the fact that symmetries lead to conserved quantities (Noether's theorem). For example, in classical mechanics, the angular momentum of a particle moving in a rotationally symmetric gravitational field is conserved. In the present context, the concept of 'symmetries of a gravitational field' is replaced by 'symmetries of the metric', and we therefore expect conserved charges associated with the presence of Killing vectors. Here are the two most important classes of examples of this phenomenon:

1. Killing Vectors, Geodesics and Conserved Charges

   Let K^μ be a Killing vector field, and $x^\mu(\tau)$ be a geodesic. Then the quantity

   $$Q_K = K_\mu \dot{x}^\mu \tag{10.1}$$

   is constant along the geodesic. Indeed,

   $$\begin{aligned} \frac{d}{d\tau} Q_K = \frac{d}{d\tau}(K_\mu \dot{x}^\mu) &= (D_\tau K_\mu)\dot{x}^\mu + K_\mu D_\tau \dot{x}^\mu \\ &= \nabla_\nu K_\mu \dot{x}^\nu \dot{x}^\mu + 0 \\ &= \tfrac{1}{2}(\nabla_\nu K_\mu + \nabla_\mu K_\nu)\dot{x}^\mu \dot{x}^\nu = 0 \ . \end{aligned} \tag{10.2}$$

   Note that this is precisely the conserved quantity Q_V (3.36) with $V \to K$ deduced from Noether's theorem and the variational principle for geodesics in section 3.2.

2. Conserved Currents from the Energy-Momentum Tensor

   Let K^μ be a Killing vector field, and $T^{\mu\nu}$ the covariantly conserved symmetric energy-momentum tensor, $\nabla_\mu T^{\mu\nu} = 0$. Then the current

   $$J_K^\mu = T^{\mu\nu} K_\nu \tag{10.3}$$

   is covariantly conserved. Indeed,

   $$\begin{aligned} \nabla_\mu J_K^\mu &= (\nabla_\mu T^{\mu\nu})K_\nu + T^{\mu\nu}\nabla_\mu K_\nu \\ &= 0 + \tfrac{1}{2} T^{\mu\nu}(\nabla_\mu K_\nu + \nabla_\nu K_\mu) = 0 \ . \end{aligned} \tag{10.4}$$

   Hence, as we now have a conserved current, we can associate with it a conserved charge in the way discussed above.

   The argument evidently does not rely on $T^{\mu\nu}$ being an energy-momentum tensor but only on the properties $T^{\mu\nu} = T^{\nu\mu}$ and $\nabla_\mu T^{\mu\nu} = 0$.

10.2 Conformal Killing Vectors and Conserved Charges

Another situation of interest occurs when one has a theory invariant under Weyl rescalings and thus a traceless energy-momentum tensor (section 7.7). In that case one can associate conserved currents not only to Killing vectors fields but also to *conformal Killing vectors* C^μ, satisfying

$$L_C g_{\mu\nu} = \nabla_\mu C_\nu + \nabla_\nu C_\mu = 2\omega(x) g_{\mu\nu} \tag{10.5}$$

for some function $\omega(x)$. Such conformal Killing vectors generate coordinate transformations that leave the metric invariant up to an overall (Weyl) rescaling.

If the theory is invariant under such Weyl rescalings, then the energy-momentum tensor is traceless and there should also be a corresponding conserved current. Indeed, we have

2' Let C^μ be a *conformal* Killing vector field, and $T^{\mu\nu}$ a covariantly conserved symmetric and *traceless* energy-momentum tensor, $\nabla_\mu T^{\mu\nu} = T^{\mu\nu} g_{\mu\nu} = 0$. Then

$$J^\mu_C = T^{\mu\nu} C_\nu \tag{10.6}$$

is a covariantly conserved current. Indeed,

$$\begin{aligned} \nabla_\mu J^\mu_C &= (\nabla_\mu T^{\mu\nu}) C_\nu + T^{\mu\nu} \nabla_\mu C_\nu \\ &= 0 + \tfrac{1}{2} T^{\mu\nu} (\nabla_\mu C_\nu + \nabla_\nu C_\mu) = \omega(x) T^{\mu\nu} g_{\mu\nu} = 0 \ . \end{aligned} \tag{10.7}$$

We will look at the example of the conformal Killing vectors of Minkowski space in more detail in section 10.3 below.

There is also a counterpart of statement 1 (conserved charges for geodesics) in the case of conformal Killing vectors, namely for null geodesics (this condition replacing the assumption in statement 2' that the energy-momentum tensor is traceless):

1' Let C^μ be a *conformal* Killing vector field, and let $x^\mu(\tau)$ be a *null* geodesic. Then the quantity

$$Q_C = C_\mu \dot{x}^\mu \tag{10.8}$$

is constant along the geodesic. Indeed, repeating the calculation leading to statement 1, for a null geodesic one has

$$\frac{d}{d\tau} Q_C = \frac{d}{d\tau}(C_\mu \dot{x}^\mu) = \tfrac{1}{2}(\nabla_\nu C_\mu + \nabla_\mu C_\nu)\dot{x}^\mu \dot{x}^\nu = \omega(x) g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = 0 \ . \tag{10.9}$$

We will make use of (10.9) in the discussion of the cosmological redshift in section 34.8.

As an aside, note that if K^μ is a true Killing vector for a metric $\tilde{g}_{\mu\nu}$, say, then it is at least a conformal Killing vector for any conformally rescaled metric

$$g_{\mu\nu} = \mathrm{e}^{2\alpha(x)}\tilde{g}_{\mu\nu} \ . \tag{10.10}$$

Indeed, writing the Killing equation in the non-covariant form (9.44) (in order to avoid having to determine the covariant derivatives or Christoffel symbols of conformally rescaled metrics)

$$K^\lambda\partial_\lambda\tilde{g}_{\mu\nu} + \partial_\mu K^\lambda\tilde{g}_{\lambda\nu} + \partial_\nu K^\lambda\tilde{g}_{\mu\lambda} = 0 \tag{10.11}$$

and expressing this in terms of the metric $g_{\mu\nu}$, one finds

$$K^\lambda\partial_\lambda g_{\mu\nu} + \partial_\mu K^\lambda g_{\lambda\nu} + \partial_\nu K^\lambda g_{\mu\lambda} = 2(K^\lambda\partial_\lambda\alpha)g_{\mu\nu} \ . \tag{10.12}$$

This is precisely the conformal Killing vector equation with

$$\omega(x) = K^\lambda\partial_\lambda\alpha(x) \ . \tag{10.13}$$

Alternatively, and more simply, we could have just used the Lie derivative directly to conclude that

$$L_K\tilde{g}_{\mu\nu} = 0 \quad \Rightarrow \quad L_K g_{\mu\nu} = 2(L_K\alpha)g_{\mu\nu} \ . \tag{10.14}$$

Either way we see thus K^μ will be a true Killing vector field for the rescaled metric if the conformal factor $\alpha(x)$ is constant along the orbits (integral curves) of K^μ, and will otherwise be a conformal Killing vector field. Conformal Killing vector fields that do not arise from true Killing vector fields in this way are called *essential*. In the Riemannian case it is known that (under some technical assumptions) metrics admitting essential conformal vector fields are conformal to the standard metric on the sphere or the Euclidean space. In the pseudo-Riemannian (Lorentzian signature) case the situation turns out to be quite different (with an interesting connection with the plane wave metrics that are the subject of section 43).[25]

More generally, by the same argument as above we can conclude that if C is a conformal Killing vector of the metric $\tilde{g}_{\mu\nu}$ it will (at the very least) be a conformal Killing vector of any conformally rescaled metric,

$$L_C\tilde{g}_{\mu\nu} = 2\omega\tilde{g}_{\mu\nu} \quad \Rightarrow \quad L_C g_{\mu\nu} = 2(L_C\alpha + \omega)g_{\mu\nu} \ . \tag{10.15}$$

[25] See e.g. F. Belgun, A. Moroianu, L. Ornea, *Essential points of conformal vector fields*, `arXiv:1002.0482 [math.DG]` and references therein, as well as W. Kühnel, H. Rademacher, *Essential conformal field in pseudo-Riemannian geometry*, `http://www.math.uni-leipzig.de/~rademacher/Paper/j-math-pures.pdf`, *Conformal transformations of pseudo-Riemannian manifolds*, `http://www.math.uni-leipzig.de/~rademacher/esi.pdf`.

As an example, let us consider 4-dimensional Minkowski space. In that case there are 5 conformal Killing vectors (in addition to the 10 true Killing vectors (9.48) generating Poincaré transformations).

- One is the generator

$$D = \xi^a \partial_a : \quad \partial_a D_b + \partial_b D_a = 2\eta_{ab} \tag{10.16}$$

  of dilatations,

$$\xi^a \to \mathrm{e}^{\lambda} \xi^a \quad \Rightarrow \quad ds^2 = \eta_{ab} d\xi^a d\xi^b \to \mathrm{e}^{2\lambda} ds^2 \quad . \tag{10.17}$$

  In this case $\omega(x) = 1$ is constant, and such a conformal symmetry is called a *homothety* (see also section 10.4 below). Provided that one has a symmetric traceless conserved energy-momentum tensor, one has a corresponding conserved current

$$J^a_D = T^a_b D^b = T^a_b \xi^b \quad . \tag{10.18}$$

- The other 4 conformal Killing vectors are

$$C^{(m)} = (2\xi^m \xi^a - \eta^{ma}\xi^2)\partial_a \tag{10.19}$$

  where $\xi^2 = \eta_{ab}\xi^a\xi^b$. Indeed, is is straightforward to see that these vector fields satisfy

$$\partial_a C^{(m)}_b + \partial_b C^{(m)}_a = 4\xi^m \eta_{ab} \quad , \tag{10.20}$$

  so that in this case there is a nontrivial conformal factor $\omega^{(m)}(\xi) = 2\xi^m$.

  The $C^{(m)}$ generate what are known as *special conformal transformations*,

$$\xi^a \to \bar{\xi}^a = \frac{\xi^a + c^a \xi^2}{1 + 2c.\xi + c^2\xi^2} \tag{10.21}$$

  with the evident short-hand notation $c.\xi = \eta_{ab}c^a\xi^b$ and $c^2 = \eta_{ab}c^a c^b$. Simple algebra shows that this transformation can be written in the form

$$\frac{\bar{\xi}^a}{\bar{\xi}^2} = \frac{\xi^a}{\xi^2} + c^a \quad . \tag{10.22}$$

  Thus a special conformal transformation can be understood as an inversion $\xi^a \to \xi^a/\xi^2$, followed by a translation (with respect to the "point at infinity"), and another inversion.

  Provided that one has a symmetric traceless conserved energy-momentum tensor, the associated conserved currents are

$$J^a_{(m)} \equiv J^a_{C^{(m)}} = T^a_b C^{(m)b} \quad . \tag{10.23}$$

The dilatation and the special conformal transformation enlarge the Poincaré algebra (9.51) of translations and Lorentz transformations to the *conformal algebra*. Adding the generators D and $C_b \equiv C^{(b)}$ to the generators P_a and M_{ab} of the Poincaré algebra, one finds the extended algebra

$$
\begin{aligned}
[P_a, P_b] &= 0 \\
[M_{ab}, P_c] &= -\eta_{ac} P_b + \eta_{bc} P_a \\
[M_{ab}, M_{cd}] &= \eta_{ad} M_{bc} + \eta_{bc} M_{ad} - \eta_{ac} M_{bd} - \eta_{bd} M_{ac} \\
[D, P_a] &= -P_a \\
[M_{ab}, D] &= 0 \\
[P_a, C_b] &= 2(\eta_{ab} D - M_{ab}) \\
[M_{ab}, C_c] &= -\eta_{ac} C_b + \eta_{bc} C_a \\
[D, C_a] &= C_a \\
[C_a, C_b] &= 0 \quad .
\end{aligned} \tag{10.24}
$$

Here

- the first three relations just define the Poincaré algebra;
- the fourth expresses the obvious fact that $P_a = \partial_{\xi^a}$ is homogeneous of degree (-1) under the dilatation generated by D;
- the fifth says that D is a scalar under Lorentz transformations;
- the seventh just expresses the fact that C_a is a Lorentz vector;
- the eighth says that C_a is homogeneous of degree (+1) under the dilatation generated by D.
- the last relation says that special conformal transformations generate an Abelian algebra (corresponding to the fact that they generate inverted translations).

Thus the only relation that is not a priori obvious is the sixth, $[P_a, C_b] = 2(\eta_{ab} D - M_{ab})$, but this follows simply from

$$
[P_a, C_b]^c = \partial_a (2\xi_b \xi^c - \delta_b^c \xi^2) = 2\eta_{ab}\xi^c - 2(\xi_a \delta_b^c - \xi_b \delta_a^c) \quad . \tag{10.25}
$$

It is perhaps also not obvious at first sight that this conformal Lie algebra is isomorphic to the Lie algebra of $SO(2,4)$, or $SO(2,D)$ in D space-time dimensions. This is the group of rotations in the $(D+2)$-dimensional pseudo-Euclidean space $\mathbb{R}^{2,D}$ preserving the metric η_{AB} with signature $(-+\ldots+-)$, i.e. the indices have the range $A = 0, 1, \ldots, D+1$, and $\eta_{DD} = -\eta_{(D+1)(D+1)} = +1$. Its Lie algebra is just the obvious counterpart of the D-dimensional Lorentz Lie algebra (9.51), namely

$$
[M_{AB}, M_{CD}] = \eta_{AD} M_{BC} + \eta_{BC} M_{AD} - \eta_{AC} M_{BD} - \eta_{BD} M_{AC} \quad , \tag{10.26}
$$

Concretely, with z^A Cartesian coordinates on $\mathbb{R}^{2,D}$, this Lie algebra can be realised as the algebra of rotational Killing vectors of the metric η_{AB}, given by

$$M_{AB} = \eta_{AC}z^C\partial_B - \eta_{BC}z^C\partial_A \equiv z_A\partial_B - z_B\partial_A = -M_{BA} \quad . \tag{10.27}$$

Returning to the conformal algebra, it is now easy to see that with the identification

$$P_a = M_{aD} + M_{a(D+1)} \quad , \quad C_a = M_{aD} - M_{a(D+1)} \quad , \quad D = M_{D(D+1)} \tag{10.28}$$

the Lie algebra relations (10.24) and (10.26) are mapped precisely into each other.

Thus, when one has a *conserved, symmetric, traceless* energy-momentum tensor, one can construct conserved currents for the entire conformal group and thus has a (at least classically) conformally invariant field theory (or *conformal field theory* for short).

As we have seen in section 7.7, when the matter action is invariant under Weyl rescalings of the metric alone, the covariant energy-momentum tensor is conserved, symmetric and traceless, and thus the specialisation of the theory to Minkowski space should define a conformal field theory.

There is an interesting twist to this story when one also needs to transform the matter fields (and modify the action by non-minimal couplings to the gravitational field) which will be discussed in section 22.3.

10.4 Homotheties and Conserved Charges

Finally, let us consider the special case that the conformal factor $\omega(x)$ in (10.5) is constant, $\omega(x) = \omega_0$,

$$\nabla_\mu C_\nu + \nabla_\nu C_\mu = 2\omega_0 g_{\mu\nu} \tag{10.29}$$

In that case, the transformation generated by the conformal Killing vector is called a *homothety.*

An example of a homothetic Killing vector is the generator of dilatations (10.16)

$$D = \xi^a\partial_a \tag{10.30}$$

in Minkowski space. Other examples of space-times admitting homotheties are for example the exact gravitational plane waves (to be discussed in detail much later, in section 43), for which the metrics take the form (43.19)

$$ds^2 = 2dudv + A_{ab}(u)x^a x^b du^2 + d\vec{x}^2 \quad , \tag{10.31}$$

with $A_{ab}(u)$ an arbitrary function of u. These metrics have the homothety

$$(u, v, x^a) \to (u, \lambda^2 v, \lambda x^a) \quad \Rightarrow \quad ds^2 \to \lambda^2 ds^2 \tag{10.32}$$

for any choice of plane wave "profile" $A_{ab}(u)$, and this homothety is generated by

$$C = 2v\partial_v + x^a \partial_{x^a} \quad . \tag{10.33}$$

Whenever one has such a homothety, there is an explicitly τ-dependent conserved quantity even for non-null geodesics:

1" Let C^μ be a *homothetic* Killing vector field, with factor ω_0, and let $x^\mu(\tau)$ be a geodesic. Then the quantity

$$Q_C = C_\mu \dot{x}^\mu - \tau\omega_0 g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \tag{10.34}$$

is constant along the geodesic. Indeed, repeating the calculation leading to statement 1', and using the fact that $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ is constant, one finds

$$\frac{d}{d\tau}Q_C = \omega_0 g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu - \omega_0 g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = 0 \quad . \tag{10.35}$$

<u>Remarks:</u>

1. Note that for a null geodesic (10.34) reduces to the conserved charge $C_\mu\dot{x}^\mu$ (10.8) in 1' above (which does not explicitly depend on τ).

2. The existence of this constant of motion can also be understood from the Noether theorem (applied now to transformations of the "fields" $x^\alpha(\tau)$ and the "coordinate" τ). Indeed, when one has a homothety, one has

$$g_{\mu\nu}dx^\mu dx^\nu \to \lambda^2 g_{\mu\nu}dx^\mu dx^\nu \quad , \tag{10.36}$$

so that the action is invariant when one also scales $\tau \to \lambda^2\tau$,

$$\tau \to \lambda^2\tau \quad \Rightarrow \quad d\tau\; g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \to d\tau\; g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \quad , \tag{10.37}$$

and (10.34) is the corresponding Noether charge.

3. Typically such explicitly τ-dependent constants of motion are somewhat trivial or tautological, in the sense that they can be written in terms of "trivial" constants of integration like initial positions and velocities.

4. For example, for the homothety of Minkowski space-time generated by D (10.30), the conserved charge is explicitly

$$Q_D = \eta_{ab}\xi^a\dot{\xi}^b - \tau\eta_{ab}\dot{\xi}^a\dot{\xi}^b \quad . \tag{10.38}$$

For a free particle with $\ddot{\xi}^a = 0$ and $\eta_{ab}\dot{\xi}^a\dot{\xi}^b = -1$ this is rather obviously conserved, since then

$$\frac{d}{d\tau}Q_D = -1 + 1 = 0 \quad . \tag{10.39}$$

Parametrising $\xi(\tau)$ as

$$\xi^a(\tau) = \xi_0^a + (p^a/m)\tau \tag{10.40}$$

with the constant momenta satisfying the mass shell condition

$$p^a p_a = -m^2 \ , \tag{10.41}$$

one finds that

$$Q_D = \eta_{ab}\xi_0^a p^b/m \tag{10.42}$$

which is about as manifestly constant as it gets.

5. Nevertheless, in other circumstances this explicitly τ-dependent constant of motion (allowing one to integrate a τ-independent 2nd order equation to a τ-dependent 1st order equation) can be useful.

10.5 Conserved Charges from Killing Tensors and Killing-Yano Tensors

When a metric possesses sufficiently many symmetries (Killing vectors), the geodesic equations (or the associated Hamilton-Jacobi equation) or, say, the Klein-Gordon equation or some other field equation in that background are separable and can hence be reduced to quadratures of ordinary differential equations. It is not uncommon, however, in particular in the context of black hole physics, to encounter space-times in which these equations can be separated even though there appear not to be enough isometries (symmetries of the metric) to explain this. In many cases, this phenomenon can be explained via (or deduced from) the existence of additional (hidden) symmetries of the problem, associated not to Killing vectors but to certain higher-rank generalisations thereof. Most prominent among them are (totally symmetric) *Killing tensors* (occasionally also called *Killing-Stäckel tensors*), and (totally anti-symmetric) *Killing-Yano tensors*.

To set the stage, recall from above that a Killing vector satisfies

$$\nabla_{(\alpha}K_{\beta)} = 0 \quad \Leftrightarrow \quad \nabla_\alpha K_\beta = \nabla_{[\alpha}K_{\beta]} \tag{10.43}$$

and that using the geodesic equation $\dot{x}^\alpha\nabla_\alpha\dot{x}^\beta = 0$ this leads to a first integral $Q_K = K_\beta\dot{x}^\beta$ of the geodesic equations of motion via the simple chain of manipulations

$$\frac{d}{d\tau}(K_\beta\dot{x}^\beta) = \dot{x}^\alpha\nabla_\alpha(K_\beta\dot{x}^\beta) = \dot{x}^\alpha\dot{x}^\beta\nabla_\alpha K_\beta = 0 \tag{10.44}$$

by symmetry of $\dot{x}^\alpha\dot{x}^\beta$ and anti-symmetry of $\nabla_\alpha K_\beta$.

This has the following two immediate (and, as it turns out, actually useful in practice) generalisations:

1. Killing Tensors (or Killing-Stäckel Tensors)

   Let $K_{\beta_1 \dots \beta_n}$ be totally symmetric rank-n tensor satisfying the Killing tensor equation

   $$\nabla_{(\alpha} K_{\beta_1 \dots \beta_n)} = 0 \ . \tag{10.45}$$

   This is evidently one possible generalisation of the Killing vector equation (10.43) to higher rank tensors (generalising the first formulation in (10.43)). Then

   $$Q_K = K_{\beta_1 \dots \beta_n} \dot{x}^{\beta_1} \dots \dot{x}^{\beta_n} \tag{10.46}$$

   is constant along the geodesic. Indeed,

   $$\frac{d}{d\tau} Q_K = (\nabla_\alpha K_{\beta_1 \dots \beta_n}) \dot{x}^\alpha \dot{x}^{\beta_1} \dots \dot{x}^{\beta_n} = 0 \tag{10.47}$$

   because evidently $\dot{x}^\alpha \dot{x}^{\beta_1} \dots \dot{x}^{\beta_n}$ is totally symmetric.

2. Killing-Yano Tensors

   Let $Y_{\beta_1 \dots \beta_n}$ be totally anti-symmetric rank-n tensor satisfying the Killing-Yano equation

   $$\nabla_{(\alpha} Y_{\beta_1) \dots \beta_n} = 0 \quad \Leftrightarrow \quad \nabla_\alpha Y_{\beta_1 \dots \beta_n} = \nabla_{[\alpha} Y_{\beta_1 \dots \beta_n]} \tag{10.48}$$

   This is evidently another possible generalisation of the Killing vector equation (10.43) to higher rank tensors. Then the tensorial charges

   $$Z_{\beta_1 \dots \beta_{n-1}} = \dot{x}^\beta Y_{\beta \beta_1 \dots \beta_{n-1}} \tag{10.49}$$

   are constant (parallel transported) along the geodesic. Indeed,

   $$\frac{d}{d\tau} Z_{\beta_1 \dots \beta_{n-1}} = \dot{x}^\alpha \dot{x}^\beta \nabla_\alpha Y_{\beta \beta_1 \dots \beta_{n-1}} = 0 \tag{10.50}$$

   because evidently $\dot{x}^\alpha \dot{x}^\beta$ is symmetric while by definition of a Killing-Yano tensor $\nabla_\alpha Y_{\beta \beta_1 \dots \beta_{n-1}}$ is totally anti-symmetric.

<u>Remarks:</u>

1. Trivial examples of Killing tensors are the metric $g_{\alpha\beta}$ (whose associated conserved quantity $g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta$ we already know), and products of Killing vectors $K_\alpha \dots K_\beta$ which do not yield any new independent constants of motion beyond those provided by the Killing vectors. New constants of motion are associated with Killing tensors that cannot be constructed from the metric and the Killing vectors alone.

   Trivial Killing-Yano tensors are Killing vectors K_α and the Levi-Civita tensor (in four dimensions $\epsilon_{\alpha\beta\gamma\delta}$).

2. There are interesting relations between Killing-Yano tensors and Killing tensors. For example, it is not difficult to check that if $Y_{\alpha\beta}$ is a rank-2 Killing-Yano tensor, then its square

$$K_{\alpha\beta} = Y_{\alpha\gamma} Y^{\gamma}_{\beta} \tag{10.51}$$

(which is symmetric) is a rank-2 Killing tensor (and squares of trivial Killing-Yano tensors give rise to trivial Killing tensors, as in $K_\alpha \to K_\alpha K_\beta$). Indeed, the totally symmetrised covariant derivative of this $K_{\alpha\beta}$ can be expressed in terms of partially symmetrised covariant derivatives of $Y_{\alpha\beta}$, but by definition of a Killing-Yano tensor its covariant derivatives are totally anti-symmetric, and hence

$$\nabla_\gamma Y_{\alpha\beta} = \nabla_{[\gamma} Y_{\alpha\beta]} \quad \Rightarrow \quad \nabla_{(\gamma} K_{\alpha\beta)} = 0 \ . \tag{10.52}$$

3. There are conformal generalisations of these Killing(-Yano) tensor equations, analogous to the conformal Killing equations (10.5),

$$\nabla_{(\alpha} C_{\beta)} = \omega(x) g_{\alpha\beta} \ , \tag{10.53}$$

and just as the latter these turn out to be useful for massless particles or fields. For example, a rank 2 conformal Killing tensor satisifies an equation of the form

$$\nabla_{(\alpha} C_{\beta\gamma)} = g_{(\alpha\beta} V_{\gamma)} \tag{10.54}$$

for some (co-)vector field V. Repeating the calculation (10.47) in the case at hand for the quantity $Q_C = C_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta$, one finds

$$\frac{d}{d\tau} Q_C = \nabla_{(\alpha} C_{\beta\gamma)} \dot{x}^\alpha \dot{x}^\beta \dot{x}^\gamma = g_{\alpha\beta} V_\gamma \dot{x}^\alpha \dot{x}^\beta \dot{x}^\gamma \tag{10.55}$$

which evidentliy vanishes for null geodesics ($g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta = 0$).

4. Historically, the discovery of (conformal) Killing and Killing-Yano tensors for the *Kerr metric*, the metric describing a rotating black hole (see section 30.1) and their relation to the separability of the geodesic and field equations in the Kerr background played a decisive role in the development of the subject.[26]

[26] For more information about and examples and applications of Killing(-Yano) tensors, see e.g. section 35.3 of H. Stephani, D. Kramer, M. MacCallum, C. Hoenslaers, E. Herlt, *Exact Solutions to Einstein's Field Equations - Second Edition* or the articles O. Santillan, *Killing-Yano tensors and some applications*, `arXiv:1108.0149 [hep-th]`, F. Larsen, C. Keeler, *Separability of Black Holes in String Theory*, `arXiv:1207.5928 [hep-th]` and the references therein.

11 Curvature II: Geometry and Curvature

In this section, we will first discuss two properties of the Riemann curvature tensor that illustrate its geometric significance and thus, *a posteriori*, justify equating the commutator of covariant derivatives with the intuitive concept of curvature. These properties are

- the *path-dependence of parallel transport* in the presence of curvature,
- the fact that the space-time metric is equivalent to the (in an obvious sense flat) Minkowski metric if and only if the Riemann curvature tensor vanishes.

We then briefly discuss some other general aspects of the relation between geometry and curvature (while the interplay between geodesics and curvature and Killing vectors and curvature will be discussed in sections 12 and 13 respectively).

11.1 Intrinsic Geometry, Curvature and Parallel Transport

The Riemann curvature tensor and its relatives, introduced above, measure the *intrinsic geometry* and curvature of a space or space-time. This means that they can be calculated by making experiments and measurements in the space itself. Such experiments might involve things like checking if the interior angles of a triangle add up to π or not.

This intrinsic geometry and curvature described above should be contrasted with the *extrinsic geometry* which depends on how the space may be embedded in some larger space. As we have no intention of embedding space-time into something higher dimensional, we will mainly be concerned with intrinsic geometry in the following. However, if you would for example be interested in the properties of spacelike hypersurfaces in space-time, then aspects of both intrinsic and extrinsic geometry of that hypersurface would be relevant. See section 18 for some further comments on this.

Let us return to intrinsic geometry. An even better method, the subject of this section, to determine the curvature is to check the properties of parallel transport. The tell-tale sign (or smoking gun) of the presence of curvature is the fact that parallel transport is path dependent, i.e. that parallel transporting a vector V from a point A to a point B along two different paths will in general produce two different vectors at B. Another way of saying this is that parallel transporting a vector around a closed loop at A will in general produce a new vector at A which differs from the initial vector.

This is easy to see in the case of the two-sphere, for which we also worked out explicitly the parallel transport in section 5.9 (see Figure 8). Since all the great circles on a two-sphere are geodesics, in particular the segments N-C, N-E, and E-C in the figure, we know that in order to parallel transport a vector along such a line we just need to

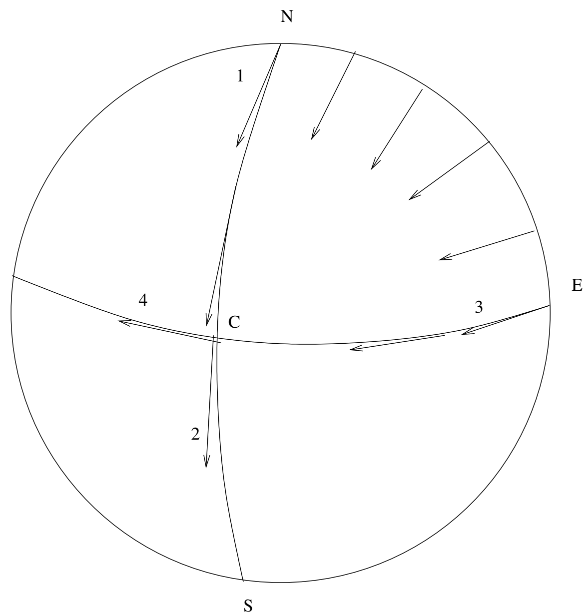


Figure 8: Figure illustrating the path dependence of parallel transport on a curved space: Vector 1 at N can be parallel transported along the geodesic N-S to C, giving rise to Vector 2. Alternatively, it can first be transported along the geodesic N-E (Vector 3) and then along E-C to give the Vector 4. Clearly these two are different. The angle between them reflects the curvature of the two-sphere.

make sure that its length and the angle between the vector and the geodesic line are constant. Thus imagine a vector 1 at the north pole N, pointing downwards along the line N-C-S. First parallel transport this along N-C to the point C. There we will obtain the vector 2, pointing downwards along C-S. Alternatively imagine parallel transporting the vector 1 first to the point E. Since the vector has to remain at a constant (right) angle to the line N-E, at the point E parallel transport will produce the vector 3 pointing westwards along E-C. Now parallel transporting this vector along E-C to C will produce the vector 4 at C. This vector clearly differs from the vector 2 that was obtained by parallel transporting along N-C instead of N-E-C.

To illustrate the claim about closed loops above, imagine parallel transporting vector 1 along the closed loop N-E-C-N from N to N. In order to complete this loop, we still have to parallel transport vector 4 back up to N. Clearly this will give a vector, not indicated in the figure, different from (and pointing roughly at a right angle to) the vector 1 we started off with.

The precise statement regarding the relation between the path dependence of parallel transport and the presence of curvature is the following. If one parallel transports a covector V_μ (I use a covector instead of a vector only to save myself a few minus signs here and there) along a closed infinitesimal loop $x^\mu(\tau)$ with, say, $x(\tau_0) = x(\tau_1) = x_0$, then one has

$$V_\mu(\tau_1) - V_\mu(\tau_0) = \tfrac{1}{2}(\oint x^\rho dx^\nu) R^\sigma_{\ \mu\rho\nu}(x_0) V_\sigma(\tau_0) \ . \tag{11.1}$$

Thus an arbitrary vector V^μ will not change under parallel transport around an arbitrary small loop at x_0 only if the curvature tensor at x_0 is zero. This can of course be extended to finite loops, but the important point is that in order to detect curvature at a given point one only requires parallel transport along infinitesimal loops.

Before turning to a proof of this result, I just want to note that intuitively it can be understood directly from the definition of the curvature tensor (8.2). Imagine that the infinitesimal loop is actually a tiny parallelogram made up of the coordinate lines x^1 and x^2. Parallel transport along x^1 is governed by the equation $\nabla_1 V^\mu = 0$, that along x^2 by $\nabla_2 V^\mu = 0$. The fact that parallel transporting first along x^1 and then along x^2 can be different from doing it the other way around is precisely the statement that ∇_1 and ∇_2 do not commute, i.e. that some of the components $R_{\mu\nu 12}$ of the curvature tensor are non-zero.

To establish (11.1) we first reformulate the condition of parallel transport,

$$D_\tau V_\mu = 0 \quad \Leftrightarrow \quad \frac{d}{d\tau} V_\mu = \Gamma^\lambda_{\ \mu\nu} \dot{x}^\nu V_\lambda \tag{11.2}$$

with the initial condition at $\tau = \tau_0$ as the integral equation

$$V_\mu(\tau) = V_\mu(\tau_0) + \int_{\tau_0}^{\tau} d\tau' \, \Gamma^\lambda_{\ \mu\nu}(x(\tau')) \dot{x}^\nu(\tau') V_\lambda(\tau') \ . \tag{11.3}$$

As usual, such an equation can be 'solved' by iteration (leading to a time-ordered exponential). Keeping only the first two non-trivial terms in the iteration, one has

$$\begin{aligned} V_\mu(\tau) \;&=\; V_\mu(\tau_0) + \int_{\tau_0}^{\tau} d\tau' \, \Gamma^\lambda_{\ \mu\nu}(x(\tau')) \dot{x}^\nu(\tau') V_\lambda(\tau_0) \\ &\quad + \int_{\tau_0}^{\tau} d\tau' \int_{\tau_0}^{\tau'} d\tau'' \, \Gamma^\lambda_{\ \mu\nu}(x(\tau')) \dot{x}^\nu(\tau') \Gamma^\sigma_{\ \lambda\rho}(x(\tau'')) \dot{x}^\rho(\tau'') V_\sigma(\tau_0) \\ &\quad + \ldots \end{aligned} \tag{11.4}$$

For sufficiently small (infinitesimal) loops, we can expand the Christoffel symbols as

$$\Gamma^\lambda_{\ \mu\nu}(x(\tau)) = \Gamma^\lambda_{\ \mu\nu}(x_0) + (x(\tau) - x_0)^\rho (\partial_\rho \Gamma^\lambda_{\ \mu\nu})(x_0) + \ldots \tag{11.5}$$

The linear term in the expansion of $V_\mu(\tau)$ arises from the zero'th order contribution $\Gamma^\lambda_{\ \mu\nu}(x_0)$ in the first order (single integral) term in (11.4),

$$[V_\mu(\tau_1) - V_\mu(\tau_0)]^{(1)} = \Gamma^\lambda_{\ \mu\nu}(x_0) V_\lambda(\tau_0) (\int_{\tau_0}^{\tau_1} d\tau' \, \dot{x}^\nu(\tau')) \ . \tag{11.6}$$

Now the important observation is that, for a closed loop, the integral in brackets is zero,

$$\int_{\tau_0}^{\tau_1} d\tau' \dot{x}^\nu(\tau') = x^\nu(\tau_1) - x^\nu(\tau_0) = 0 \ . \tag{11.7}$$

Thus the change in $V_\mu(\tau)$, when transported along a small loop, is at least of second order. Such second order terms arise in two different ways, from the first order term in the expansion of $\Gamma^\lambda_{\mu\nu}(x)$ in the first order term in (11.4), and from the zero'th order terms $\Gamma^\lambda_{\mu\nu}(x_0)$ in the quadratic (double integral) term in (11.4),

$$\begin{aligned}[V_\mu(\tau_1) - V_\mu(\tau_0)]^{(2)} &= (\partial_\rho \Gamma^\lambda_{\mu\nu})(x_0) V_\lambda(\tau_0)(\int_{\tau_0}^{\tau_1} d\tau'\ (x(\tau') - x_0)^\rho \dot{x}^\nu(\tau')) \\ &+ (\Gamma^\lambda_{\mu\nu}\Gamma^\sigma_{\lambda\rho})(x_0) V_\sigma(\tau_0) \int_{\tau_0}^{\tau_1} d\tau' \int_{\tau_0}^{\tau'} d\tau''\ \dot{x}^\nu(\tau')\dot{x}^\rho(\tau'') \end{aligned} \quad (11.8)$$

The τ''-integral can be performed explicitly,

$$\int_{\tau_0}^{\tau_1} d\tau' \int_{\tau_0}^{\tau'} d\tau''\ \dot{x}^\nu(\tau')\dot{x}^\rho(\tau'') = \int_{\tau_0}^{\tau_1} d\tau'\ \dot{x}^\nu(\tau')(x(\tau') - x_0)^\rho = \int_{\tau_0}^{\tau_1} d\tau'\ \dot{x}^\nu(\tau')x^\rho(\tau') \quad (11.9)$$

and therefore we find

$$V_\mu(\tau_1) - V_\mu(\tau_0) \approx (\partial_\rho \Gamma^\sigma_{\mu\nu} + \Gamma^\lambda_{\mu\nu}\Gamma^\sigma_{\lambda\rho})(x_0) V_\sigma(\tau_0)(\int_{\tau_0}^{\tau_1} d\tau'\ \dot{x}^\nu(\tau')x^\rho(\tau')) \quad (11.10)$$

The final observation we need is that the remaining integral is anti-symmetric in the indices ν, ρ, which follows immediately from

$$\int_{\tau_0}^{\tau_1} d\tau'\ (\dot{x}^\nu(\tau')x^\rho(\tau') + x^\nu(\tau')\dot{x}^\rho(\tau')) = \int_{\tau_0}^{\tau_1} d\tau'\ \frac{d}{d\tau'}(x^\nu(\tau')x^\rho(\tau')) = 0 \quad . \quad (11.11)$$

It now follows from (11.10) and the definition of the Riemann tensor that

$$V_\mu(\tau_1) - V_\mu(\tau_0) = \tfrac{1}{2}(\oint x^\rho dx^\nu) R^\sigma_{\mu\rho\nu}(x_0) V_\sigma(\tau_0) \quad . \quad (11.12)$$

Simply by raising and lowering of the indices, and using the symmetry properties of the Riemann tensor, we can deduce that the corresponding equation for the parallel tansport of vectors is

$$V^\mu(\tau_1) - V^\mu(\tau_0) = -\tfrac{1}{2}(\oint x^\rho dx^\nu) R^\mu_{\sigma\rho\nu}(x_0) V^\sigma(\tau_0) \quad . \quad (11.13)$$

As an example, recall that in section 8.6 we already determined explicitly the parallel transport of vectors on the 2-sphere along the circles with fixed $\theta = \theta_0$. Choosing θ_0 infinitesimal corresponds to an infinitesimal loop around the north pole. Expanding the result (5.119) for small θ_0, in particular using

$$\sin(2\pi\cos\theta_0) \approx \sin(2\pi(1 - \theta_0^2/2)) \approx -\tfrac{1}{2}(2\pi)\theta_0^2 \quad , \quad (11.14)$$

one finds complete agreement between (11.13) and the components of the Riemann tensor of the 2-sphere, determined in (8.70),

$$r^\theta{}_{\phi\theta\phi} = \sin^2\theta \quad , \quad r^\phi{}_{\theta\phi\theta} = 1 \quad , \quad (11.15)$$

evaluated for $\theta_0 \to 0$. In verifying this, some care should be taken with the fact that $\theta = 0$ is a coordinate singularity so that one should never strictly set $\theta_0 = 0$. Alternatively, and to be on the safe side, one can rewrite (11.13) as an equation for orthonormal frame components and use the result (5.124) for the parallel transport of the frame components (which is not sensitive to coordinate singularities).

11.2 Vanishing Riemann Tensor and Existence of Flat Coordinates

We are now finally in a position to prove the converse to the statement that the Minkowski metric has vanishing Riemann tensor. Namely, we will see that when the Riemann tensor of a metric vanishes, locally there are coordinates in which the metric is the standard Minkowski metric. Since the opposite of curved is flat, this then allows one to unambiguously refer to the Minkowski metric as *the* flat metric (locally at least), and to Minkowski space as *flat space(-time)*.

So let us assume that we are given a metric with vanishing Riemann tensor. Then, by the above, parallel transport is path independent and we can, in particular, extend a vector $V^\mu(x_0)$ to a vector field everywhere in space-time: to define $V^\mu(x_1)$ we choose any path from x_0 to x_1 and use parallel transport along that path. In particular, the vector field V^μ, defined in this way, will be covariantly constant or *parallel*, $\nabla_\mu V^\nu = 0$. We can also do this for four linearly independent vectors V_a^μ at x_0 and obtain four covariantly constant (parallel) vector fields which are linearly independent at every point.

An alternative way of saying or seeing this is the following: The integrability condition for the equation $\nabla_\mu V^\lambda = 0$ is

$$\nabla_\mu V^\lambda = 0 \;\;\Rightarrow\;\; [\nabla_\mu, \nabla_\nu]V^\lambda = R^\lambda_{\;\sigma\mu\nu}V^\sigma = 0 \;\; . \tag{11.16}$$

This means that the (4×4) matrices $M(\mu, \nu)$ with coefficients $M(\mu,\nu)^\lambda_{\;\sigma} = R^\lambda_{\;\sigma\mu\nu}$ have a zero eigenvalue. If this integrability condition is satisfied, a solution to $\nabla_\mu V^\lambda$ can be found. If one wants four linearly independent parallel vector fields, then the matrices $M(\mu, \nu)$ must have four zero eigenvalues, i.e. they are zero and therefore $R^\lambda_{\;\sigma\mu\nu} = 0$. If this condition is satisfied, all the integrability conditions are satisfied and there will be four linearly independent covariantly constant vector fields - the same conclusion as above.

We will now use this result in the proof, but for covectors instead of vectors. Clearly this makes no difference: if V^μ is a parallel vector field, then $g_{\mu\nu}V^\nu$ is a parallel covector field.

Fix some point x_0. At x_0, there will be an invertible matrix e^a_μ such that

$$g^{\mu\nu}(x_0)e^a_\mu e^b_\nu = \eta^{ab} \;\; . \tag{11.17}$$

Now we solve the equations

$$\nabla_\nu E^a_\mu = 0 \Leftrightarrow \partial_\nu E^a_\mu = \Gamma^\lambda_{\ \mu\nu} E^a_\lambda \tag{11.18}$$

with the initial condition $E^a_\mu(x_0) = e^a_\mu$. This gives rise to four linearly independent parallel covectors E^a_μ.

Now it follows from (11.18) that

$$\partial_\mu E^a_\nu = \partial_\nu E^a_\mu \quad . \tag{11.19}$$

Therefore locally there are four scalars ξ^a such that

$$E^a_\mu = \frac{\partial \xi^a}{\partial x^\mu} \quad . \tag{11.20}$$

These are already the flat coordinates we have been looking for. To see this, consider the expression $g^{\mu\nu} E^a_\mu E^b_\nu$. This is clearly constant because the metric and the E^a_μ are covariantly constant,

$$\partial_\lambda (g^{\mu\nu} E^a_\mu E^b_\nu) = \nabla_\lambda (g^{\mu\nu} E^a_\mu E^b_\nu) = 0 \quad . \tag{11.21}$$

At x_0, this is just the flat metric and thus

$$(g^{\mu\nu} E^a_\mu E^b_\nu)(x) = (g^{\mu\nu} E^a_\mu E^b_\nu)(x_0) = \eta^{ab} \quad . \tag{11.22}$$

Summing this up, we have seen that, starting from the assumption that the Riemann curvature tensor of a metric $g_{\mu\nu}$ is zero, we have proven the existence of coordinates ξ^a in which the metric takes the Minkowski form,

$$g_{\mu\nu} = \frac{\partial \xi^a}{\partial x^\mu} \frac{\partial \xi^b}{\partial x^\nu} \eta_{ab} \quad . \tag{11.23}$$

The argument given above is local in the sense that the existence of these coordinates ξ^a is only guaranteed locally, i.e. in the neighbourhood of some point. Whether or not these coordinates can be used to cover the space-time globally depends on gobal (topological) properties of the space-time which are not captured by the intrinsic local and locally determined Riemann tensor.

For example, imagine starting with Minkowski space $\mathbb{R}^{1,3}$ with inertial coordinates ξ^a, and then making a periodic identification of ξ^1, say,

$$\xi^1 \sim \xi^1 + 2\pi L_1 \quad \Rightarrow \quad \mathbb{R}^{1,3} \to \mathbb{R}^{1,2} \times S^1 \quad . \tag{11.24}$$

Since the Minkowksi metric is translation-invariant, it gives rise to a well-defined metric on the periodically identified space-time, and the metric of this space-time still has zero curvature tensor. Nevertheless, in this case

- in the new space-time the coordinate ξ^1, which is now an angular variable, is not globally well defined,
- and the space-time looks like Minkowski space only locally, not globally.

11.3 Curvature of Surfaces: Euler, Gauss(-Bonnet) and Liouville

We can generalise the example of the curvature of the 2-sphere, discussed in section 8.6, somewhat, in this way connecting our considerations with the classical realm of the differential geometry of surfaces, in particular with the *Gauss Curvature*, the *Euler characteristic*, the *Gauss-Bonnet* theorem and the *Liouville Equation*.

For any 2-dimensional metric g_{ab} it is a simple exercise to derive the relation between the one independent component, say R_{1212}, of the Riemann tensor, and the scalar curvature. First of all, the Ricci tensor is

$$R_{ab} = R^c{}_{acb} = R^1{}_{a1b} + R^2{}_{a2b} \tag{11.25}$$

so that the scalar curvature is

$$R(g_{ab}) = g^{ab}R_{ab} = g^{11}R^2{}_{121} + g^{12}R^1{}_{112} + g^{21}R^2{}_{221} + g^{22}R^1{}_{212} \quad . \tag{11.26}$$

Using the fact that in 2 dimensions the components of the inverse metric are explicitly given by

$$\left(g^{ab}\right) = \frac{1}{g_{11}g_{22} - g_{12}g_{21}} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{pmatrix} \tag{11.27}$$

and the (anti-)symmetry properties (1) and (2) of the Riemann tensor, one finds

$$R(g_{ab}) = \frac{2}{g_{11}g_{22} - g_{12}g_{21}} R_{1212} \quad . \tag{11.28}$$

This is precisely the relation (8.55) between the Riemann tensor and Ricci scalar. The factor of 2 in this equation is a consequence of our (and the conventional) definition of the Riemann curvature tensor, and is responsible for the fact that the scalar curvature of the unit 2-sphere is $R = +2$. We can also write this result as

$$R_{abcd} = \tfrac{1}{2}(g_{ac}g_{bd} - g_{ad}g_{bc})R \quad \Leftrightarrow \quad R^a{}_{bcd} = \tfrac{1}{2}(\delta^a{}_c g_{bd} - \delta^a{}_d g_{bc})R \quad . \tag{11.29}$$

In two dimensions, it is often convenient and natural to absorb this ubiquitous factor of 2 into the definition of the (scalar) curvature, and what one then gets is the classical Gauss Curvature

$$K := \frac{1}{2}R(g_{ab}) \tag{11.30}$$

of a two-dimensional surface.

It follows from (11.29) that the Ricci tensor is related to the Ricci scalar by

$$R_{ab} = (R/2)g_{ab} = Kg_{ab} \quad . \tag{11.31}$$

This generalises the result for the standard metric on the 2-sphere found by explicit calculation in section 8.6. It shows that in complete generality the Ricci tensor of a two-dimensional space or space-time, thought of as the linear map

$$R^a_{\ b} = K\delta^a_{\ b} : \quad v^a \mapsto R^a_{\ b}\, v^b \quad , \tag{11.32}$$

has only one (double) eigenvalue, namely the Gauss curvature K. It can also be interpreted as saying that in 2 dimensions the Einstein tensor (8.108) is identically zero,

$$G_{ab} = R_{ab} - \tfrac{1}{2}g_{ab}R = 0 \ . \tag{11.33}$$

We will now briefly look at two important and interesting consequences of the above formulae, one related to the Euler characteristic of a surface and its integral representation (the Gauss-Bonnet theorem), and the other to the Liouville equation describing metrics with constant Gauss curvauter $K = k = \pm 1$.

1. Euler Characteristic and the Gauss-Bonnet Theorem

   Let us consider a compact closed surface S, i.e. topologically something like a sphere, or a torus (a sphere with one handle), or a sphere with several handles. A surface S_h with h handles is called a surface of *genus* h.

   Given a metric on S_h, we can associate to S_h its area with respect to the metric,

$$A(S_h) := \int_{S_h} \sqrt{g} d^2x \ . \tag{11.34}$$

   Clearly, this areas depends on a choice of metric, and under a variation δ_g of the metric it transforms as

$$\delta_g A(S_h) = \int_{S_h} \delta_g \sqrt{g} d^2x = \tfrac{1}{2} \int_{S_h} \sqrt{g} d^2x \ g^{ab} \delta g_{ab} \ . \tag{11.35}$$

   Given a metric on S_g, we can also naturally associate to it the real number

$$\chi(S_h) = \frac{1}{2\pi} \int \sqrt{g} d^2x \ K = \frac{1}{4\pi} \int \sqrt{g} d^2x \ R \ . \tag{11.36}$$

   Remarkably, this number turns out to be independent of the metric, in particular *$\chi(S_h)$ is invariant under variations of the metric g_{ab}*,

$$\delta_g \chi(S_h) = 0 \ . \tag{11.37}$$

   Here are two rather explicit ways of establishing this remarkable result:

   (a) The variation of the integrand is

$$\delta_g(\sqrt{g}R) = \delta_g(\sqrt{g}g^{ab}R_{ab}) \ . \tag{11.38}$$

   Since $R = g^{ab}R_{ab}$, it is convenient to express variations of the metric in terms of variations of the inverse metric, with (5.88)

$$\delta\sqrt{g} = \tfrac{1}{2}g^{ab}\delta g_{ab} = -\tfrac{1}{2}g_{ab}\delta g^{ab} \ . \tag{11.39}$$

Then one finds

$$\begin{aligned}\delta_g(\sqrt{g}R) &= (\delta_g\sqrt{g})g^{ab}R_{ab} + \sqrt{g}(\delta g^{ab})R_{ab} + \sqrt{g}g^{ab}\delta_g R_{ab} \\ &= \sqrt{g}(-\tfrac{1}{2}Rg_{ab} + R_{ab})\delta g^{ab} + \sqrt{g}g^{ab}\delta_g R_{ab} \\ &= \sqrt{g}G_{ab}\delta g^{ab} + \sqrt{g}g^{ab}\delta_g R_{ab} \quad .\end{aligned} \tag{11.40}$$

Now, as shown above, in 2 dimensions the Einstein tensor is identically zero, $G_{ab} = 0$. Moreover, in section 20.2 we will show that in any dimension the 2nd term is a total derivative (20.18),

$$g^{ab}\delta_g R_{ab} = \nabla_a B^a \tag{11.41}$$

for some well-defined B^a built from the covariant derivatives of the variations of the metric, as in (20.19). Taken together, these two facts imply that for a closed surface S_h (without boundary) one has

$$\delta_g\chi(S_h) = \frac{1}{4\pi}\int \sqrt{g}d^2x(G_{ab}\delta g^{ab} + \nabla_a B^a) = 0 \quad , \tag{11.42}$$

as was to be shown.

(b) Alternatively, somewhat less covariantly but very explicitly, one can show that the integrand $\sqrt{g}K$ or $\sqrt{g}R$ can itself locally be written as a total derivative. Indeed, using (11.29) to write

$$R^1{}_{212} = \tfrac{1}{2}g_{11}R \quad \Leftrightarrow \quad K = \frac{1}{g_{11}}R^2{}_{121} \tag{11.43}$$

and simply writing out explicitly this Riemann curvature tensor component in terms of the Christoffel symbols,

$$R^2{}_{121} = \partial_2\Gamma^2{}_{11} - \partial_1\Gamma^2{}_{12} + \Gamma^2{}_{2a}\Gamma^a{}_{11} - \Gamma^2{}_{1a}\Gamma^a{}_{12} \quad , \tag{11.44}$$

one finds that $\sqrt{g}K$ can be written as

$$\sqrt{g}K = \in^{ab}\partial_a\beta_b \tag{11.45}$$

with $\in^{12} = -\in^{21} = 1$ the Levi-Civita symbol, and

$$\beta_b = -\frac{\sqrt{g}}{g_{11}}\Gamma^2{}_{1b} \quad . \tag{11.46}$$

The fact that the integrand is locally a total derivative does not mean that the integral is zero (because of the non-tensorial nature of β_a, which will typically exhibit coordinate singularities). It does mean however, that the integral of the metric variation of this expression is zero (because that is tensorial and well-defined on S_h),

$$\sqrt{g}K = \varepsilon^{ab}\partial_a\beta_b \quad \Rightarrow \quad \delta_g\chi(S_h) = \frac{1}{2\pi}\int d^2x\delta_g(\sqrt{g}K) = 0 \quad . \tag{11.47}$$

Either way we have seen that the real number $\chi(S_h)$ is independent of the metric one uses to calculate it. For example, for $h = 0$ and for the standard metric on the sphere S^2 one finds

$$\chi(S_{h=0}) = \chi(S^2) = \frac{1}{4\pi}\int \sqrt{g}R = \frac{1}{2\pi}\int \sqrt{g} = 2 \ , \tag{11.48}$$

and this will therefore be the result for any metric on S^2. Likewise, for $h = 1$, i.e. a torus, by choosing the flat metric on T^2 (see e.g. the discussion and construction in section 18.1), one finds

$$\chi(S_{h=1}) = \chi(T^2) = 0 \ , \tag{11.49}$$

and this will therefore be the result for any metric on T^2 (and it is instructive to check this explicitly for the non-trivial, non-flat metric on T^2 induced by its embedding into $\mathbb{R}^3$ constructed in section 18.1). I am not aware of an equally elementary calculation to determine $\chi(S_h)$ for $h > 1$ in this way but fact of the matter is that

$$\chi(S_h) = 2 - 2h \tag{11.50}$$

is the *Euler characteristic* of S_h, which can also be defined purely combinatorially as the number

$$\chi(S) = n_F - n_E + n_V \tag{11.51}$$

of faces minus vertices plus edges of any cubist rendition of a surface S (and $\chi(S)$ is independent of such a cubist realisation or *triangulation*). The remarkable fact that this *topological invariant* of a surface S can be calculated in terms of differential geometric quantities, namely as the integral of the curvature scalar, is known as the *Gauss-Bonnet theorem.*

2. Constant Curvature and the Liouville Equation

   When we specialise the above to the class of *conformally flat* metrics with line element

$$ds^2 = \mathrm{e}^{2h(x,y)}(dx^2 + dy^2) \quad \Leftrightarrow \quad g_{ab} = \exp 2h(x,y)\delta_{ab} \tag{11.52}$$

   the calculation of the Riemann tensor is particularly simple and one finds the (easy to memorise) results

$$R^x{}_{yxy} = -\Delta h \tag{11.53}$$

   and

$$K = -\mathrm{e}^{-2h}\Delta h \tag{11.54}$$

   where $\Delta = \partial_x^2 + \partial_y^2$ is the 2-dimensional Laplacian with respect to the flat Euclidean metric $dx^2 + dy^2$. Thus a surface with *constant curvature* $K = k$ is given by a solution to the non-linear differential equation

$$\Delta h + k\mathrm{e}^{2h} = 0 \ . \tag{11.55}$$

This is the (in-)famous *Liouville equation*, which plays a fundamental role in many branches of mathematics (and mathematical physics).

In terms of the intrinsic Laplacian Δ_g associated to the metric g_{ab}, the Gaussian cuvature and the Liouville equation can also simply be written as

$$K = -\Delta_g h \qquad , \qquad \Delta_g h + k = 0 \ , \tag{11.56}$$

since, due to the peculiarities of 2 dimensions, $\sqrt{g}g^{ab}$ in independent of h, i.e. is conformally invariant (as we already observed in a different context in section 7.7, cf. (7.120)),

$$\begin{aligned} &\sqrt{g}g^{ab} = \mathrm{e}^{2h}\mathrm{e}^{-2h}\delta^{ab} = \delta^{ab} \\ \Rightarrow \quad &\Delta_g = \frac{1}{\sqrt{g}}\partial_a(\sqrt{g}g^{ab}\partial_b) = \frac{1}{\sqrt{g}}\partial_a(\delta^{ab}\partial_b) = \mathrm{e}^{-2h}\Delta \ . \end{aligned} \tag{11.57}$$

I will not attempt to say anything about the general (local) solution of this equation (which roughly speaking depends on an arbitrary meromorphic function of the complex coordinate $z = x + iy$), but close this section with some special (and particularly prominent) solutions of this equation.

(a) It is easy to see that

$$\mathrm{e}^{2h(x,y)} = y^{-2} \quad \Leftrightarrow \quad h(x,y) = -\ln y \tag{11.58}$$

solves the Liouville equation with $k = -1$. The corresponding space of constant negative curvature is the *Poincaré upper-half plane* model of the hyperbolic geometry,

$$ds^2 = \frac{dx^2 + dy^2}{y^2} \qquad (\ (x,y) \in \mathbb{R}^2,\ y > 0\ )\ . \tag{11.59}$$

By the coordinate transformation $y = e^z$ this is mapped to the equivalent metric

$$ds^2 = dz^2 + \mathrm{e}^{-2z}dx^2 \tag{11.60}$$

on the entire (x,z)-plane.

(b) Another solution (for any k) is the rotationally invariant function

$$\mathrm{e}^{2h(x,y)} = 4(1 + k(x^2+y^2))^{-2} \quad \Leftrightarrow \quad h = -\ln(1 + k(x^2+y^2)) + \text{const.} \tag{11.61}$$

i. For $k = 0$ one finds the flat (zero curvature) Euclidean metric on $\mathbb{R}^2$.

ii. For $k = +1$, one obtains the metric

$$ds^2 = 4\frac{dx^2 + dy^2}{(1 + x^2 + y^2)^2} \ . \tag{11.62}$$

This is the constant positive curvature metric on the *Riemann sphere* one gets by stereographic projection of the standard metric on the two-sphere S^2 to the (x, y)-plane.
In terms of polar coordinates (r, ϕ) on the Euclidean plane, this metric takes the form

$$ds^2 = 4\frac{dr^2 + r^2 d\phi^2}{(1 + r^2)^2} \quad , \tag{11.63}$$

and the further change of variables $r = \tan\theta/2$ shows that this is indeed the standard line element $d\Omega^2$ on the 2-sphere,

$$r = \tan\theta/2 \quad \Rightarrow \quad ds^2 = d\theta^2 + \sin^2\theta d\phi^2 = d\Omega^2 \quad . \tag{11.64}$$

Read backwards, this can also be read as the statement that via the above change of variables the Euclidean metric on $\mathbb{R}^2$ can be written as

$$dr^2 + r^2 d\phi^2 = \frac{(1 + r(\theta)^2)^2}{4}(d\theta^2 + \sin^2\theta d\phi^2) = \frac{1}{4\cos^4\theta/2} d\Omega^2 \quad . \tag{11.65}$$

In this process the points "at infinity" in $\mathbb{R}^2$ (where $r \to \infty$) are mapped to the south pole $\theta = \pi$ of the sphere (where the conformal factor in front of the line element of the sphere diverges accordingly). This exhibits S^2 as the *conformal compactification* of $\mathbb{R}^2$.

iii. For $k = -1$, one finds

$$ds^2 = 4\frac{dx^2 + dy^2}{(1 - (x^2 + y^2))^2} \qquad (\ \{x, y\} \in \mathbb{R}^2, x^2 + y^2 < 1\ ) \quad . \tag{11.66}$$

This is the *Poincaré disc* model of the hyperbolic geometry, defined in the interior of the unit disc in $\mathbb{R}^2$. In terms of polar coordinates, it can also be written as

$$ds^2 = 4\frac{dr^2 + r^2 d\phi^2}{(1 - r^2)^2} \qquad (\ 0 \le r < 1\ ) \tag{11.67}$$

The two metrics (11.59) and (11.66) are isometric, i.e. related by a (albeit not completely evident) coordinate transformation.

(c) As our final example, one other solution (given here only for $k = -1$) is

$$\mathrm{e}^{2h(x,y)} = \mathrm{e}^{2h(x)} = 4\frac{\mathrm{e}^{2x}}{(1 - \mathrm{e}^{2x})^2} = \frac{1}{\sinh^2 x} \quad . \tag{11.68}$$

While this may look obscure, it is straightforward to verify that $h = -\log|\sinh x|$ satisfies the Liouville equation with $k = -1$,

$$h = -\log|\sinh x| \quad \Rightarrow \quad \partial_x^2 h = \mathrm{e}^{2h} \quad . \tag{11.69}$$

This leads to the form (2.31) of the unit metric on H^2,

$$ds^2 = d\sigma^2 + \sinh^2\sigma\, d\phi^2 \quad . \tag{11.70}$$

Indeed, we can write this in conformally flat form as

$$ds^2 = \sinh^2 \sigma(x)(dx^2 + d\phi^2) \tag{11.71}$$

where

$$x = \log \tanh(\sigma/2) \quad \Rightarrow \quad dx = d\sigma / \sinh \sigma \ . \tag{11.72}$$

Performing the exponential / hyperbolic gymnastics required to write the conformal factor $\sinh^2 \sigma$ as a function of x, one finds

$$\sinh^2 \sigma(x) = \frac{1}{\sinh^2 x} \ , \tag{11.73}$$

giving rise to the solution given in (11.68).

It is worth remarking that the Poincaré upper-half plane model of a space with constant negative curvature readily generalises to arbitrary dimensions and signature. Thus

$$ds^2 = \frac{d\vec{x}^2 + dy^2}{y^2} \quad , \quad d\vec{x}^2 = \delta_{ab} dx^a dx^b \quad \text{or} \quad d\vec{x}^2 = \eta_{ab} dx^a dx^b \tag{11.74}$$

is the metric of a $D = (d+1)$-dimensional space(-time) with constant negative curvature.

The Lorentzian metric will reappear later as a solution to the Einstein equations with a negative cosmological constant, and is in this context known as the *anti-de Sitter metric* (in Poincaré coordinates, which cover only a part of the complete space-time), and we will discuss this solution in some detail in section 39.

11.4 The Weyl Tensor and its Uses

In section 8 from the Riemann tensor we have extracted its traces, the Ricci tensor and the Ricci scalar, as well as a particular linear combination of them, the Einstein tensor. We can therefore also explicitly decompose the Riemann tensor into these trace parts and the remaining traceless part.

We noted in section 8.5 that for $D = 2$ and $D = 3$ the Riemann tensor would be "pure trace", i.e. could be written entirely in terms of the Ricci tensor and Ricci scalar. For $D = 2$ we have already established this explicitly by proving the relation (11.29),

$$D = 2: \quad R_{\alpha\beta\gamma\delta} = \tfrac{1}{2}(g_{\alpha\gamma} g_{\beta\delta} - g_{\alpha\delta} g_{\beta\gamma}) R \tag{11.75}$$

in section 11.3.

We now look at this issue for $D \geq 3$. Simply by linear algebra one finds, for $D \geq 3$, the decomposition

$$\begin{aligned} R_{\mu\nu\rho\sigma} &= C_{\mu\nu\rho\sigma} \\ &\quad + \frac{1}{D-2}(g_{\mu\rho} R_{\nu\sigma} + R_{\mu\rho} g_{\nu\sigma} - g_{\nu\rho} R_{\mu\sigma} - R_{\nu\rho} g_{\mu\sigma}) \\ &\quad - \frac{1}{(D-1)(D-2)} R (g_{\mu\rho} g_{\nu\sigma} - g_{\nu\rho} g_{\mu\sigma}) \ . \end{aligned} \tag{11.76}$$

This definition is such that $C_{\mu\nu\rho\sigma}$ has all the symmetries of the Riemann tensor (this is manifest) and such that all of its traces are zero, i.e.

$$C^{\mu}_{\ \nu\mu\sigma} = 0 \ , \tag{11.77}$$

as is easily verified. This traceless part $C_{\mu\nu\rho\sigma}$ of the Riemann tensor is called the *Weyl tensor*.

Occasionally it is more convenient and transparent to decompose the Riemann tensor not into the Weyl tensor, the Ricci tensor and the Ricci scalar, but to perform an orthogonal decomposition (with respect to the metric) into the Weyl tensor, the traceless part $S_{\mu\nu}$ of the Ricci tensor,

$$S_{\mu\nu} = R_{\mu\nu} - \frac{1}{D} g_{\mu\nu} R \ , \tag{11.78}$$

and the trace R. Then the decomposition becomes

$$\begin{aligned} R_{\mu\nu\rho\sigma} &= C_{\mu\nu\rho\sigma} \\ &+ \frac{1}{D-2}(g_{\mu\rho}S_{\nu\sigma} + S_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}S_{\mu\sigma} - S_{\nu\rho}g_{\mu\sigma}) \\ &+ \frac{1}{D(D-1)} R(g_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}g_{\mu\sigma}) \ . \end{aligned} \tag{11.79}$$

One other common and convenient decomposition is in terms of a tensor $P_{\mu\nu}$ such that (11.76) takes the form

$$R_{\mu\nu\rho\sigma} = C_{\mu\nu\rho\sigma} + (g_{\mu\rho}P_{\nu\sigma} + P_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}P_{\mu\sigma} - P_{\nu\rho}g_{\mu\sigma}) \ . \tag{11.80}$$

Comparison with (11.76) shows that this is accomplished by the choice

$$P_{\mu\nu} = \frac{1}{D-2}\left(R_{\mu\nu} - \frac{1}{2(D-1)} g_{\mu\nu} R\right) \ . \tag{11.81}$$

This tensor $P_{\mu\nu}$ is known as the *Schouten Tensor*.

Regardless of how we write the trace part of the Riemann tensor, it turns out that for $D = 3$ the Weyl tensor vanishes identically,

$$D = 3: \quad C_{\alpha\beta\gamma\delta} \equiv 0 \tag{11.82}$$

(I will give an elementary proof of this momentarily). Therefore, for $D = 3$ one has the decomposition

$$\begin{aligned} D = 3: \quad R_{\alpha\beta\gamma\delta} &= (g_{\alpha\gamma}R_{\beta\delta} + R_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}R_{\beta\gamma} - R_{\alpha\delta}g_{\beta\gamma}) \\ &+ \tfrac{1}{2}(g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta})R \end{aligned} \tag{11.83}$$

This is precisely the result claimed previously in (8.56).

To establish (11.82), in order to trivialise the algebra let us fix a point x_0 and choose coordinates there such that $g_{\alpha\beta}(x_0) = \delta_{\alpha\beta}$ (or $\eta_{\alpha\beta}$, depending on the signature of the metric, but let us assume that we are in the case of Euclidean signature - the same argument works in the Lorentzian case). Now the proof consists of the following elementary steps:

- Since we are in $D = 3$, at least two of the indices in $C_{\alpha\beta\gamma\delta}$ must be equal. Since the Weyl tensor has all the symmetries of the Riemann tensor, if more than two indices are equal, the Weyl tensor component is zero. Thus we only need to consider the components where 2 indices are equal and we can without loss of generality choose these to be $C_{1\beta 1\delta}$, say, with $\beta, \delta \neq 1$.

- Because the Weyl tensor is traceless, one has the relation

$$C_{1\beta 1\delta} = -C_{2\beta 2\delta} - C_{3\beta 3\delta} \ . \tag{11.84}$$

  This implies that $C_{1213} = C_{1312} = 0$, so a non-zero component requires $\beta = \delta$.

- For $\beta = \delta$, one derives from this

$$C_{1212} = -C_{3232} \quad \text{and} \quad C_{1313} = -C_{2323} \tag{11.85}$$

  and likewise for $\alpha = \gamma = 2$ and $\alpha = \gamma = 3$.

- Thus all in all the Weyl tensor can have only 3 independent non-zero components, namely $C_{1212} = C_{2121}, C_{1313} = C_{3131}, C_{2323} = C_{3232}$, and they are all required to be pairwise negatives of each other. This is impossible for non-trivial $C_{\alpha\beta\gamma\delta}$,

$$C_{1212} = -C_{3232} = +C_{3131} = -C_{2121} = -C_{1212} \quad \Rightarrow \quad C_{1212} = 0 \ , \tag{11.86}$$

  and implies that all of the components of the Weyl tensor are identically zero in $D = 3$.

Thus the Weyl tensor is only non-trivial for $D \geq 4$. Using the Bianchi identies discussed in section 8.8, in particular also (8.105),

$$\nabla^\mu R_{\mu\nu\rho\sigma} = \nabla_\rho R_{\nu\sigma} - \nabla_\sigma R_{\nu\rho} \tag{11.87}$$

from (11.76) one finds a simple expression for the divergence, namely

$$\nabla^\mu C_{\mu\nu\rho\sigma} = (D-3)\left(\nabla_\rho P_{\nu\sigma} - \nabla_\sigma P_{\nu\rho}\right) \ . \tag{11.88}$$

The tensor appearing on the right-hand side also has its own name. It is called the *Cotton Tensor* $C_{\nu\rho\sigma}$,

$$C_{\nu\rho\sigma} = \nabla_\rho P_{\nu\sigma} - \nabla_\sigma P_{\nu\rho} \ . \tag{11.89}$$

The content of (11.88) is evidently trivial in $D = 3$, but the Cotton tensor itself is not (and I will briefly come back to this below).

The Weyl tensor plays an important role in many aspects of gravitational physics:

1. For example, the Weyl tensor has traditionally been one of the central objects of interest in the invariant algebraic classification of gravitational fields and in

the characterisation of what are known as algebraically special solutions to the Einstein equations (the so-called *Petrov classification* and related procedures). Originally, this was (of course) developed for $D=4$, and this case has a number of special features. It is based on the classification of the properties of the eigenvalues λ of the Weyl tensor (at a point x_0), thought of as a map on the space of anti-symmetric $(2,0)$-tensors (*bivectors*),

$$\tfrac{1}{2}C^{\alpha\beta}{}_{\gamma\delta}X^{\gamma\delta} = \lambda X^{\alpha\beta} \tag{11.90}$$

or

$$C^A_B X^B = \lambda X^A \tag{11.91}$$

with $C_{\alpha\beta\gamma\delta} \equiv C_{AB}$ thought of as a symmetric (6×6) matrix.[27]

An equivalent (as it turns out) classification arises from determining the number and multiplicity of linearly independent null vectors ℓ satisfying the condition

$$\ell_{[\gamma}C_{\alpha]\mu\nu[\beta}\ell_{\delta]}\ell^\mu\ell^\nu = 0 \ . \tag{11.92}$$

Such ℓ^α are called the *principal null directions* of the Weyl tensor. More recently, this classification scheme (based on the latter approach) has been (partially) extended to higher dimensions.[28]

2. As we will see in section 19.6, the Einstein equations imply that the Weyl tensor describes the gravitational field in vacuum. Specifically, when (or where) the energy-momentum tensor is zero, the Riemann curvature tensor is equal to the Weyl tensor,

$$T_{\alpha\beta}(x) = 0 \quad \Rightarrow \quad R_{\alpha\beta\gamma\delta}(x) = C_{\alpha\beta\gamma\delta}(x) \ . \tag{11.93}$$

The Weyl tensor thus encodes the information about things like gravitational waves and the asymptotic behaviour of a gravitational field and has been studied extensively from this point of view.

3. In the presence of matter, on the other hand, (11.88), in conjunction with the Einstein equations, becomes an evolution equation for these vacuum components of the gravitational field in terms of the sources - see equations (19.57) and (19.58).

The Weyl tensor also plays an important role in geometry, as it is *conformally invariant*, i.e. $C^\mu_{\nu\rho\sigma}$ is invariant under conformal (Weyl) rescalings of the metric,

$$g_{\mu\nu}(x) \to \mathrm{e}^{2f(x)}g_{\mu\nu}(x) \quad \Rightarrow \quad C^\mu_{\nu\rho\sigma} \to C^\mu_{\nu\rho\sigma} \ , \tag{11.94}$$

[27] See e.g. sections 7.16 and 11 of J. Plebanski, A. Krasinski, *An Introduction to General Relativity and Cosmology*, or section 4 of H. Stephani, D. Kramer, M. MacCallum, C. Hoenslaers, E. Herlt, *Exact Solutions to Einstein's Field Equations* (2nd Edition) for recent expositions of this subject.

[28] See A. Coley, R. Milson, V. Pravda, A. Pravdova, *Classification of the Weyl Tensor in Higher Dimensions*, `arXiv:arXiv:gr-qc/0401008`; A. Coley, *Classification of the Weyl Tensor in Higher Dimensions and Applications*, `arXiv:0710.1598 [gr-qc]`; M. Ortaggio, V. Pravda, A. Pravdova, *Algebraic classification of higher dimensional spacetimes based on null alignment*, `arXiv:1211.7289 [gr-qc]`.

equivalently

$$g_{\mu\nu}(x) \to \mathrm{e}^{2f(x)} g_{\mu\nu}(x) \quad \Rightarrow \quad C_{\mu\nu\rho\sigma} \to \mathrm{e}^{2f(x)} C_{\mu\nu\rho\sigma} \quad . \tag{11.95}$$

In particular, the Weyl tensor is zero if the metric is *conformally flat*, i.e. related by a conformal transformation to the flat metric $\eta_{\mu\nu}$ (of any signature),

$$g_{\mu\nu}(x) = \mathrm{e}^{2f(x)} \eta_{\mu\nu}(x) \quad \Rightarrow \quad C^{\mu}_{\ \nu\rho\sigma} = 0 \quad . \tag{11.96}$$

This can be established by brute force calculation and is not *per se* particularly enlightning.

Conversely for $D \geq 4$ vanishing of the Weyl tensor is also a sufficient condition for a metric to be (locally) conformal to the flat metric. This is a non-trivial result because at face value one seems to obtain a completely overdetermined system of equations for the single function f, of the form

$$P_{\mu\nu} = \{\text{1st and 2nd derivatives of } f\}_{\mu\nu} \quad . \tag{11.97}$$

However, it turns out that the integrability conditions for this system of equations are equivalent to the vanishing of the Weyl tensor, and then a variant of the *Frobenius integrability theorem* (mentioned in a different context in section 15.5) can be used to establish the local existence of a solution f.

For $D = 3$, the situation is slightly (but not fundamentally) different. We see from (11.89) that for any $D \geq 4$ conformal flatness implies vanishing of the Cotton tensor. It turns out that for $D = 3$ the Cotton tensor takes over the role of the Weyl tensor (which, as proven above, is itself trivial for $D = 3$), i.e. one has the statement that for $D = 3$ a metric is (locally) conformally flat if and only if the Cotton tensor vanishes.

11.5 Generalisations: Torsion and Non-Metricity

In section 5.4 we had seen that the Levi-Civita connection (defined by the Christoffel symbols) is characterised by the fact that

1. the metric is covariantly constant, $\nabla_\mu g_{\nu\lambda} = 0$, and
2. the torsion is zero, i.e. the second covariant derivatives of a scalar commute.

It is of course possible to relax either of the conditions (1) or (2), or both of them and, in particular, connections with torsion (relaxation of condition 2) are popular in certain circles and/or arise naturally in certain generalised (gauge) theories of gravity and in string theory.

To discuss this a bit more systematically, we consider a general connection

$$\tilde{\Gamma}^{\mu}_{\ \nu\lambda} = \Gamma^{\mu}_{\ \nu\lambda} + C^{\mu}_{\ \nu\lambda} \tag{11.98}$$

with $\Gamma^{\mu}_{\ \nu\lambda}$ the canonical Levi-Civita connection, and $C^{\mu}_{\ \nu\lambda}$ a $(1,2)$-tensor. We will also use the corresponding $(0,3)$-tensor

$$C_{\mu\nu\lambda} = g_{\mu\rho} C^{\rho}_{\ \nu\lambda} \quad . \tag{11.99}$$

Associated with $\tilde{\Gamma}^{\mu}_{\ \nu\lambda}$ we have the covariant derivative $\tilde{\nabla}_{\mu}$. Since $\tilde{\Gamma}^{\mu}_{\ \nu\lambda}$ will in general not be symmetric in its lower indices, in this section we need to be particularly careful with (and choose a convention for) the ordering of the lower indices in the covariant derivative. We will choose the convention that the last index always refers to the direction along which one is differentiating, i.e.

$$\tilde{\nabla}_{\mu} V^{\nu} = \partial_{\mu} V^{\nu} + \tilde{\Gamma}^{\nu}_{\ \lambda\mu} V^{\lambda} \tag{11.100}$$

etc. The reason for this choice is that one should think of the collection of objects $\Gamma^{\nu}_{\ \lambda\mu}$ (and $\tilde{\Gamma}^{\nu}_{\ \lambda\mu}$) as the coefficients of a matrix-valued 1-form (cf. section 4.6) $\Gamma^{\nu}_{\ \lambda} = \Gamma^{\nu}_{\ \lambda\mu} dx^{\mu}$, the matrices acting by rotation on vectors (and more general tensors), as in (5.21).

We now define the *torsion tensor* $T^{\lambda}_{\mu\nu}$ by

$$[\tilde{\nabla}_{\mu}, \tilde{\nabla}_{\nu}]\phi = T^{\lambda}_{\mu\nu} \partial_{\lambda}\phi \quad , \quad T_{\lambda\mu\nu} = g_{\lambda\rho} T^{\rho}_{\mu\nu} \quad , \tag{11.101}$$

and the *non-metricity tensor* $Q_{\nu\lambda\mu}$ by

$$\tilde{\nabla}_{\mu} g_{\nu\lambda} = -Q_{\nu\lambda\mu} \quad . \tag{11.102}$$

In terms of the $C^{\mu}_{\nu\lambda}$ these tensors can be written as

$$\begin{aligned} T_{\lambda\mu\nu} &= C_{\lambda\mu\nu} - C_{\lambda\nu\mu} = 2C_{\lambda[\mu\nu]} \\ Q_{\nu\lambda\mu} &= C_{\nu\lambda\mu} + C_{\lambda\nu\mu} = 2C_{(\nu\lambda)\mu} \quad . \end{aligned} \tag{11.103}$$

Thus the torsion is zero iff $C^{\lambda}_{\mu\nu}$ (and hence $\tilde{\Gamma}^{\lambda}_{\ \mu\nu}$) is symmetric in its lower indices, and the connection is compatible with the metric iff $C_{\nu\lambda\mu}$ is anti-symmetric in its first two indices. In particular, if the torsion is zero and the connection is metric-compatible, one has

$$C_{\lambda\mu\nu} = C_{\lambda\nu\mu} \quad \text{and} \quad C_{\lambda\mu\nu} = -C_{\mu\lambda\nu} \quad \Rightarrow \quad C_{\lambda\mu\nu} = 0 \ , \tag{11.104}$$

as one can see by the gymnastics

$$C_{\lambda\mu\nu} = C_{\lambda\nu\mu} = -C_{\nu\lambda\mu} = -C_{\nu\mu\lambda} = C_{\mu\nu\lambda} = C_{\mu\lambda\nu} = -C_{\lambda\mu\nu} \quad . \tag{11.105}$$

Conversely, since the absence of torsion and non-metricity characterises the Levi-Civita connection, it should be possible to express the deviation $C_{\lambda\mu\nu}$ from the Levi-Civita connection entirely in terms of torsion and non-metricity. This is indeed the case. By repeating the calculation (5.44) in this more general context, one finds

$$2C_{\lambda(\mu\nu)} = Q_{\nu\lambda\mu} + Q_{\mu\lambda\nu} - Q_{\mu\nu\lambda} - T_{\mu\lambda\nu} - T_{\nu\lambda\mu} \quad . \tag{11.106}$$

Combining this with $2C_{\lambda[\mu\nu]} = T_{\lambda\mu\nu}$, one obtains

$$\begin{aligned} C_{\lambda\mu\nu} &= \tfrac{1}{2}(T_{\lambda\mu\nu} - T_{\mu\lambda\nu} - T_{\nu\lambda\mu}) + \tfrac{1}{2}(Q_{\mu\lambda\nu} + Q_{\nu\lambda\mu} - Q_{\mu\nu\lambda}) \\ &\equiv \tilde{T}_{\lambda\mu\nu} + \tilde{Q}_{\lambda\mu\nu} \quad , \end{aligned} \tag{11.107}$$

with

$$\tilde{T}_{\lambda\mu\nu} = \tfrac{1}{2}(T_{\lambda\mu\nu} - T_{\mu\lambda\nu} - T_{\nu\lambda\mu}) = -\tilde{T}_{\mu\lambda\nu} \tag{11.108}$$

and

$$\tilde{Q}_{\lambda\mu\nu} = \tfrac{1}{2}(Q_{\mu\lambda\nu} + Q_{\nu\lambda\mu} - Q_{\mu\nu\lambda}) = \tilde{Q}_{\lambda\nu\mu} \quad . \tag{11.109}$$

Thus we can now split a general connection more informatively into the 3 pieces

$$\tilde{\Gamma}^{\lambda}_{\ \mu\nu} = \Gamma^{\lambda}_{\ \mu\nu} + \tilde{T}^{\lambda}_{\ \mu\nu} + \tilde{Q}^{\lambda}_{\ \mu\nu} \quad . \tag{11.110}$$

<u>Remarks:</u>

1. The tensor $\tilde{T}_{\lambda\mu\nu}$ is known as the *contorsion tensor* (frequently (mis-)spelled as "contortion" tensor). I am not aware of a commonly used name for $\tilde{Q}_{\lambda\mu\nu}$, and will not try to invent one. The contorsion tensor is the linear combination of components of the torsion tensor that appear as the connection coefficents of a general metric-compatible connection with torsion,

$$Q_{\lambda\mu\nu} = 0 \quad \Rightarrow \quad \tilde{\Gamma}^{\lambda}_{\ \mu\nu} = \Gamma^{\lambda}_{\ \mu\nu} + \tilde{T}^{\lambda}_{\ \mu\nu} \quad . \tag{11.111}$$

   In general it can have both symmetric and anti-symmetric components,

$$\tilde{T}_{\lambda(\mu\nu)} = -\tfrac{1}{2}(T_{\mu\lambda\nu} + T_{\nu\lambda\mu}) = T_{(\mu\nu)\lambda} \quad , \quad \tilde{T}^{\lambda}_{\ [\mu\nu]} = \tfrac{1}{2}T^{\lambda}_{\ \mu\nu} \quad , \tag{11.112}$$

   but it cannot be symmetric (if the contorsion were symmetric, the torsion, and hence the contorsion, would be zero). If its symmetric part vanishes, then $\tilde{T}_{\lambda\mu\nu}$ is completely anti-symmetric,

$$\tilde{T}^{\lambda}_{\ (\mu\nu)} = 0 \quad \Rightarrow \quad \tilde{T}_{\lambda\mu\nu} = \tilde{T}_{[\lambda\mu\nu]} \quad . \tag{11.113}$$

2. Since $\Gamma_{\lambda\mu\nu}$ contains a part that is antisymmetric in the first 2 indices,

$$\Gamma_{\lambda\mu\nu} = \tfrac{1}{2}g_{\lambda\mu,\nu} + \tfrac{1}{2}(g_{\nu\lambda,\mu} - g_{\nu\mu,\lambda}) \quad , \tag{11.114}$$

   one might be tempted to think that that part can be cancelled (or absorbed) by a metric-compatible $C_{\lambda\mu\nu} = C_{[\lambda\mu]\nu}$, so that a very simple metric-compatible connection would be

$$\tilde{\Gamma}_{\lambda\mu\nu} \stackrel{?}{=} \tfrac{1}{2}g_{\lambda\mu,\nu} \quad , \quad \tilde{\Gamma}^{\lambda}_{\ \mu\nu} \stackrel{?}{=} \tfrac{1}{2}g^{\lambda\rho}g_{\rho\mu,\nu} \quad . \tag{11.115}$$

   However, the term that one has cancelled (or absorbed) is *not* a tensor. Therefore, this candidate "connection" does not transform as (and therefore does not qualify as) a connection and cannot be used to define a covariant derivative.

3. In general, for a connection $\tilde{\nabla}$, the notions of autoparallels (section 5.8),

$$\tilde{\nabla}_\tau X^\mu = 0 \quad \Leftrightarrow \quad \ddot{x}^\mu + \tilde{\Gamma}^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0 \;, \tag{11.116}$$

(i.e. curves characterised by the fact that their tangent vectors are parallel transported along the curve - this depends on a choice of connection) no longer coincides with the notion of geodesics (which are obtained by extremising proper time or distance, and which always lead to the Levi-Civita connection). However, this difference disappears if $C^\mu_{\nu\lambda}$ happens to be anti-symmetric in its lower indices (e.g. for a metric-compatible connection with totally anti-symmetric contorsion tensor), as one then has

$$\ddot{x}^\mu + \tilde{\Gamma}^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda = \ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda \;. \tag{11.117}$$

We have defined the Riemann tensor via the commutator of covariant derivatives (8.2)

$$[\nabla_\mu, \nabla_\nu]V^\lambda = R^\lambda_{\ \sigma\mu\nu}V^\sigma \tag{11.118}$$

associated to the Levi-Civita connection (Christoffel symbols), or, equivalently, by the relation (8.5) (now being careful with the positioning of the lower indices)

$$R^\lambda_{\ \sigma\mu\nu} = \partial_\mu\Gamma^\lambda_{\ \sigma\nu} - \partial_\nu\Gamma^\lambda_{\ \sigma\mu} + \Gamma^\lambda_{\ \rho\mu}\Gamma^\rho_{\ \sigma\nu} - \Gamma^\lambda_{\ \rho\nu}\Gamma^\rho_{\ \sigma\mu} \;. \tag{11.119}$$

In order to show explicitly (rather than by appealing to (11.118)) that this transforms as a tensor, all that one needs is the characteristic non-tensorial transformation behaviour of the Christoffel symbols $\Gamma^\lambda_{\ \mu\nu}$. As discussed in section 5.4 and above, an arbitrary connection $\tilde{\Gamma}^\lambda_{\ \mu\nu}$ that can be used to define a tensorial covariant derivative has the same non-tensorial transformation behaviour. Therefore

$$\tilde{R}^\lambda_{\ \sigma\mu\nu} \equiv R^\lambda_{\ \sigma\mu\nu}(\tilde{\Gamma}) = \partial_\mu\tilde{\Gamma}^\lambda_{\ \sigma\nu} - \partial_\nu\tilde{\Gamma}^\lambda_{\ \sigma\mu} + \tilde{\Gamma}^\lambda_{\ \rho\mu}\tilde{\Gamma}^\rho_{\ \sigma\nu} - \tilde{\Gamma}^\lambda_{\ \rho\mu}\tilde{\Gamma}^\rho_{\ \sigma\mu} \tag{11.120}$$

defines a tensor for any connection, namely the curvature tensor of the connection $\tilde{\Gamma}^\lambda_{\ \mu\nu}$. It is related to the commutator of covariant derivatives by

$$[\tilde{\nabla}_\mu, \tilde{\nabla}_\nu]V^\lambda = \tilde{R}^\lambda_{\ \sigma\mu\nu}V^\sigma + (\tilde{\Gamma}^\rho_{\ \mu\nu} - \tilde{\Gamma}^\rho_{\ \nu\mu})\nabla_\rho V^\lambda = \tilde{R}^\lambda_{\ \sigma\mu\nu}V^\sigma + T^\rho_{\mu\nu}\nabla_\rho V^\lambda \;, \tag{11.121}$$

where $T^\rho_{\mu\nu}$ is the torsion tensor. As before, one can also define the Ricci tensor and Ricci scalar by

$$\tilde{R}_{\mu\nu} \equiv R_{\mu\nu}(\tilde{\Gamma}) = \tilde{R}^\lambda_{\ \mu\lambda\nu} \quad , \quad \tilde{R} \equiv R(\tilde{\Gamma}) = g^{\mu\nu}\tilde{R}_{\mu\nu} \;. \tag{11.122}$$

However, it is crucial to keep in mind that the symmetry properties and Bianchi identities satisfied by these generalised curvature tensors will in general differ from those of the Riemann-Christoffel tensor. This should be clear from the way we derived the symmetries of the Riemann tensor in section 8.3, where we related the symmetries to the properties (metricity, no torsion) that characterise the canonical Levi-Civita connection (Christoffel symbols). For example, in general the Ricci tensor will not be symmetric, the Bianchi identity $R_{\alpha[\beta\gamma\delta]} = 0$ will be replaced by an identity relating $\tilde{R}_{\alpha[\beta\gamma\delta]}$ to the torsion (and its covariant derivative), etc.[29]

[29] For more on this and related topics, see e.g. section 1 of T. Ortin, *Gravity and Strings*.

For some further discussion of connections with non-metricity or torsion and their curvature tensors see section 20.7.

12 Curvature III: Curvature and Geodesic Congruences

In section 8.4 we had already encountered the so-called *geodesic deviation equation* (8.45),

$$(D_\tau)^2\delta x^\mu = R^\mu_{\ \nu\lambda\rho}\dot{x}^\nu\dot{x}^\lambda\delta x^\rho \quad , \tag{12.1}$$

describing the evolution of a separation (or deviation) vector along a given geodesic. There we had seen one non-covariant derivation, and one covariant derivation based on the identity (8.16) for the commutator of covariant derivatives along families of curves. In this section we will rederive this result in a somewhat more general covariant manner and then subsequently use the same covariant framework to discuss the extension of these results to the so-called *Raychaudhuri equation*, which descibes the focussing properties of congruences of geodesics.

12.1 Covariant Derivation of the Geodesic Deviation Equation

The starting point is an affinely parametrised geodesic with tangent vector field u^α,

$$u^\alpha u_\alpha = \pm 1, 0 \quad , \quad u^\beta\nabla_\beta u^\alpha = 0 \quad , \tag{12.2}$$

and a deviation vector field $\delta x^\alpha = \xi^\alpha$ characterised by the condition

$$[u,\xi]^\alpha = u^\beta\nabla_\beta\xi^\alpha - \xi^\beta\nabla_\beta u^\alpha = 0 \quad \Leftrightarrow \quad D_\tau\xi^\alpha = \xi^\beta\nabla_\beta u^\alpha \quad . \tag{12.3}$$

The rationale for this condition is that, if $x^\alpha(\tau,s)$ is a family of such affinely parametrised geodesics labelled by s, one has the identifications

$$u^\alpha = \frac{\partial}{\partial\tau}x^\alpha(\tau,s) \quad , \quad \xi^\alpha = \frac{\partial}{\partial s}x^\alpha(\tau,s) \quad . \tag{12.4}$$

Since second partial derivatives commute, this implies the relation

$$\frac{\partial}{\partial\tau}\xi^\alpha(\tau,s) = \frac{\partial}{\partial s}u^\alpha(\tau,s) \quad , \tag{12.5}$$

(implicit in the identification $\delta\dot{x} = (d/d\tau)\delta x$ employed in the derivation in section 8.4). Condition (12.3) is nothing other than the covariant way of writing (12.5).

Introducing the tensor

$$B_{\alpha\beta} = \nabla_\beta u_\alpha \tag{12.6}$$

we can write (12.3) as

$$D_\tau\xi^\alpha = B^\alpha_{\ \beta}\xi^\beta \quad , \tag{12.7}$$

so $B_{\alpha\beta}$ describes the evolution and deformation of the deviation vector ξ^α along the geodesic. It will play an important role in the following, and we will take a closer look at its properties in sections 12.2 and 12.4 for timelike and null geodesic congruences respectively. For now, all we need is that

$$u_\alpha B^\alpha_{\ \beta}\xi^\beta = \xi^\beta u_\alpha \nabla_\beta u^\alpha = \tfrac{1}{2}\xi^\beta \nabla_\beta (u^\alpha u_\alpha) = 0 \tag{12.8}$$

because $u^\alpha u_\alpha$ is constant over the entire family of geodesic.

As a consequence one has

$$u_\alpha D_\tau \xi^\alpha = u_\alpha B^\alpha_{\ \beta}\xi^\beta = 0 \ . \tag{12.9}$$

(i.e. $D_\tau \xi^\alpha$ is transverse to u^α) and therefore

$$\frac{d}{d\tau}(u_\alpha \xi^\alpha) = D_\tau(u_\alpha \xi^\alpha) = u_\alpha D_\tau \xi^\alpha = 0 \ . \tag{12.10}$$

This means that the u-component of a geodesic deviation vector ξ in the sense of $u_\alpha \xi^\alpha$ is simply constant and contains no interesting information about the geodesic itself.

In the timelike case this means that a vector of the form $\xi^\alpha = \xi u^\alpha$ is a deviation vector only if ξ is constant, and then ξ^α is simply a translation along the geodesic and therefore not a deviation vector of interest (and certainly anyhow not a vector of the kind one has in mind when thinking about a deviation vector, which should point away from the geodesic). In the null case, the interpretation is slightly different (and we will return to this in section 12.4), but the fact that $u_\alpha \xi^\alpha$ is simply constant for a deviation vector remains, and we can without loss of information choose the deviation vector to satisfy the condition $\xi^\alpha u_\alpha = 0$.

Given this set-up, we now want to calculate

$$\begin{aligned}(D_\tau)^2 \xi^\alpha &= (D_\tau B^\alpha_{\ \beta})\xi^\beta + B^\alpha_{\ \beta} D_\tau \xi^\beta \\ &= (D_\tau B^\alpha_{\ \gamma} + B^\alpha_{\ \beta} B^\beta_{\ \gamma})\xi^\gamma \ .\end{aligned} \tag{12.11}$$

Note that, along with $D_\tau \xi^\alpha$, also $D_\tau^2 \xi^\alpha$ is automatically transverse to u^α, $u_\alpha D_\tau^2 \xi^\alpha = 0$, regardless of whether or not one imposes the condition $\xi^\alpha u_\alpha = 0$.

For the term in brackets we find, using the geodesic equation for u^α,

$$\begin{aligned} D_\tau B^\alpha_{\ \gamma} + B^\alpha_{\ \beta} B^\beta_{\ \gamma} &= u^\beta \nabla_\beta \nabla_\gamma u^\alpha + (\nabla_\gamma u^\beta)\nabla_\beta u^\alpha \\ &= u^\beta \nabla_\beta \nabla_\gamma u^\alpha + \nabla_\gamma (u^\beta \nabla_\beta u^\alpha) - u^\beta \nabla_\gamma \nabla_\beta u^\alpha \\ &= u^\beta (\nabla_\beta \nabla_\gamma - \nabla_\gamma \nabla_\beta) u^\alpha = R^\alpha_{\ \delta\beta\gamma} u^\beta u^\delta \ , \end{aligned} \tag{12.12}$$

and plugging this back into (12.11), we obtain straightaway the covariant version (8.45) of the geodesic deviation equation in the form

$$(D_\tau)^2 \xi^\alpha = R^\alpha_{\ \delta\beta\gamma} u^\delta u^\beta \xi^\gamma \ . \tag{12.13}$$

Note that this result is automatically transverse to u^α,

$$u_\alpha (D_\tau)^2 \xi^\alpha = R_{\alpha\delta\beta\gamma} u^\alpha u^\delta u^\beta \xi^\gamma = 0 \ . \tag{12.14}$$

REMARKS:

1. The object we have called $B_{\alpha\beta}$ in (12.6) and its evolution equation (12.12) will play a central role in our derivation of the Raychaudhuri equation below.

2. When considering a null geodesic, the condition $\xi^\alpha u_\alpha = 0$ does not eliminate the component of ξ^α tangent to the geodesic. In that case it is convenient to introduce an auxiliarly linearly independent null vector field in order to be able to project the deviation vector ξ^α and its derivatives $D_\tau \xi^\alpha$ and $D^2_\tau \xi^\alpha$ into some spatial codimension 2 plane transverse to these null directions. This *transverse null geodesic equation* will be derived and discussed in section 12.3.

3. If the curve is not a geodesic (but still parametrised by proper time, so that $u^\alpha u_\alpha = -1$), then the above derivation shows that in addition to the force exerted by the space-time curvature the deviation vector feels a force proportional to the change of the acceleration $a^\alpha = u^\beta \nabla_\beta u^\alpha$ along the curve,

$$(D_\tau)^2 \xi^\alpha = R^\alpha_{\ \delta\beta\gamma} u^\delta u^\beta \xi^\gamma + D_\tau a^\alpha \quad . \tag{12.15}$$

   In flat space, only the last term is present and describes the (tidal) forces arising from the possible non-uniformity of the external force acting on the particle (or, better: on the extended object described by a family of worldlines) to produce the acceleration a^α. Thus, in precise analogy with the Newtonian situation, the gravitational (i.e. here now Riemann curvature tensor) contribution to the geodesic deviation equation should be interpreted as the *gravitational tidal force*.

12.2 Raychaudhuri Equation for Timelike Geodesic Congruences

A *congruence of curves* is a (locally) space-time filling family of curves, i.e. it is such that locally for any space-time point there is a unique curve passing through that point. A (timelike) *geodesic congruence* is then a congruence of (timelike) geodesics.

Manipulations similar to those leading to (12.13) allow one to derive an equation for the rate of change of the divergence $\nabla_\alpha u^\alpha$ of a family of geodesics along the geodesics. This simple result, known as the *Raychaudhuri equation*, has important implications and ramifications in general relativity, in particular in the context of the so-called *singularity theorems* of Penrose, Hawking and others, none of which will, however, be explored here (see footnote 98 of section 29.3 for some references).

Thus u^α now denotes a tangent vector field to an affinely parametrised geodesic congruence, $u^\alpha \nabla_\alpha u^\beta = 0$ (and $u^\alpha u_\alpha = -1$ or $u^\alpha u_\alpha = 0$ everywhere for a timelike or null congruence). As in section 12.1, we introduce the tensor field (12.6)

$$B_{\alpha\beta} = \nabla_\beta u_\alpha \quad . \tag{12.16}$$

Because u^α is affinely geodesic, $B_{\alpha\beta}$ satisfies

$$B_{\alpha\beta}u^\beta = u^\beta \nabla_\beta u_\alpha = 0 \tag{12.17}$$

and

$$u^\alpha B_{\alpha\beta} = \tfrac{1}{2}\nabla_\beta(u^\alpha u_\alpha) = 0 \tag{12.18}$$

(for a space-time filling congruence $u^\alpha u_\alpha$ is simply a constant). Thus $B_{\alpha\beta}$ is transverse to u^α is the sense that it only has components in the directions transverse to u^α. Its trace

$$\theta = B^\alpha_{\ \alpha} = g^{\alpha\beta}B_{\alpha\beta} = \nabla_\alpha u^\alpha \tag{12.19}$$

is the divergence of u^α and is known as the *expansion* of the (affinely parametrised) geodesic congruence.

The key equation governing the evolution of $B_{\alpha\beta}$ along the integral curves of the geodesic vector field is (12.12)

$$D_\tau B^\alpha_{\ \gamma} + B^\alpha_{\ \beta}B^\beta_{\ \gamma} = R^\alpha_{\ \delta\beta\gamma}u^\beta u^\delta \quad . \tag{12.20}$$

By taking the trace of this equation, we evidently obtain an evolution equation for the expansion θ, namely

$$\frac{d}{d\tau}\theta = -(\nabla_\alpha u_\beta)(\nabla^\beta u^\alpha) - R_{\alpha\beta}u^\alpha u^\beta \quad . \tag{12.21}$$

Note that this equation, written in the form

$$u^\beta\nabla_\beta(\nabla_\alpha u^\alpha) + (\nabla_\alpha u_\beta)(\nabla^\beta u^\alpha) + R_{\alpha\beta}u^\alpha u^\beta = 0 \quad . \tag{12.22}$$

is a special case of the "master equation" (8.64) for $V^\alpha \to u^\alpha$ with $u^\beta\nabla_\beta u^\alpha = 0$.

To gain some more insight into the geometric significance of this equation, we now consider the case that the geodesic congruence u^α is timelike and normalised in the standard way as $u^\alpha u_\alpha = -1$ (so that τ is proper time).

Given this timelike geodesic congruence, we can introduce the tensor

$$h_{\alpha\beta} = g_{\alpha\beta} + u_\alpha u_\beta \quad . \tag{12.23}$$

The properties of this tensor are closely related to those of the (induced metric) tensor $h_{\alpha\beta} = g_{\alpha\beta} - \epsilon N_\alpha N_\beta$ (16.1) studied in section 16.1 in the context of hypersurfaces. The main difference in the present context is that u_α is not necessarily hypersurface-orthogonal (section 15.5) and therefore, in particular, not necessarily a normal vector field to a family of spacelike hypersurfaces. Therefore $h_{\alpha\beta}$ does not necesarily have an interpretation as the induced metric on some hypersurface. Nevertheless, pointwise it can be interpreted as a metric on the space of vectors transverse to the geodesic and its purely algebraic properties are identical to those of the induced metric.

In particular,

- $h_{\alpha\beta}$ has the characteristic property that it is orthogonal to u^α,

$$u^\alpha h_{\alpha\beta} = h_{\alpha\beta}u^\beta = 0 \ . \tag{12.24}$$

- It can therefore be interpreted as the spatial projection of the metric in the directions orthogonal to the timelike vector field u^α. This can be seen more explicitly in terms of the projectors

$$\begin{aligned} h^\alpha_{\ \beta} &= \delta^\alpha_{\ \beta} + u^\alpha u_\beta \\ h^\alpha_{\ \beta} h^\beta_{\ \gamma} &= h^\alpha_{\ \gamma} \ . \end{aligned} \tag{12.25}$$

  On directions tangential to u^α they act as

$$h^\alpha_{\ \beta} u^\beta = 0 \ , \tag{12.26}$$

  whereas on vectors ξ^α orthogonal to u^α, $u_\alpha \xi^\alpha = 0$ (spacelike vectors), one has

$$h^\alpha_{\ \beta} \xi^\beta = \xi^\alpha \ . \tag{12.27}$$

- Thus, acting on an arbitrary vector field V^α, $v^\alpha = h^\alpha_{\ \beta} V^\beta$ is the projection of this vector into the plane orthogonal to u^α. In the same way one can project an arbitrary tensor to a spatial or transverse tensor. E.g.

$$t_{\alpha\ldots\beta} = T_{\gamma\ldots\delta}\, h^\gamma_{\ \alpha} \ldots h^\delta_{\ \beta} \tag{12.28}$$

  satisfies

$$u^\alpha t_{\alpha\ldots\beta} = \ldots = u^\beta t_{\alpha\ldots\beta} = 0 \ . \tag{12.29}$$

- In particular, the projection of the metric is

$$g_{\alpha\beta} \to g_{\gamma\delta} h^\gamma_{\ \alpha} h^\delta_{\ \beta} = g_{\alpha\beta} + u_\alpha u_\beta = h_{\alpha\beta} \ , \tag{12.30}$$

  as anticipated above. Whereas for the space-time metric one obviously has $g^{\alpha\beta} g_{\alpha\beta} = 4$, the trace of $h_{\alpha\beta}$ is (in the 4-dimensional case)

$$g^{\alpha\beta} h_{\alpha\beta} = g^{\alpha\beta} g_{\alpha\beta} + g^{\alpha\beta} u_\alpha u_\beta = 4 - 1 = 3 = h^{\alpha\beta} h_{\alpha\beta} \ . \tag{12.31}$$

Thus for an affinely parametrised congruence the properties (12.17) and (12.18) show that $B_{\alpha\beta}$ is automatically a spatial or transverse tensor in the sense above,

$$b_{\alpha\beta} \equiv h^\gamma_\alpha h^\delta_\beta B_{\gamma\delta} = B_{\alpha\beta} \ . \tag{12.32}$$

Note that the affine parametrisation of the timelike geodesic congruence, expressed by the normalisation condition $u^\alpha u_\alpha = -1$, is crucial for this entire set-up, since the projection operator requires a unit vector field. This is to be contrasted with the situation for null geodesic congruences ℓ^α, to be discussed below, where the property

$\ell^\alpha \ell_\alpha = 0$ is independent of the parametrisation and one can (and we will) also consider the case of non-affine parametrisations.

In the spirit of elasticity theory, we now decompose $b_{\alpha\beta}$ into its anti-symmetric, symmetric traceless and trace part,

$$b_{\alpha\beta} = \omega_{\alpha\beta} + \sigma_{\alpha\beta} + \tfrac{1}{3}\theta h_{\alpha\beta} \quad , \tag{12.33}$$

with

$$\begin{aligned}
\omega_{\alpha\beta} &= \tfrac{1}{2}(b_{\alpha\beta} - b_{\alpha\beta}) \\
\sigma_{\alpha\beta} &= \tfrac{1}{2}(b_{\alpha\beta} + b_{\alpha\beta}) - \tfrac{1}{3}\theta h_{\alpha\beta} \\
\theta &= h^{\alpha\beta} b_{\alpha\beta} = g^{\alpha\beta} B_{\alpha\beta} = \nabla_\alpha u^\alpha \quad .
\end{aligned} \tag{12.34}$$

The quantities $\omega_{\alpha\beta}$, $\sigma_{\alpha\beta}$ and θ are known as the *rotation tensor*, *shear tensor*, and *expansion* of the congruence (family) of geodesics defined by u^α.

In terms of these quantities we can write the evolution equation (12.7) for deviation vectors as

$$D_\tau \xi^\alpha = \omega^\alpha_{\ \beta}\xi^\beta + \sigma^\alpha_{\ \beta}\xi^\beta + \tfrac{1}{3}\theta\xi^\alpha \quad , \tag{12.35}$$

and the evolution equation (12.21) for the expansion θ as

$$\frac{d}{d\tau}\theta = -\tfrac{1}{3}\theta^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta} - R_{\alpha\beta}u^\alpha u^\beta \quad . \tag{12.36}$$

This is the *Raychaudhuri equation* for timelike geodesic congruences.

REMARKS:

1. The expansion θ can be written as

$$\begin{aligned}
\theta &= h^{\alpha\beta} b_{\alpha\beta} = h^{\alpha\beta} B_{\alpha\beta} = h^{\alpha\beta}\nabla_\beta u_\alpha \\
&= \tfrac{1}{2} h^{\alpha\beta}(\nabla_\beta u_\alpha + \nabla_\alpha u_\beta) = \tfrac{1}{2} h^{\alpha\beta} L_u g_{\alpha\beta}
\end{aligned} \tag{12.37}$$

   where L_u denotes the Lie derivative along the vector field u. Substituting $g_{\alpha\beta} = h_{\alpha\beta} - u_\alpha u_\beta$, one finds

$$\theta = \tfrac{1}{2} h^{\alpha\beta} L_u (h_{\alpha\beta} - u_\alpha u_\beta) = \tfrac{1}{2} h^{\alpha\beta} L_u h_{\alpha\beta} \quad . \tag{12.38}$$

   Recalling the formula

$$\delta\sqrt{g} = \sqrt{g} g^{\alpha\beta}\delta g_{\alpha\beta}/2 \tag{12.39}$$

   for the variation of a volume element induced by a variation of the metric, morally speaking the above equation says that θ measures the change of a transverse (cross-sectional) volume of the congruence with volume element $\sqrt{h}$ as one moves along the geodesics,

$$\theta = \frac{1}{\sqrt{h}}\frac{d}{d\tau}\sqrt{h} \quad . \tag{12.40}$$

The statement as such is correct and provides the correct intuition for the meaning of θ, but the definition of the cross-sectional volume and its volume element require a bit of care. When the congruence is hypersurface orthogonal, with the induced metric (16.6)

$$h_{ab} = E_a^\alpha E_b^\beta h_{\alpha\beta} \quad , \tag{12.41}$$

then (12.40) with $h = \det(h_{ab})$ follows from (12.38), because

$$\begin{aligned} h^{ab} L_u h_{ab} &= h^{ab} L_u (E_a^\alpha E_b^\beta h_{\alpha\beta}) \\ &= h^{ab} E_a^\alpha E_b^\beta L_u h_{\alpha\beta} = h^{\alpha\beta} L_u h_{\alpha\beta} \quad . \end{aligned} \tag{12.42}$$

Here we have made use of the fact that u and E_a have vanishing Lie bracket, because (introducing τ and y^a as coordinates, instead of the x^α)

$$u^\alpha = \frac{\partial x^\alpha}{\partial \tau} \quad , \quad E_a^\alpha = \frac{\partial x^\alpha}{\partial y^a} \tag{12.43}$$

and the Lie bracket gives the commutator of the second partical derivatives of x^α.

When the congruence is not hypersurface orthogonal, one can still construct a transverse cross-sectional volume, but one can only choose it to be orthogonal at a given geodesic. Introducing in a neighbourhood of a point on this geodesic coordinates y^a labelling the geodesics, as well as the parameter τ along the geodesic, the above calculation will then still go through.[30]

2. If required and desired, from (12.20) similar (but somewhat less transparent) equations can be derived for the evolution of the shear and rotation tensors along the geodesic congruence, i.e. for $(d/d\tau)\sigma_{\alpha\beta}$ and $(d/d\tau)\omega_{\alpha\beta}$.

3. Since $\omega_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ are purely spatial tensors, their squares are non-negative,

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} \geq 0 \quad , \quad \omega^{\alpha\beta}\omega_{\alpha\beta} \geq 0 \ , \tag{12.44}$$

with $\sigma^{\alpha\beta}\sigma_{\alpha\beta} = 0$ only for $\sigma_{\alpha\beta} = 0$ (and likewise for the rotation). They thus enter the Raychaudhuri equation with opposite signs.

4. In the presence of both these terms it is difficult to say something general about the evolution of θ. Since the first term $(-\theta^2/3)$ is non-positive, an important special case of the Raychaudhuri equation arises when the rotation is zero, $\omega_{\alpha\beta} = 0$. This happens for example when $u_\alpha = \partial_\alpha S$ is the gradient co-vector of some function S. In this case u_α is orthogonal to the level-surfaces of S. In fact, more generally we have the statement that

$$\omega_{\alpha\beta} = 0 \quad \Leftrightarrow \quad u^\alpha \quad \text{hypersurface orthogonal} \ . \tag{12.45}$$

[30] For a more careful proof of this statement see the discussion in section 2.4.8 of E. Poisson, *A Relativist's Toolkit*.

Indeed, assume that u_α is hypersurface orthogonal, i.e. (15.55)

$$u_{[\alpha}\nabla_\beta u_{\gamma]} = 0 \quad \Leftrightarrow \quad \omega_{\alpha\beta}u_\gamma + \omega_{\beta\gamma}u_\alpha + \omega_{\gamma\alpha}u_\beta = 0 \ . \tag{12.46}$$

Contracting this with u^γ and using $u^\gamma u_\gamma = -1$ and $u^\gamma \omega_{\gamma\beta} = 0$, only the first term survives and one finds on the nose that $\omega_{\alpha\beta} = 0$,

$$u_{[\alpha}\nabla_\beta u_{\gamma]} = 0 \quad \Rightarrow \quad \omega_{\alpha\beta} = 0 \ , \tag{12.47}$$

and the Frobenius theorem provides one with the converse statement. Alternatively, $\omega_{\alpha\beta} = 0$ follows from assuming that u_α has the explicit hypersurface-orthogonal form $u_\alpha = f\partial_\alpha S$. Then one has (15.52)

$$\omega_{\beta\alpha} = \nabla_{[\alpha}u_{\beta]} = (\nabla_{[\alpha}\log f)u_{\beta]} \tag{12.48}$$

and by contraction with u^α one deduces

$$\partial_\alpha f = -(u^\beta\partial_\beta f)u_\alpha \sim u_\alpha \quad \Rightarrow \qquad \omega_{\alpha\beta} = 0 \ . \tag{12.49}$$

5. Either way, for a hypersurface orthogonal congruence of timelike geodesics one has

$$\frac{d}{d\tau}\theta = -\tfrac{1}{3}\theta^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} - R_{\alpha\beta}u^\alpha u^\beta \ . \tag{12.50}$$

   The first two terms on the right hand side are manifestly non-positive (recall that $\sigma_{\alpha\beta}$ is a spatial tensor and hence $\sigma_{\alpha\beta}\sigma^{\alpha\beta} \geq 0$). Thus, if one assumes that the geometry is such that

$$R_{\alpha\beta}u^\alpha u^\beta \geq 0 \tag{12.51}$$

   (by the Einstein equations to be discussed in the section 19, this translates into a positivity condition on the energy-momentum tensor known as the *strong energy condition*, cf. section 22.1), one finds

$$\frac{d}{d\tau}\theta = -\tfrac{1}{3}\theta^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} - R_{\alpha\beta}u^\alpha u^\beta \leq 0 \ . \tag{12.52}$$

   This means that the divergence (convergence) of geodesics will decrease (increase) in time. The interpretation of this result is that gravity is an *attractive* force (for matter satisfying the strong energy condition) whose effect is to focus geodesics.

6. According to (12.52), $d\theta/d\tau$ is not only negative but actually bounded from above by

$$\frac{d}{d\tau}\theta \leq -\tfrac{1}{3}\theta^2 \ . \tag{12.53}$$

   Rewriting this equation as

$$\frac{d}{d\tau}\frac{1}{\theta} \geq \frac{1}{3} \ , \tag{12.54}$$

   one deduces immediately that

$$\frac{1}{\theta(\tau)} \geq \frac{1}{\theta(0)} + \frac{\tau}{3} \ . \tag{12.55}$$

This has the rather dramatic implication that, if $\theta(0) < 0$ (i.e. the geodesics are initially converging), then $\theta(\tau) \to -\infty$ within finite proper time $\tau \leq 3/|\theta(0)|$,

$$\theta(\tau) \to -\infty \quad \text{for} \quad \tau \leq 3/|\theta(0)| \tag{12.56}$$

provided that the geodesics can be extended that far.

7. If one thinks of the geodesics as trajectories of physical particles, this is obviously a rather catastrophic situation in which these particles will be infinitely squashed. In general, however, the divergence of θ only indicates that the family of geodesics develops what is known as a *caustic* where different geodesics meet.

8. Simple non-catastrophic examples of such caustics are e.g. the poles of a sphere where great circles meet, or even just the origin in Euclidean space $\mathbb{R}^n$ when considering the family of radial geodesics passing through the origin. E.g. in the latter case the tangent vector field is simply ∂_r, and its divergence is

$$\nabla_\alpha(\partial_r)^\alpha = \frac{1}{\sqrt{g}}\partial_\alpha(\sqrt{g}(\partial_r)^\alpha) \sim r^{-1} \quad , \tag{12.57}$$

which diverges as $r \to 0$. This divergent behaviour is strictly related to the breakdown at the origin $r = 0$ of spherical coordinates adapted to this congruence (cf. also the related discussion in section 43.6).

9. Nevertheless, the above result plays a crucial role in establishing the occurrence of true singularities in general relativity if supplemented e.g. by conditions which ensure that such "harmless" caustics cannot appear, as this means that the geodesic cannot be extended to where one would find $\theta \to -\infty$. This kind of argument (leading to the conclusion of *geodesic incompleteness* of a space-time) is one of the typical ingredients of the singularity theorems of general relativity (see footnote 98 of section 29.3 for some references).

10. The adaptation of this formalism in general and the Raychaudhuri equation in particular to congruences of null geodesics requires some more care (and is ultimately expressed in terms of 2-dimensional rather than 3-dimensional spatial tensors), and we will discuss this in section 12.4 below.

12.3 Transverse Null Geodesic Deviation Equation

In section 12.4 we will derive the null counterpart of the Raychaudhuri equation for timelike geodesic congruences discussed in section 12.2 above. The set-up we will use is a suitable combination of that for timelike geodesics and the formalism of projectors adapted to null directions. As a preparation for this, and a useful by-product, in this section we will first derive a variant of the geodesic deviation equation for null geodesics, the *transverse null geodesic deviation equation*.

Thus we consider a null geodesic (or congruence of null geodesics), with tangent vector field ℓ^α, and we will initially choose these null geodesics to be affinely parametrised so that one has

$$\ell^\alpha \ell_\alpha = 0 \quad , \quad \ell^\alpha \nabla_\alpha \ell^\beta = 0 \ . \tag{12.58}$$

The affine parameter along the null geodesics of this congruence will (for lack of imagination) be called τ.

Now recall from the discussion of the geodesic deviation equation in section 12.1 that for any geodesic deviation vector ξ^α, i.e. a vector satisfying the condition

$$D_\tau \xi^\alpha = \xi^\beta \nabla_\beta u^\alpha \ , \tag{12.59}$$

the quantity $\xi^\alpha u_\alpha$ is constant along the geodesic,

$$\frac{d}{d\tau}(\xi^\alpha u_\alpha) = (D_\tau \xi^\alpha) u_\alpha = \xi^\beta u^\alpha \nabla_\beta u^\alpha = 0 \tag{12.60}$$

because $u^\alpha u_\alpha = \epsilon$ is constant, regardless of whether u^α is timelike or null. However, the interpretation of this and its implications depend on whether one is dealing with a timelike geodesic (congruence) or a null geodesic (congruence):

1. In the timelike case the quantity $(-\xi^\alpha u_\alpha)$ is the uninteresting component of ξ along the geodesic u^α, and there was clearly no point in not setting it to zero. The transversality condition $\xi^\alpha u_\alpha = 0$ on the deviation vector could be consistently imposed and was sufficient to remove this component.

2. In the null case, however, the condition $\xi^\alpha \ell_\alpha = 0$ does not accomplish this, i.e. does not remove the component of ξ pointing in the direction of ℓ^α because it imposes no condition precisely on that component. Thus we expect the deviation vector ξ to have *two* uninteresting components in the null case, $\xi^\alpha \ell_\alpha$ and the component of ξ in the direction of ℓ^α:

   - Indeed, as recalled above, we already know in general that $\xi^\alpha \ell_\alpha$ is constant,

   $$\frac{d}{d\tau}(\xi^\alpha \ell_\alpha) = 0 \ . \tag{12.61}$$

   - To verify this also for a component of ξ along ℓ^α, we impose the condition that ξ be a deviation vector on the vector

   $$\xi^\alpha = \xi \ell^\alpha \tag{12.62}$$

   (note once again that this ξ is completely unrelated to $\xi^\alpha \ell_\alpha$). Since ℓ_α is a geodesic, we have

   $$D_\tau \xi^\alpha = (\frac{d}{d\tau}\xi)\ell^\alpha \stackrel{!}{=} \xi^\beta \nabla_\beta \ell^\alpha = \xi \ell^\beta \nabla_\beta \ell^\alpha = 0 \ . \tag{12.63}$$

   Thus $\xi^\alpha = \xi \ell^\alpha$ is a (boring non-) deviation vector iff ξ is constant, and one should impose the condition $\xi = 0$.

Therefore it is natural to project out both these components from ξ_α. In order to construct a suitable projection operator, one can proceed as in section 17.4 and introduce a complementary null vector (field) n^α with

$$n^\alpha n_\alpha = 0 \quad , \quad n^\alpha \ell_\alpha = -1 \ . \tag{12.64}$$

Then

$$\xi^\alpha = \xi \ell^\alpha + \dots \quad \Rightarrow \quad \xi = -\xi^\alpha n_\alpha \tag{12.65}$$

and we can elininate both boring components by imposing the transversality conditions

$$\xi^\alpha \ell_\alpha = \xi^\alpha n_\alpha = 0 \tag{12.66}$$

on the deviation vector ξ^α, in addition to the deviation vector condition (12.3),

$$\ell^\alpha \nabla_\alpha \xi^\beta = \xi^\alpha \nabla_\alpha \ell^\beta \ . \tag{12.67}$$

As in (12.6) we introduce

$$B_{\alpha\beta} = \nabla_\beta \ell_\alpha \ , \tag{12.68}$$

so that the above condition that ξ^α is a deviation vector can be written as

$$D_\tau \xi^\alpha = B^\alpha_{\ \beta} \xi^\beta \ . \tag{12.69}$$

Because ℓ^α is null and (affinely) geodesic, one has

$$\ell^\alpha B_{\alpha\beta} = B_{\alpha\beta} \ell^\beta = 0 \ , \tag{12.70}$$

but $B_{\alpha\beta}$ is not automatically orthogonal to n^α (and we will come back to and rectify this below). Exactly the same calculation as (12.12) in section 12.1 now shows that

$$D_\tau B_{\alpha\beta} + B^\alpha_{\ \gamma} B^\gamma_{\ \beta} = R^\alpha_{\ \delta\gamma\beta} \ell^\gamma \ell^\delta \tag{12.71}$$

so that one also has the null counterpart of (12.13), namely

$$(D_\tau)^2 \xi^\alpha = R^\alpha_{\ \delta\gamma\beta} \ell^\gamma \ell^\delta \xi^\beta \ . \tag{12.72}$$

Again this equation is automatically transverse to ℓ, in the sense that

$$\ell_\alpha (D_\tau)^2 \xi^\alpha = R_{\alpha\delta\gamma\beta} \ell^\alpha \ell^\gamma \ell^\delta \xi^\beta = 0 \ , \tag{12.73}$$

but not transverse to n^α.

In order to pick up only the transverse components of $D_\tau \xi^\alpha$ in (12.69) (and in subsequent equations), we thus need to project $B_{\alpha\beta}$ onto its transverse components. The construction of this projector is identical to that in section 17.4:

- Associated with a choice of n^α we have a decomposition of the metric into a longitudinal and a transverse spatial part,

$$g_{\alpha\beta} = s_{\alpha\beta} - (\ell_\alpha n_\beta + \ell_\beta n_\alpha) \quad , \tag{12.74}$$

with the properties

$$s_{\alpha\beta}\ell^\beta = s_{\alpha\beta}n^\beta = 0 \tag{12.75}$$

and

$$g^{\alpha\beta}s_{\alpha\beta} = s^{\alpha\beta}s_{\alpha\beta} = s^{\alpha}_{\ \alpha} = 2 \quad . \tag{12.76}$$

(we are considering the case $D = 4$).

- We then also have the corresponding transverse (spatial) projectors

$$\begin{aligned} s^\alpha_{\ \beta} = \delta^\alpha_{\ \beta} + (\ell^\alpha n_\beta + n^\alpha \ell_\beta): \quad & s^\alpha_{\ \gamma}s^\gamma_{\ \beta} = s^\alpha_{\ \beta} \\ & s^\alpha_{\ \beta}\ell^\beta = s^\alpha_{\ \beta}n^\beta = 0 \quad , \end{aligned} \tag{12.77}$$

With the aid of these projectors, we can now write the fully projected version of (12.69) as

$$s^\alpha_{\ \beta}D_\tau(s^\beta_{\ \gamma}\xi^\gamma) = b^\alpha_{\ \beta}\xi^\beta \tag{12.78}$$

where $b_{\alpha\beta}$ is the projection of $B_{\alpha\beta}$,

$$b_{\alpha\beta} = s_\alpha^{\ \gamma}s_\beta^{\ \delta}B_{\gamma\delta} \quad . \tag{12.79}$$

Likewise the purely transverse (to ℓ and n) variant of the null geodesic equation (12.72) can be written as

$$s^\alpha_{\ \beta}(D_\tau)^2(s^\beta_{\ \gamma}\xi^\gamma) = s^\alpha_{\ \beta}s^\gamma_{\ \delta}R^\beta_{\ \mu\nu\gamma}\ell^\mu\ell^\nu\xi^\delta \quad . \tag{12.80}$$

While this is essentially the final result, it is not particularly transparent yet. We will put this equation into a somewhat more attractive form below, in which manifestly only the transverse components of the deviation vector and $R^\beta_{\ \mu\nu\gamma}\ell^\mu\ell^\nu$ appear.

First of all, note that the auxiliary normal vector n^α is not unique. For a fixed choice of ℓ^α, at a point on the geodesic, that is for a given value of τ, it is uniquely determined up to null rotations around ℓ^α,

$$\ell \to \ell \quad , \quad n \to n + \beta^a E_a + \tfrac{1}{2}\beta^2 \ell \quad , \quad E_a \to E_a + \beta_a \ell \quad , \tag{12.81}$$

where E_a are (necessarily spatial) vectors orthogonal to ℓ and n, and $\beta_a = \beta_a(\tau)$. This ambiguity does not affect any of the results in this section, so in principle one can make any choice of auxiliary n^α. In particular, one can choose n^α to be parallel-transported along ℓ^α,

$$\ell^\alpha \nabla_\alpha n^\beta = 0 \quad . \tag{12.82}$$

Then the properties of parallel transport obtained in section 5.8 imply that the conditions (12.64) on n^α hold everywhere along the null geodesic (or congruence of null geodesics) if they are satisified initially. This reduces the ambiguity in (12.81) to τ-independent null rotations.

In fact, one can do even better than that and choose (see also the discussion at the end of section 17.4, in particular around (17.57)) an entire pseudo-orthonormal frame

$$\{E_A\} = \{E_+ = \ell, E_- = n, E_a\} : \quad g_{\alpha\beta}E^\alpha_A E^\beta_B = \eta_{AB} \tag{12.83}$$

where

$$\eta_{++} = \eta_{--} = 0 \quad , \quad \eta_{+-} = -1 \quad , \quad \eta_{a+} = \eta_{a-} = 0 \quad , \quad \eta_{ab} = \delta_{ab} \quad . \tag{12.84}$$

If one selects such a frame at one point along the geodesic and then parallel transports the frame along the geodesic, the orthogonality relations (12.83) will hold everywhere along the geodesic. Thus we can always choose a basis E_A such that

$$D_\tau E^\alpha_A = 0 \quad , \quad g_{\alpha\beta}E^\alpha_A E^\beta_B = \eta_{AB} \quad . \tag{12.85}$$

With this choice, a transverse geodesic deviation vector is simply one which has components only in the E_a-directions,

$$\xi^\alpha \ell_\alpha = \xi^\alpha n_\alpha = 0 \quad \Leftrightarrow \quad \xi^\alpha = \xi^a E^\alpha_a \quad , \tag{12.86}$$

or simply

$$\xi = \xi^a E_a \quad . \tag{12.87}$$

Since $D_\tau E_a$=0, one has

$$D_\tau E_a = 0 \quad \Rightarrow \quad D_\tau \xi^a = \frac{d}{d\tau}\xi^a \quad , \tag{12.88}$$

so that with respect to this basis covariant differentiation along the curve reduces to ordinary differentiation of the components. Then the null geodesic deviation equation (12.80) can simply be written as

$$\frac{d^2}{d\tau^2}\xi^a = R^a{}_{++b}\xi^b = -R^a{}_{+b+}\xi^b \quad , \tag{12.89}$$

where the R_{a+b+} are the frame components of the Riemann tensor,

$$R_{a+b+} = E^\alpha_a E^\mu_+ E^\beta_b E^\nu_+ R_{\alpha\mu\beta\nu} = E^\alpha_a E^\beta_b R_{\alpha\mu\beta\nu}\ell^\mu \ell^\nu \quad . \tag{12.90}$$

Thus the transverse null geodesic deviation equation has the form of a $(D-2)$-dimensional (transverse) harmonic oscillator equation,

$$\frac{d^2}{d\tau^2}\xi^a = -(\Omega^2)^a{}_b \xi^b \quad , \tag{12.91}$$

with the time-dependent symmetric frequency matrix

$$(\Omega^2)(\tau)^a{}_b = R^a{}_{+b+}(\tau) \quad . \tag{12.92}$$

The notation used here is perhaps suggestive but it is not meant to imply that Ω^2 is necessarily positive - the frequencies can be real or imaginary. Using the decomposition (11.76) of the Riemann tensor into its traceless and trace parts, we can (with $D = 4$, $\eta_{a+} = \eta_{++} = 0, \eta_{ab} = \delta_{ab}$) decompose R_{a+b+} as

$$R_{a+b+} = C_{a+b+} + \tfrac{1}{2}\delta_{ab}R_{++} \quad . \tag{12.93}$$

In particular, if the Ricci tensor is zero (as we will see this means that the metric solves the vacuum Einstein equations), the frequency matrix Ω^2 is symmetric traceless and thus necesarily has positive and negative eigenvalues (corresponding to real and imaginary frequencies).

12.4 Raychaudhuri Equation for Affine Null Geodesic Congruences

We now consider a null geodesic congruence, with tangent vector field again denoted by ℓ^α, and we will initially choose these null geodesics to be affinely parametrised so that one has

$$\ell^\alpha \ell_\alpha = 0 \quad , \quad \ell^\alpha \nabla_\alpha \ell^\beta = 0 \quad . \tag{12.94}$$

We use the same framework as in the previous setion, with an auxiliary null vector field n^α with $\ell^\alpha n_\alpha = -1$, the associated projectors etc.

In (12.79) we introduced the transverse projection $b_{\alpha\beta}$ of the tensor $B_{\alpha\beta} = \nabla_\beta \ell_\alpha$,

$$b_{\alpha\beta} = s_\alpha^{\ \gamma} s_\beta^{\ \delta} B_{\gamma\delta} \quad . \tag{12.95}$$

Performing this projection explicitly, one sees that this spatial projection $b_{\alpha\beta}$ is equal to

$$b_{\alpha\beta} = s_\alpha^{\ \gamma} s_\beta^{\ \delta} B_{\gamma\delta} = B_{\alpha\beta} + \ell_\alpha n^\gamma B_{\gamma\beta} + \ell_\beta n^\delta B_{\alpha\delta} + \ell_\alpha \ell_\beta n^\gamma n^\delta B_{\gamma\delta} \quad . \tag{12.96}$$

This has two useful immediate consequences that we will make use of in the following, namely

- that the spatial trace of $b_{\alpha\beta}$ with respect to $s_{\alpha\beta}$ is equal to the space-time trace of $B_{\alpha\beta}$ (with respect to $g_{\alpha\beta}$),

$$g^{\alpha\beta} B_{\alpha\beta} = g^{\alpha\beta} b_{\alpha\beta} = s^{\alpha\beta} b_{\alpha\beta} \quad , \tag{12.97}$$

- and that the square of $b_{\alpha\beta}$ is identical to that of $B_{\alpha\beta}$,

$$B^{\alpha\beta} B_{\beta\alpha} = b^{\alpha\beta} b_{\beta\alpha} \quad . \tag{12.98}$$

We can now, as in the timelike case, decompose $b_{\alpha\beta}$ orthogonally into its irreducible (trace, symmetric traceless, anti-symmetric) parts,

$$\begin{aligned} b_{\alpha\beta} &= \tfrac{1}{2}\theta_\ell s_{\alpha\beta} + \tfrac{1}{2}(b_{\alpha\beta} + b_{\beta\alpha} - \theta_\ell s_{\alpha\beta}) + \tfrac{1}{2}(b_{\alpha\beta} - b_{\beta\alpha}) \\ &= \tfrac{1}{2}\theta_\ell s_{\alpha\beta} + \sigma_{\alpha\beta} + \omega_{\alpha\beta} \quad . \end{aligned} \tag{12.99}$$

Here θ_ℓ is the *expansion*

$$\theta_\ell = s^{\alpha\beta}b_{\alpha\beta} = s^{\alpha\beta}\nabla_\beta \ell_\alpha = g^{\alpha\beta}\nabla_\alpha \ell_\beta = \nabla^\alpha \ell_\alpha \quad , \tag{12.100}$$

REMARKS:

1. As in the timelike case (12.38), the expansion θ_ℓ can be written as

$$\theta_\ell = \tfrac{1}{2} s^{\alpha\beta} L_\ell s_{\alpha\beta} \quad . \tag{12.101}$$

   and leads to an analogous interpretation of θ_ℓ as measuring the change in the cross-sectional area element $\sqrt{s}$ of the congruence (12.40),

$$\theta_\ell = \frac{1}{\sqrt{s}} L_\ell \sqrt{s} \quad . \tag{12.102}$$

2. The equivalence between the spatial and space-time traces of $\nabla_\alpha \ell_\beta$ in the above equation is due to the fact that we have chosen ℓ to be affinely parametrised. We will always define θ to be the spatial trace (divergence) of $\nabla_\alpha \ell_\beta$, even when ℓ is not affinely parametrised, but in that case θ and $\nabla_\alpha \ell^\alpha$ are no longer equal (see (12.126)). We will return to this issue below.

3. As regards the other terms, $\sigma_{\alpha\beta}$ and $\omega_{\alpha\beta}$ are again known as the *shear tensor* and *rotation tensor* respectively.

4. As in the timelike case, the rotation is zero if ℓ^α is hypersurface orthogonal. We will establish this result below.

5. Because the above decomposition is orthogonal, we have

$$B^{\alpha\beta}B_{\beta\alpha} = b^{\alpha\beta}b_{\beta\alpha} = +\tfrac{1}{2}\theta_\ell^2 + \sigma^{\alpha\beta}\sigma_{\alpha\beta} - \omega^{\alpha\beta}\omega_{\alpha\beta} \quad . \tag{12.103}$$

6. Because the tensors appearing on the right-hand side of this equation are spatial tensors, their squares are non-negative,

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} \geq 0 \quad , \quad \omega^{\alpha\beta}\omega_{\alpha\beta} \geq 0 \quad . \tag{12.104}$$

We now want to determine

$$\frac{d}{d\tau}\theta_\ell = \ell^\alpha \nabla_\alpha \theta_\ell \quad . \tag{12.105}$$

We can do this either by again deriving an evolution equation for all of $b_{\alpha\beta}$, as in (12.71), or by calculating directly the derivative of θ_ℓ along ℓ. Adopting the latter procedure here, just following one's tensor calculus nose one finds

$$\begin{aligned}\frac{d}{d\tau}\theta_\ell &= \ell^\alpha \nabla_\alpha (\nabla_\beta \ell^\beta) \\ &= \ell^\alpha [\nabla_\alpha, \nabla_\beta]\ell^\beta + \ell^\alpha \nabla_\beta \nabla_\alpha \ell^\beta \\ &= -R_{\alpha\beta}\ell^\alpha \ell^\beta - (\nabla^\beta \ell^\alpha)\nabla_\alpha \ell_\beta + \nabla_\beta(\ell^\alpha \nabla_\alpha \ell^\beta) \quad .\end{aligned} \tag{12.106}$$

The 2nd term is just $-B^{\alpha\beta}B_{\beta\alpha}$ and the 3rd term is zero because ℓ is geodesic. Thus one finds the *Raychaudhuri equation for null congruences*

$$\frac{d}{d\tau}\theta_\ell = -R_{\alpha\beta}\ell^\alpha\ell^\beta - \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta} \ . \tag{12.107}$$

Using (12.102) in the form

$$\frac{d}{d\tau}\sqrt{s} = \theta_\ell\sqrt{s} \ , \tag{12.108}$$

we can also write this as an equation for the change in the expansion rate of the cross-sectional area $\sqrt{s}$ of the congruence. This leads to an additional $+\theta_\ell^2$ in the evolution equation, and thus flips the sign of the 2nd term of (12.107), resulting in

$$\frac{d^2}{d\tau^2}\sqrt{s} = \left(-R_{\alpha\beta}\ell^\alpha\ell^\beta + \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta}\right)\sqrt{s} \ . \tag{12.109}$$

REMARKS:

1. If the geometry is such that

$$R_{\alpha\beta}\ell^\alpha\ell^\beta \geq 0 \tag{12.110}$$

(by the Einstein equations to be discussed in the section 19, this translates into a positivity condition on the energy-momentum tensor known as the *null energy condition*, cf. section 22.1), this gives a negative contribution to the right-hand side of the Raychaudhuri equation (indicating the focussing effect of gravity on lightrays).

2. If moreover the rotation $\omega_{\alpha\beta}$ is zero, for an affinely parametrised null congruence one has

$$\frac{d}{d\tau}\theta_\ell = -R_{\alpha\beta}\ell^\alpha\ell^\beta - \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} \tag{12.111}$$

and therefore, in particular,

$$\frac{d}{d\tau}\theta_\ell \leq -\tfrac{1}{2}\theta_\ell^2 \leq 0 \ . \tag{12.112}$$

In Minkowski space-time (with zero curvature) we would have an equality instead of the first inequality sign, so (12.112) is the (reasonable) statement that the expansion is smaller/slower than in Minkowski space, i.e. that (reasonable) matter has the tendency to focus lightrays.

3. Analogously to the timelike case, (12.112) has the consequence that if one has an initially converging null congruence, $\theta_\ell(\tau_0) < 0$, then because of

$$\frac{d}{d\tau}\theta_\ell \leq -\tfrac{1}{2}\theta_\ell^2 \quad \Rightarrow \quad \frac{1}{\theta_\ell(\tau)} \geq \frac{1}{\theta_\ell(\tau_0)} + \frac{\tau-\tau_0}{2} \tag{12.113}$$

$1/\theta(\tau) \to 0_-$ or $\theta(\tau) \to -\infty$ at the latest at

$$\theta_\ell(\tau) \to -\infty \quad \text{for} \quad \tau \leq \tau_0 + 2/|\theta_\ell(\tau_0)| \tag{12.114}$$

(if the geodesics can be extended that far). As in the timelike case, this usually indicates the formation of a (harmless) *caustic* where these null geodesics cross.

4. E.g. in Minkowski space radially outgoing lightrays $\ell = \partial_v$, $v = t + r$, have expansion

$$\theta_\ell = \nabla_\alpha(\partial_v)^\alpha = \frac{1}{r^2}\partial_\alpha(r^2(\partial_t + \partial_r)^\alpha) = +\frac{2}{r} > 0 \ , \tag{12.115}$$

while radially ingoing lightrays $n = \partial_u$, $u = t - r$, have expansion

$$\theta_n = \nabla_\alpha(\partial_u)^\alpha = \frac{1}{r^2}\partial_\alpha(r^2(\partial_t - \partial_r)^\alpha) = -\frac{2}{r} < 0 \ , \tag{12.116}$$

reflecting the fact that outgoing lightrays expand while ingoing lightrays contract. Both expansions diverge as $r \to 0$, but this evidently does not indicate a pathology of Minkowski space-time, but only of the chosen congruences at the origin (where all the lightrays of the congruence meet).

Note also that e.g. θ_ℓ satisfies the Raychaudhuri equation

$$\frac{d}{d\tau}\theta_\ell = \ell^\alpha\nabla_\alpha\theta_\ell = \partial_r(+2/r) = -2/r^2 = -\tfrac{1}{2}\theta_\ell^2 \ , \tag{12.117}$$

as it should (no curvature, no shear, no rotation). With cross-sectional area $\sqrt{s} \sim r^2$, one also has

$$\frac{1}{\sqrt{s}}\frac{d^2}{d\tau^2}\sqrt{s} = \frac{1}{r^2}(\partial_r)^2 r^2 = \frac{2}{r^2} = +\tfrac{1}{2}\theta_\ell^2 \ , \tag{12.118}$$

illustrating the variant (12.109) of the Raychaudhuri equation: while $\theta_\ell \sim r^{-1} \to 0$ as $r \to \infty$, indicating that the cross-sectional spheres become flatter and flatter for large r, the cross-sectional area grows like $\sqrt{s} \sim r^2$, leading to an acceleration of its growth.

5. An argument similar to that in the timelike case shows that the rotation vanishes if (and locally only if, by Frobenius) ℓ^α is hypersurface orthogonal

$$\omega_{\alpha\beta} = 0 \quad \Leftrightarrow \quad \ell^\alpha \quad \text{hypersurface-orthogonal} \ . \tag{12.119}$$

In order to establish this we begin with the condition

$$\ell_{[\alpha}\nabla_\beta\ell_{\gamma]} = 0 \quad \Leftrightarrow \quad B_{[\alpha\beta]}\ell_\gamma + B_{[\beta\gamma]}\ell_\alpha + B_{[\gamma\alpha]}\ell_\beta = 0 \ . \tag{12.120}$$

Contracting this with ℓ^γ, we would get $0 = 0$, which is true but unenlightning. To get something non-trivial and more useful, we contract with n^γ instead. Then we deduce

$$B_{[\alpha\beta]} - B_{[\beta\gamma]}n^\gamma\ell_\alpha - B_{[\gamma\alpha]}n^\gamma\ell_\beta = 0 \ . \tag{12.121}$$

Disentangling this, one sees that this is precisely the statement that the anti-symmetric part of $b_{\alpha\beta}$ is zero, i.e. that $\omega_{\alpha\beta} = 0$, so we have established the desired result

$$\ell_{[\alpha}\nabla_\beta\ell_{\gamma]} = 0 \quad \Rightarrow \quad \omega_{\alpha\beta} = 0 \ . \tag{12.122}$$

6. The expansion properties of families of null geodesics play a crucial role both in the singularity theorems of general relativity (where for example so-called *trapped surfaces* are characterised by negative expansions for both ingoing and "outgoing" families of lightrays), and in the study of black holes and the laws governing the evolution of their event horizons (where the interest is in the null geodesic congruences generating the horizon). In particular, in the latter case the Raychaudhuri equation is one crucial ingredient in the proof of the statement (Hawking's theorem) that under reasonable conditions the cross-sectional area of the event horizon of a black hole cannot decrease.

12.5 Raychaudhuri Equation for Non-affinely Parametrised Null Geodesics

Let us now look at the case when the null geodesic congruence is not affinely parametrised, i.e. when, instead of (12.94), the starting point is a null vector field ℓ^α satisfying

$$\ell^\alpha \ell_\alpha = 0 \quad , \quad \ell^\alpha \nabla_\alpha \ell^\beta = \kappa_\ell \ell^\beta \quad , \tag{12.123}$$

with κ_ℓ the inaffinity. Then a couple of things change in the derivation, but the end result (12.129) turns out to differ from (12.107) by only one term (and I will give an alternative and much quicker derivation of the result below).

As before, we can choose an auxiliary null vector field n^α, construct the projectors $s^\alpha_{\ \beta}$ etc. Defining again $B_{\alpha\beta} = \nabla_\beta \ell_\alpha$, one still has $\ell^\alpha B_{\alpha\beta} = 0$ (because this is implied by $\ell^\alpha \ell_\alpha = 0$), but instead of $B_{\alpha\beta}\ell^\beta = 0$ one now has

$$B_{\alpha\beta}\ell^\beta = \kappa_\ell \ell_\alpha \quad . \tag{12.124}$$

While the projection (12.96) remains unchanged, i.e. the relation between $b_{\alpha\beta}$ and $B_{\alpha\beta}$ has the same form as in (12.96), the equations (12.97) and (12.98) for the trace and square of $B_{\alpha\beta}$ differ. Instead of (12.97) one has

$$\begin{aligned} s^{\alpha\beta} b_{\alpha\beta} &= (g^{\alpha\beta} + \ell^\alpha n^\beta + \ell^\beta n^\alpha) B_{\alpha\beta} \\ &= \nabla_\alpha \ell^\alpha + \kappa_\ell n^\alpha \ell_\alpha = \nabla_\alpha \ell^\alpha - \kappa_\ell \quad . \end{aligned} \tag{12.125}$$

Thus, if we define the expansion θ_ℓ as the spatial trace of $b_{\alpha\beta}$, instead of (12.100) we find

$$\nabla_\alpha \ell^\alpha = \theta_\ell + \kappa_\ell \quad . \tag{12.126}$$

Analogously, for the square of $B_{\alpha\beta}$ one finds, instead of (12.98),

$$B^{\alpha\beta} B_{\beta\alpha} = b^{\alpha\beta} b_{\beta\alpha} + \kappa_\ell^2 \quad . \tag{12.127}$$

Putting everything together and calculating $(d/d\tau)\nabla_\alpha \ell^\alpha$ as in (12.106), one then finds

$$\begin{aligned}\frac{d}{d\tau}(\theta_\ell + \kappa_\ell) &= -R_{\alpha\beta}\ell^\alpha\ell^\beta - (\nabla^\beta \ell^\alpha)\nabla_\alpha \ell_\beta + \nabla_\beta(\ell^\alpha \nabla_\alpha \ell^\beta) \\ &= -R_{\alpha\beta}\ell^\alpha\ell^\beta - B_{\alpha\beta}B^{\beta\alpha} + \nabla_\beta(\kappa_\ell \ell^\beta) \\ &= -R_{\alpha\beta}\ell^\alpha\ell^\beta - b_{\alpha\beta}b^{\beta\alpha} - \kappa_\ell^2 + \frac{d}{d\tau}\kappa_\ell + \kappa_\ell(\theta_\ell + \kappa_\ell) \\ &= -R_{\alpha\beta}\ell^\alpha\ell^\beta - b_{\alpha\beta}b^{\beta\alpha} + \frac{d}{d\tau}\kappa_\ell + \kappa_\ell\theta_\ell \quad .\end{aligned} \tag{12.128}$$

Thus the net effect of dealing with a non-affinely parametrised null congruence is that one just picks up one additional term on the right-hand side of the Raychaudhuri equation,

$$\frac{d}{d\tau}\theta_\ell = \kappa_\ell\theta_\ell - R_{\alpha\beta}\ell^\alpha\ell^\beta - \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta} \quad . \tag{12.129}$$

A quick(er) way to derive (12.129) is from the result (12.107) for affinely parametrised null geodesics, by determining how the quantities appearing in (12.107) change under a reparametrisation

$$\ell^\alpha \to \tilde{\ell}^\alpha = f\ell^\alpha \tag{12.130}$$

(with ℓ^α affinely parametrised, say).

- On the one hand, for the inaffinity one has

$$\tilde{\ell}^\alpha \nabla_\alpha \tilde{\ell}^\beta = \left(\frac{d}{d\tilde{\tau}}\log|f|\right)\tilde{\ell}^\beta \equiv \kappa_{\tilde{\ell}}\,\tilde{\ell}^\beta \quad , \tag{12.131}$$

  where $\tilde{\tau}$ is the non-affine parameter along ℓ^α,

$$\frac{d}{d\tilde{\tau}}f = \tilde{\ell}^\alpha\nabla_\alpha f \tag{12.132}$$

- On the other hand for the expansion parameter etc one deduces from

$$\tilde{B}_{\alpha\beta} = \nabla_\beta \tilde{\ell}_\alpha = fB_{\alpha\beta} + \ell_\alpha\nabla_\beta f \tag{12.133}$$

  that the transverse projection is simply

$$\tilde{b}_{\alpha\beta} = fb_{\alpha\beta} \tag{12.134}$$

  which implies

$$\tilde{b}_{\alpha\beta} = fb_{\alpha\beta} \quad \Rightarrow \quad \left\{\begin{array}{ll}\theta_{\tilde{\ell}} & = f\theta_\ell \\ \tilde{\sigma}_{\alpha\beta} & = f\sigma_{\alpha\beta} \\ \tilde{\omega}_{\alpha\beta} & = f\omega_{\alpha\beta}\end{array}\right. \tag{12.135}$$

Plugging these results into (12.107) one obtains on the nose (12.129) (with $\tau \to \tilde{\tau}, \ell \to \tilde{\ell}$ etc.).

12.6 Expansions and Inaffinities of Radial Null Congruences

In this section, we look at some general properties of radial null congruences in a spherically symmetric space-time. All of the results of the previous sections 12.4 and 12.5 are of course also valid in this case, but the spherically symmetric case also has some special and simplifying features.

Thus we consider a spherically symmetric metric. Such a metric could always be written in the form

$$ds^2 = -A(t,r)dt^2 + B(t,r)dr^2 + r^2 d\Omega^2 \tag{12.136}$$

by a suitable choice of coordinates. However, we will not need to commit ourselves to this particular choice of coordinates. By making an arbitrary coordinate transformation

$$(t,r) \to z^a(t,r) \tag{12.137}$$

preserving the manifest spherical symmetry, this metric can be written in the form

$$ds^2 = g_{ab}(z)dz^a dz^b + r(z)^2 d\Omega^2 \tag{12.138}$$

for some 2-dimensional Lorentzian metric $g_{ab}(z)$, and with $r = r(z^a)$ now a function of the new coordinates.

We now consider two linearly independent radial and spherically symmetric null vector fields ℓ^α and n^α, which we choose to be cross normalised such that

$$\ell^\alpha \ell_\alpha = n^\alpha n_\alpha = 0 \quad , \quad \ell^\alpha n_\alpha = -1 \ . \tag{12.139}$$

Remarks:

1. Here "radial" means that it has components only in the ∂_{z^a}-directions transverse to the sphere, and "spherically symmetric" that the coefficients only depend on the z^a and not on the coordinates of the sphere (this can of course also, if desired, be phrased in a more coordinate-independent way, e.g. as the statement that the Lie derivatives of ℓ^α and n^α along the Killing vectors generating the rotational symmetry vanish, but for present purposes not much is gained by this).

2. In concrete applications we will choose n^α to be ingoing (in the sense that future directed null rays tangent to n^α will move towards smaller values of r) and ℓ^α to be (asymptotically) outgoing.

3. The minus sign in the cross normalisation is such that both vector fields are either future or past oriented (and we will of course choose the former).

4. Note that the individual normalisation of the ℓ^α and n^α is not fixed by the above conditions, i.e. one can still perform the boost

$$\ell^\beta \to \mathrm{e}^{+\alpha(x)}\ell^\beta \quad , \quad n^\beta \to \mathrm{e}^{-\alpha(x)} n^\beta \ . \tag{12.140}$$

This can e.g. be used to select a preferred normalisation for one of them. If ℓ^α has been fixed, then, in spherical symmetry and with the assumption that n^α is also purely radial (longitudinal), n^α is uniquely determined by the 2 conditions $n^\alpha n_\alpha = 0$ and $n^\alpha \ell_\alpha = -1$. This should be contrasted with the situation without spherical symmetry where, as discussed in section 12.3, there is still the additional freedom to perform null rotations on n^α.

Spherical symmetry (and the choice of spherically symmetric null vector fields) also has other implications. For instance, it follows from spherical symmetry that $\ell^\alpha \nabla_\alpha \ell^\beta$ will be some linear combination of ℓ^β and n^β (i.e. no component tangent to the transverse sphere),

$$\ell^\alpha \nabla_\alpha \ell^\beta = A\ell^\beta + Bn^\beta \tag{12.141}$$

(and likewise for n^α). Taking the scalar product with ℓ_α and using

$$(\nabla_\alpha \ell^\beta)\ell_\beta = \tfrac{1}{2}\nabla_\alpha(\ell^\beta \ell_\beta) = 0 \quad , \tag{12.142}$$

one finds that $B = 0$, so that automatically $\ell^\alpha \nabla_\alpha \ell^\beta \sim \ell^\beta$. Thus ℓ^α is automatically a geodesic null vector field, but perhaps not affinely parametrised. The proportionality constant A is then the *inaffinity* $A = \kappa_\ell$ (and likewise for n^α). Thus one has

$$\ell^\alpha \nabla_\alpha \ell^\beta = \kappa_\ell \ell^\beta \quad , \quad n^\alpha \nabla_\alpha n^\beta = \kappa_n n^\beta \quad . \tag{12.143}$$

The boost freedom can then e.g. be used to choose either ℓ^α or n^α to be affinely parametrised (but usually not both of them simultaneously).

The same argument as above leads to the conclusion that necessarily $n^\alpha \nabla_\alpha \ell^\beta \sim \ell^\beta$ as well, and in this case the constant of proportionality is fixed by taking the scalar product with n_β (and likewise with $\ell \leftrightarrow n$), leading to the conclusions

$$\ell^\alpha \nabla_\alpha n^\beta = -\kappa_\ell n^\beta \quad , \quad n^\alpha \nabla_\alpha \ell^\beta = -\kappa_n \ell^\beta \quad . \tag{12.144}$$

In particular, if $\kappa_\ell = 0$, say, this implies that n^α is parallel transported along ℓ, $\ell^\alpha \nabla_\alpha n^\beta = 0$.

It follows from either (12.143) or (12.144) that κ_ℓ and κ_n can be written as

$$\begin{aligned} \kappa_\ell &= -\ell^\alpha n^\beta \nabla_\alpha \ell_\beta = \ell^\alpha \ell^\beta \nabla_\alpha n_\beta \\ \kappa_n &= -n^\alpha \ell^\beta \nabla_\alpha n_\beta = n^\alpha n^\beta \nabla_\alpha \ell_\beta \quad , \end{aligned} \tag{12.145}$$

or

$$\begin{aligned} \kappa_\ell &= \tfrac{1}{2}\ell^\alpha \ell^\beta(\nabla_\alpha n_\beta + \nabla_\beta n_\alpha) = \tfrac{1}{2}\ell^\alpha \ell^\beta L_n g_{\alpha\beta} \\ \kappa_n &= \tfrac{1}{2}n^\alpha n^\beta(\nabla_\alpha \ell_\beta + \nabla_\beta \ell_\alpha) = \tfrac{1}{2}n^\alpha n^\beta L_\ell g_{\alpha\beta} \end{aligned} \tag{12.146}$$

Here L_ℓ and L_n are the Lie derivatives. Thus the inaffinities encode the information about the longitudinal projections of the derivatives $\nabla_\alpha \ell_\beta$ and $\nabla_\alpha n_\beta$, or of the Lie derivatives $L_\ell g_{\alpha\beta}$ and $L_n g_{\alpha\beta}$.

Other useful information is contained in the transverse (i.e. parallel to the sphere) projections of these objects. To define them, note that, as in section 12.3, associated with a choice of ℓ^α and n^α we have the decomposition of the metric

$$g_{\alpha\beta} = s_{\alpha\beta} - (\ell_\alpha n_\beta + \ell_\beta n_\alpha) \tag{12.147}$$

with $s_{\alpha\beta}$ the transverse spatial metric (on the sphere),

$$s_{\alpha\beta}dx^\alpha dx^\beta = r(z)^2 d\Omega^2 \quad , \tag{12.148}$$

but that in the current context this decomposition and the corresponding projectors $s^\alpha_{\ \beta}$ are now unique as the combination $\ell_\alpha n_\beta$ is boost-invariant.

The expansions of ℓ^α and n^α are defined as the transverse spatial projections of the divergence of ℓ^α respectively n^α, i.e.

$$\theta_\ell = s^{\alpha\beta}\nabla_\alpha \ell_\beta \quad , \quad \theta_n = s^{\alpha\beta}\nabla_\alpha n_\beta \ . \tag{12.149}$$

As in (12.101) and (12.102) of section 12.4, these can be written as

$$\begin{aligned} \theta_\ell &= \tfrac{1}{2}s^{\alpha\beta}L_\ell s_{\alpha\beta} = \frac{1}{\sqrt{s}}L_\ell\sqrt{s} \\ \theta_n &= \tfrac{1}{2}s^{\alpha\beta}L_n s_{\alpha\beta} = \frac{1}{\sqrt{s}}L_n\sqrt{s} \end{aligned} \tag{12.150}$$

With $\sqrt{s} = r(z)^2\sin\theta$, one finds more explicitly

$$\theta_\ell = \frac{2}{r}\ell^\alpha\partial_\alpha r \quad , \quad \theta_n = \frac{2}{r}n^\alpha\partial_\alpha r \ . \tag{12.151}$$

If one works with r as one of the coordinates, then this can also succinctly be written as

$$\theta_\ell = \frac{2}{r}\ell^r \quad , \quad \theta_n = \frac{2}{r}n^r \ . \tag{12.152}$$

As in (12.126) of section 12.5, we also have the relations

$$\nabla^\alpha \ell_\alpha = \theta_\ell + \kappa_\ell \quad , \quad \nabla^\alpha n_\alpha = \theta_n + \kappa_n \ . \tag{12.153}$$

Turning now to the Raychaudhuri equation for a spherically symmetric radial null congruence ℓ, the general result (12.129) (for $\kappa_\ell \neq 0$), i.e.

$$\frac{d}{d\tau}\theta_\ell = \kappa_\ell\theta_\ell - R_{\alpha\beta}\ell^\alpha\ell^\beta - \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta} \tag{12.154}$$

simplifies considerably. Spherical symmetry implies that the spatial shear and rotation tensors are zero (a spatial rotationally invariant 2-tensor is proportional to δ_{ik} which has neither a traceless nor an anti-symmetric part),

$$\sigma_{\alpha\beta} = \omega_{\alpha\beta} = 0 \ . \tag{12.155}$$

The vanishing of the rotation can also be deduced from the fact that ℓ is hypersurface orthogonal (specifically orthogonal to the family of null hypersurfaces generated by ℓ). Thus the Rauchaudhuri equation reduces to

$$\frac{d}{d\tau}\theta_\ell = \kappa_\ell\theta_\ell - R_{\alpha\beta}\ell^\alpha\ell^\beta - \tfrac{1}{2}\theta_\ell^2 \ . \tag{12.156}$$

13 Curvature IV: Curvature and Killing Vectors

In (pseudo-)Riemannian geometry the rich interplay between symmetries and geometry is reflected in relations between the curvature tensor and Killing vectors of a metric. Here we will explore some of these relations and their consequences.

13.1 Useful Identities Relating Curvature and Killing Vectors

Using the defining relation of the Riemann curvature tensor,

$$(\nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu)V_\lambda = -R^\rho{}_{\lambda\mu\nu}V_\rho \tag{13.1}$$

and its cyclic symmetry (8.25), it is possible to deduce that for a Killing vector K^μ,

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \ \ , \tag{13.2}$$

one has the following basic identity relating Killing vectors and the curvature tensor,

$$\nabla_\lambda\nabla_\mu K_\nu = R^\rho{}_{\lambda\mu\nu}\, K_\rho \ \ . \tag{13.3}$$

Indeed, proceeding as in the proof of the cyclic permutation identity (8.25), we deduce that

$$\nabla_{[\mu}\nabla_\nu K_{\lambda]} \sim R^\rho{}_{[\lambda\mu\nu]}K_\rho = 0 \ \ . \tag{13.4}$$

Since $\nabla_\nu K_\lambda$ is anti-symmetric, the total anti-symmetrisation is equivalent to cyclic permutation, and we therefore have

$$\nabla_\mu\nabla_\nu K_\lambda + \nabla_\nu\nabla_\lambda K_\mu + \nabla_\lambda\nabla_\mu K_\nu = 0 \ \ . \tag{13.5}$$

Using the Killing property in the second term, we can write this as

$$\nabla_\lambda\nabla_\mu K_\nu = -[\nabla_\mu,\nabla_\nu]K_\lambda = R^\rho{}_{\lambda\mu\nu}K_\rho \tag{13.6}$$

which is (13.3).

This identity can be interpreted as the statement (and can alternatively be derived from the fact) that the Lie derivative of the Christoffel symbols of a metric along a Killing vector of the metric is zero.

Indeed, first of all it is easy to see that under a general variation of the metric, the induced variation of the Christoffel symbol can be written as (20.14)

$$\delta\Gamma^\mu{}_{\nu\lambda} = \tfrac{1}{2}g^{\mu\rho}(\nabla_\nu\delta g_{\rho\lambda} + \nabla_\lambda\delta g_{\rho\nu} - \nabla_\rho\delta g_{\nu\lambda}) \ \ . \tag{13.7}$$

(this is easy to derive and also easy to remember as it takes exactly the same form as the definition of the Christoffel symbol, only with the metric replaced by the metric variation and the partial derivatives by covariant derivatives - see section 20.2 for a derivation and discussion of this identity). In particular, this exhibits the fact that the metric variation of the Christoffel symbols is a tensor (as could have been anticipated from the fact that the non-tensorial term in the transformation of the Christoffel symbols is independent of the metric), and additionally provides us with an explicit expression for this tensor.

Next, if the variation $\delta g_{\mu\nu} = L_\xi g_{\mu\nu}$ is the Lie derivative, i.e. the variation in the metric induced by an infinitesimal coordinate transformation $\delta x^\mu = \xi^\mu$, one can write this as

$$L_\xi \Gamma^\mu_{\ \nu\lambda} = \tfrac{1}{2} g^{\mu\rho}(\nabla_\nu L_\xi g_{\rho\lambda} + \nabla_\lambda L_\xi g_{\rho\nu} - \nabla_\rho L_\xi g_{\nu\lambda}) \ . \tag{13.8}$$

Note that in general the Lie derivative of a non-tensorial quantity is not well defined (or at least its definition requires a bit more thought). Here, however, it is natural to use the general formula (13.7) for the variation of the Christoffel symbols under metric variations to in particular define their Lie derivative (as the change in the Christoffel symbols induced by the Lie derivative of the metric).

Thus, adopting this definition and using $L_\xi g_{\mu\nu} = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\nu$, the right-hand side can (upon using the definition and cyclic symmetry of the Riemann tensor) be written as

$$\begin{aligned} L_\xi \Gamma^\mu_{\ \nu\lambda} &= \nabla_\lambda \nabla_\nu \xi^\mu - R^\mu_{\ \nu\lambda\rho}\xi^\rho \\ &= \nabla_\nu \nabla_\lambda \xi^\mu + R^\mu_{\ \lambda\rho\nu}\xi^\rho \ . \end{aligned} \tag{13.9}$$

In particular, if $\xi^\mu = K^\mu$ is a Killing vector, one has

$$L_K g_{\mu\nu} = 0 \quad \Rightarrow \quad L_K \Gamma^\mu_{\ \nu\lambda} = 0 \quad \Leftrightarrow \quad \nabla_\lambda \nabla_\nu K_\mu = R_{\mu\nu\lambda\rho} K^\rho \ , \tag{13.10}$$

which is equivalent to (13.3).

Contracting (13.3) over λ and ν, one obtains the next useful and frequently used identity

$$\nabla^\nu \nabla_\mu K_\nu = K^\nu R_{\mu\nu} \ . \tag{13.11}$$

In turn, one immediate consequence of this identity is (contract with K^μ, "integrate by parts" and use the anti-symmetry of $\nabla_\mu K_\nu$)

$$R_{\mu\nu} K^\mu K^\nu = (\nabla^\mu K^\nu)(\nabla_\mu K_\nu) + \nabla_\nu (K^\mu \nabla_\mu K^\nu) \ . \tag{13.12}$$

Note that this can also be deduced directly from (8.64) for $V^\mu \to K^\mu$ a Killing vector.

We will now look at various consequences of the identities (13.3), (13.11) and (13.12) which are useful and interesting in their own right. The implications of these identities for *maximal symmetry* and *maximally symmetric spaces* will be discussed separately in section 14 below.

13.2 Killing Vectors form a Lie algebra

As the first application, we will explicitly prove the assertion (9.45) of section 9.5 that the Lie bracket of two Killing vectors is again a Killing vector. While this follows from the general property (9.34) of the Lie derivative, which itself can (with some work) be deduced from the general definition of the Lie derivative (as the generator of the action of coordinate transformations on tensors), it is instructive and reassuring to verify this by an explicit calculation, also because similar manipulations are required when extending the analysis from Killing vectors to Killing tensors or Killing-Yano tensors briefly mentioned in section 10.5.[31]

Thus consider two Killing vectors A^μ and B^μ, say, i.e. vector fields satisfying

$$\nabla_\mu A_\nu + \nabla_\nu A_\mu = \nabla_\mu B_\nu + \nabla_\nu B_\mu = 0 \tag{13.13}$$

or, equivalently,

$$\nabla_\mu A_\nu = \nabla_{[\mu} A_{\nu]} \quad , \quad \nabla_\mu B_\nu = \nabla_{[\mu} B_{\nu]} \quad . \tag{13.14}$$

Explicitly, from (9.18) their Lie bracket is the vector field

$$C^\mu = [A, B]^\mu = A^\nu \nabla_\nu B^\mu - B^\nu \nabla_\nu A^\mu \quad , \tag{13.15}$$

and the claim is that C^μ is also Killing, i.e. that $\nabla_\mu C_\nu$ is anti-symmetric,

$$C^\mu = [A, B]^\mu \quad \Rightarrow \quad \nabla_\mu C_\nu = \nabla_{[\mu} C_{\nu]} \quad . \tag{13.16}$$

In calculating $\nabla_\mu C_\nu$ one encounters new first derivatives of A^μ and B^μ (which can be manipulated by the Killing equations), as well as second covariant derivatives, which can be reduced to zero-derivative terms by using the fundamental identity (13.3),

$$\begin{aligned} \nabla_\mu C_\nu &= (\nabla_\mu A_\lambda)\nabla^\lambda B_\nu - (\nabla_\mu B_\lambda)\nabla^\lambda A_\nu + A_\lambda \nabla_\mu \nabla^\lambda B_\nu - B_\lambda \nabla_\mu \nabla^\lambda A_\nu \\ &= -(\nabla_\lambda A_\mu)\nabla^\lambda B_\nu + (\nabla^\lambda B_\mu)\nabla_\lambda A_\nu + R_{\rho\mu\lambda\nu}(A^\lambda B^\rho - B^\lambda A^\rho) \quad . \end{aligned} \tag{13.17}$$

The first two terms are already manifestly anti-symmetric (the second being the anti-symmetrisation of the first), and by the cyclic identity and other symmetries of the Riemann tensor, so is the last term,

$$R_{\rho\mu\lambda\nu} - R_{\lambda\mu\rho\nu} = R_{\mu\lambda\rho\nu} - R_{\mu\rho\lambda\nu} = R_{\mu\nu\lambda\rho} = -R_{\nu\mu\lambda\rho} \quad . \tag{13.18}$$

Thus the Lie bracket of two Killing vectors is indeed again a Killing vector, as claimed.

[31] The interesting question if or when Killing-Yano tensors form a Lie algebra, extending and generalising the Lie algebra of the isometry group generated by the Killing vectors, is analysed in D. Kastor, S. Ray, J. Traschen, *Do Killing-Yano tensors form a Lie Algebra?*, `arXiv:0705.0535 [hep-th]`.

13.3 On the Isometry Algebra of a Compact Riemannian Space

In this section we will look at one immediate application of the identity (13.12),

$$R_{\mu\nu}K^\mu K^\nu = (\nabla^\mu K^\nu)(\nabla_\mu K_\nu) + \nabla_\nu(K^\mu \nabla_\mu K^\nu) \ , \tag{13.19}$$

namely an analogue of the Bochner-Yano type argument (given in remark 8 of section 8.5) for Killing vectors. Again, in order to be able to say something of substance we assume that the space we are dealing with is compact without boundary, and Riemannian, i.e. equipped with a positive-definite metric. In spite of this, the result we will derive is relevant also for physics, at least as long as one is willing to entertain the possibility that some higher-dimensional generalisations of general relativity (such as Kaluza-Klein theories discussed in section 44) plays a role in some more fundamental description of nature.

With the above assumptions, the first term on the right-hand side of (13.19) is non-negative and the second is a total derivative term that vanishes upon integration. Therefore for a Killing vector to exist on a compact Riemannian space, the integral of $R_{\mu\nu}V^\mu V^\nu$ must be non-negative as well.

This has two immediate implications:

1. If the Ricci tensor (regarded as a quadratic form) of a Riemannian metric on a compact space is negative, that metric can have no continuous isometries whatsoever.

2. If the metric on a compact Riemannian space has vanishing Ricci tensor, $R_{\mu\nu} = 0$, then any Killing vector is covariantly constant, $\nabla_\mu K_\nu = 0$.

Since the Lie bracket of two covariantly constant vector fields is zero,

$$\nabla_\mu V_\nu = \nabla_\mu W_\nu = 0 \quad \Rightarrow \quad [V,W]^\mu = V^\nu \nabla_\nu W^\mu - W^\nu \nabla_\nu V^\mu = 0 \ , \tag{13.20}$$

this means that continuous isometries of a space with vanishing Ricci tensor can at most be Abelian. An example is provided by the torus T^n equipped with the flat metric it inherits from regarding T^n as the periodic identification of $\mathbb{R}^n$. This metric has vanishing Ricci tensor (because evidently even the Riemann tensor is zero), but there are n linearly-independent (covariantly) constant translational Killing vectors (inherited from $\mathbb{R}^n$) that generate the Abelian isomtery group $U(1)^n$.

In Kaluza-Klein theory, one of the basic ideas is that gauge symmetries arise from isometries of the "internal" space living in the extra dimensions. This internal space is usually assumed to be compact (so as to be sufficiently small to have escaped our attention). Thus, if one wants to generate non-Abelian gauge theories in this way the above results provide one of the most basic constraints on the internal geometry, namely that the Ricci tensor should not be non-positive (but it does not have to be strictly positive everywhere).

13.4 Invariance of the Curvature along Killing Directions

It should be obvious and obviously true that for any Killing vector of a metric the scalar curvature of the metric inherits the corresponding symmetries of the metric, i.e. that it does not change along the orbits of that Killing vector,

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \quad \Rightarrow \quad K^\mu \nabla_\mu R = 0 \ , \tag{13.21}$$

or

$$L_K g_{\mu\nu} = 0 \quad \Rightarrow \quad L_K R = 0 \tag{13.22}$$

(with analogous statements for the Riemann and Ricci tensors).

While true, a covariant proof of this is a bit roundabout:

- One can start with the contracted Bianchi identity $\nabla^\mu G_{\mu\nu} = 0$, and contract it with K^ν to find

$$0 = (\nabla^\mu G_{\mu\nu})K^\nu = (\nabla^\mu R_{\mu\nu})K^\nu - \tfrac{1}{2} K^\nu \nabla_\nu R \ . \tag{13.23}$$

- Using the Killing equations, i.e. the anti-symmetry of $\nabla_\mu K_\nu$, and the symmetry of the Ricci tensor, one can write this as

$$K^\nu \nabla_\nu R = 2\nabla^\mu (K^\nu R_{\mu\nu}) \ . \tag{13.24}$$

- Using (13.11), one finally arrives at

$$K^\nu \nabla_\nu R = 2\nabla^\mu (\nabla^\nu \nabla_\mu K_\nu) = [\nabla^\mu, \nabla^\nu] \nabla_\mu K_\nu = 0 \tag{13.25}$$

  by (8.57).

Alternatively (and more quickly but somewhat less covariantly) one could have simply locally introduced an adapted coordinate system (9.59) in which $K = \partial_y$ and $\partial_y g_{\mu\nu} = 0$, to immediately deduce that then necessarily also $\partial_y R = 0$. However, on general grounds, and with an eye towards possible generalisations, it is always useful to have different arguments at one's disposal, in particular among them one which is covariant.

Analogously one can prove, either covariantly or non-covariantly (recommended in this case), that the Lie derivative of the Riemann tensor and the Ricci tensor along a Killing vector are zero,

$$L_K g_{\mu\nu} = 0 \quad \Rightarrow \quad L_K R_{\lambda\sigma\mu\nu} = 0 \quad , \quad L_K R_{\mu\nu} = 0 \ . \tag{13.26}$$

Since $\nabla_\mu K_\nu$ is anti-symmetric, one can write (13.11) more explicitly with the help of the formula (5.66) for the covariant divergence of an anti-symmetric tensor as

$$R^\mu_{\ \nu} K^\nu = \frac{1}{\sqrt{g}} \partial_\nu (\sqrt{g} \nabla^\mu K^\nu) \ . \tag{13.27}$$

This can be a quite efficient way to calculate certain components of the Ricci tensor of a metric, namely those which are of the form $R^\mu_{\ \nu} K^\nu$ for some Killing vector (the components referred to glibly as the "Killing components of the Ricci tensor" in the heading). In spite of this, this shortcut does not appear to be widely known or commonly used.

As an illustration of how this works, consider again the general static spherically symmetric metric (3.22),

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2 d\Omega^2 \ . \tag{13.28}$$

Among the Killing vectors of this metric is the vector field $\xi = \partial_t$ generating time-translations, and thus we can use (13.27) to determine the components $R^\mu_{\ t}$ of the Ricci tensor.

Since the only non-trivial component of ξ_μ is $\xi_t = -A(r)$, one has

$$\nabla_\mu \xi_\nu = \partial_\mu \xi_\nu - \Gamma^\lambda_{\ \mu\nu} \xi_\lambda = \partial_\mu \xi_\nu + A(r) \Gamma^t_{\ \mu\nu} \ , \tag{13.29}$$

and since according to (3.25) the only non-trivial component of $\Gamma^t_{\ \mu\nu}$ is

$$\Gamma^t_{\ rt} = \Gamma^t_{\ tr} = \frac{A'}{2A} \quad , \quad \Gamma^t_{\ \mu\nu} = 0 \quad \text{otherwise} \ , \tag{13.30}$$

the only non-trivial components of $\nabla_\mu \xi_\nu$ are

$$\nabla_t \xi_r = -\nabla_r \xi_t = A'/2 \quad , \quad \nabla_\mu \xi_\nu = 0 \quad \text{otherwise} \ . \tag{13.31}$$

Thus the only non-zero components of $\nabla^\mu \xi^\nu$ are

$$\nabla^r \xi^t = -\nabla^t \xi^r = A'/2AB \ . \tag{13.32}$$

With

$$\sqrt{g} = \sqrt{AB} r^2 \sin\theta \tag{13.33}$$

one then has

$$R^\mu_{\ t} = \frac{1}{\sqrt{AB} r^2} \partial_r (\sqrt{AB} r^2 \nabla^\mu \xi^r) \tag{13.34}$$

so that evidently

$$R^k_{\ t} = 0 \quad \text{with} \quad x^k = (r, \theta, \phi) \tag{13.35}$$

while

$$R^t{}_t = -\frac{1}{2\sqrt{AB}r^2}\partial_r(r^2 A'/\sqrt{AB}) \quad . \tag{13.36}$$

Explicitly one can write this as

$$R_{tt} = -AR^t{}_t = \frac{A''}{2B} - \frac{A'}{4B}(\frac{A'}{A} + \frac{B'}{B}) + \frac{A'}{rB} \quad . \tag{13.37}$$

<u>REMARKS:</u>

1. As you can check for yourself, this way of determining R_{tt} is much quicker than working it out from the general formula for the Ricci tensor involving the Christoffel symbols squared as well as their derivatives. In fact it is the quickest and slickest way to obtain R_{tt} by a calculation in coordinate components that I am aware of.
2. In the same way, one can also determine the angular components $R^\mu{}_\phi$, say, using the Killing vector $\eta = \partial_\phi$.
3. The only not obviously vanishing component of $R_{\mu\nu}$ (see the discussion in section 24.3) that cannot be obtained in this way is R_{rr}.

13.6 Killing Vectors as Solutions to the Maxwell Equations

A cute application of the identity (13.11) is the following. Recall that in the covariant Lorenz gauge

$$\nabla_\mu A^\mu = 0 \tag{13.38}$$

the vacuum Maxwell equations

$$\nabla^\mu F_{\mu\nu} = \nabla^\mu(\nabla_\mu A_\nu - \nabla_\nu A_\mu) = 0 \tag{13.39}$$

can be written as (8.126)

$$\nabla^\nu \nabla_\nu A^\mu = R^\mu_\nu A^\nu \quad , \tag{13.40}$$

while a Killing vector automatically satisfies

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \quad \Rightarrow \quad \nabla_\mu K^\mu = 0 \tag{13.41}$$

and the identity (13.11),

$$\nabla^\nu \nabla_\mu K^\nu = R_{\mu\nu} K^\nu \quad \Leftrightarrow \quad \nabla^\nu \nabla_\nu K^\mu = -R^\mu_\nu K^\nu \quad . \tag{13.42}$$

Thus the sign of the Ricci tensor in (13.40) and (13.42) is different, but evidently this difference disappears for a metric with vanishing Ricci tensor. This does not imply at all that the Riemann tensor is zero. Indeed, we will learn in section 19 that the vacuum Einstein equations (i.e. the gravitational field equations without or outside of matter sources) are simply the "Ricci flatness" conditions $R_{\mu\nu} = 0$ (19.36).

In that case (13.42) reduces to $\nabla^\nu \nabla_\nu K_\mu = 0$ which is of the same form as (13.40). Alternatively, because of the anti-symmetry of $\nabla_\mu K_\nu$, we can equivalently write (13.42) as

$$R_{\mu\nu} = 0 \quad \Rightarrow \quad \nabla^\mu(\nabla_\mu K_\nu - \nabla_\nu K_\mu) = 0 \ , \tag{13.43}$$

which, with the dictionary

$$A_\mu = K_\mu \quad , \quad F_{\mu\nu} = \nabla_\mu K_\mu - \nabla_\nu K_\mu \quad \Rightarrow \quad \nabla^\mu F_{\mu\nu} = 0 \ , \tag{13.44}$$

is identical to (13.39).

This means that any Killing vector of a solution to the vacuum Einstein equations automatically gives rise to a solution of the vacuum Maxwell equations in that gravitational background. Depending on the Killing vector this may or may not be a non-trivial ($F_{\mu\nu} \neq 0$) solution to the Maxwell equations, "rotational" Killing vectors typically giving rise to non-trivial solutions while for "translational" Killing vectors A_μ is pure gauge (see section 14.1 for a more precise characterisation of what is meant by "rotational" and "translational" Killing vectors at a given point).

For example, taking the general Killing vector (9.48) of Minkowski space (which certainly has vanishing Ricci tensor),

$$K^\alpha = \omega^\alpha_{\ \beta}\, x^\beta + \epsilon^\alpha \quad \Rightarrow \quad A_\alpha \equiv K_\alpha = \omega_{\alpha\beta}\, x^\beta + \epsilon_\alpha \ , \tag{13.45}$$

one finds that the associated Maxwell field strength tensor is

$$F_{\alpha\beta} = \partial_\alpha K_\beta - \partial_\beta K_\alpha = -2\omega_{\alpha\beta} \ . \tag{13.46}$$

Thus it vanishes for a purely translational Killing vector while a boost is associated with a constant electric field ($E_k \sim \omega_{0k}$) and a spatial rotation gives rise to a constant magnetic field ($B_k \sim \epsilon_{kij}\omega_{ij}$).

13.7 Killing Vectors and Komar Currents

Because the Einstein tensor $G_{\mu\nu}$ (8.108) is symmetric and conserved (the contracted Bianchi identity (8.107)), to any Killing vector one can associate (cf. the discussion in section 10.1) the conserved current

$$J_1^\mu = G^\mu_{\ \nu} K^\nu = R^\mu_{\ \nu} K^\nu - \tfrac{1}{2} R \delta^\mu_{\ \nu} K^\nu = R^\mu_{\ \nu} K^\nu - \tfrac{1}{2} R K^\mu \ . \tag{13.47}$$

However, because K^μ is Killing, one has $\nabla_\mu K^\mu = 0$ identically, as well as $K^\mu \nabla_\mu R = 0$ (as shown above), and hence one has a conserved current

$$J_a^\mu = R^\mu_{\ \nu} K^\nu - \tfrac{1}{2} a\, R K^\mu \quad , \quad \nabla_\mu J_a^\mu = 0 \ , \tag{13.48}$$

for any value of the real parameter a. Among this 1-parameter family of conserved currents, the choice

$$J^\mu_{a=0} \equiv J^\mu(K) = R^{\mu\nu} K_\nu \tag{13.49}$$

(the *Komar current*) is singled out by the fact that, by (13.11), it is not only conserved but can actually be written as the divergence of an anti-symmetric tensor,

$$J^\mu(K) = \nabla_\nu A^{\mu\nu} \quad , \quad A^{\mu\nu} = -A^{\nu\mu} = \nabla^\mu K^\nu \tag{13.50}$$

(which also shows directly, by (8.58), that $\nabla_\mu J^\mu(K) = 0$).

Thus the corresponding conserved charge, written as a hypersurface integral, can actually be written as a *surface integral* of components of $A_{\mu\nu}$. These define the so-called *Komar charges* associated to symmetries of the metric. They will make a brief appearance in section 23.4.

As an aside, note that while in the above we started off with Killing vectors, a similar story is actually true for any vector field. Namely, for any vector field ξ^μ define the current

$$J^\mu(\xi) = \nabla_\nu(\nabla^{[\mu}\xi^{\nu]}) = \tfrac{1}{2}\nabla_\nu(\nabla^\mu\xi^\nu - \nabla^\nu\xi^\mu) \ . \tag{13.51}$$

Note that this reduces to (13.50) for $\xi^\mu = K^\mu$ a Killing vector. Moreover, by (8.58) this current is conserved,

$$\nabla_\mu J^\mu(\xi) = 0 \quad \forall\, \xi^\mu \ . \tag{13.52}$$

When $\xi^\mu = K^\mu$ is a Killing vector, the current can alternatively be written in terms of the Ricci tensor as in (13.49). For a general ξ^μ one has, instead,

$$\begin{aligned} J^\mu(\xi) &= \tfrac{1}{2}\nabla_\nu(\nabla^\mu\xi^\nu - \nabla^\nu\xi^\mu) = \nabla_\nu\nabla^\mu\xi^\nu - \tfrac{1}{2}\nabla_\nu(\nabla^\nu\xi^\mu + \nabla^\mu\xi^\nu) \\ &= [\nabla^\nu, \nabla^\mu]\xi_\nu + \nabla^\mu(\nabla_\nu\xi^\nu) - \tfrac{1}{2}\nabla_\nu(\nabla^\nu\xi^\mu + \nabla^\mu\xi^\nu) \\ &= R^\mu_{\ \nu}\xi^\nu + \tfrac{1}{2}(g^{\alpha\beta}g^{\mu\nu} - g^{\mu\alpha}g^{\nu\beta})\nabla_\nu(\nabla_\alpha\xi_\beta + \nabla_\beta\xi_\alpha) \ , \end{aligned} \tag{13.53}$$

where we made use of (8.51). Note that this indeed reduces to (13.49) for a Killing vector, for which the second term on the right-hand side is absent.

The existence of these identically conserved currents and the corresponding surface charge densities $\nabla^{[\mu}\xi^{\nu]}$ reflects the fact that in general relativity (more generally, in any generally covariant theory) all vector fields can be considered as the generators of symmetries (in the sense of coordinate transformations). Indeed, the currents $J^\mu(\xi)$ can be shown to be precisely the corresponding Noether currents arising from the Lagrangian formulation of general relativity to be discussed in section 20. We will establish this result in section 20.6. Nevertheless, the currents and charges associated with Killing vectors turn out to play a privileged role, and we will in particular relate the Komar charge for a timelike Killing vector to the ADM mass of an isolated (asymptotically) static system in section 23.4.

14 Curvature V: Maximal Symmetry and Constant Curvature

As a preparation for our discussion of cosmology in sections 33 - 38, in this section we will discuss some aspects of what are known as *maximally symmetric spaces*. These are spaces that admit the maximal number of Killing vectors (which turns out to be $n(n+1)/2$ for an n-dimensional space or space-time).

As we will discuss later on, in the context of the *Cosmological Principle*, such spaces, which are simultaneously homogeneous ("the same at every point") and isotropic ("the same in every direction") provide an (admittedly highly idealised) description of space in a cosmological space-time.

If you already know (or are willing to believe) that in any spatial dimension n there are essentially only 3 such spaces, namely the Euclidean space $\mathbb{R}^n$, the sphere S^n, and its negative curvature counterpart, the hyperbolic space H^n (all equipped with their standard metrics), you can skip this section, and may just want to refer to section 14.3 where it is shown that these 3 standard metrics can be written in a unified way as

$$ds^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega_{n-1}^2 \tag{14.1}$$

for $k = 0, \pm 1$ respectively.

The discussion of maximally symmetric spaces and, in particular, space-times will be taken up again and continued in section 39 (which can also be read as a direct sequel to this section, without the intervening sections 33 - 38 on cosmology).

14.1 Homogeneous, Isotropic and Maximally Symmetric Spaces

In order to understand how to define and characterise maximally symmetric spaces, we will need to obtain some more information about how Killing vectors can be classified.

Our starting point is, as in the previous section, the identity (13.3), reproduced here with the explicit x-dependence included for present purposes,

$$\nabla_\lambda \nabla_\mu K_\nu(x) = R^\rho{}_{\lambda\mu\nu}(x) K_\rho(x) \quad . \tag{14.2}$$

In particular, this shows that the second derivatives of the Killing vector at a point x_0 are again expressed in terms of the value of the Killing vector itself at that point. This means (think of Taylor expansions) that, remarkably, a Killing vector field $K^\mu(x)$ is completely and uniquely determined everywhere by the values of $K_\mu(x_0)$ and $\nabla_\mu K_\nu(x_0)$ at a single point x_0.

A set of Killing vectors $\{K_\mu^{(i)}(x)\}$ is said to be *linearly independent* if any linear relation of the form

$$\sum_i c_i K_\mu^{(i)}(x) = 0 \quad , \tag{14.3}$$

with *constant* coefficients c_i implies $c_i = 0$ (the reason for insisting on constant coefficients rather than functions $c_i(x)$ in this definition is of course that if K^μ is a Killing vector, then so is cK^μ iff c is constant).

Since, in an n-dimensional space(-time) there can be at most n linearly independent vectors $(K_\mu(x_0))$ at a point, and at most $n(n-1)/2$ independent anti-symmetric matrices $(\nabla_\mu K_\nu(x_0))$, we reach the conclusion that an n-dimensional space(-time) can have at most

$$n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} \tag{14.4}$$

independent Killing vectors. A space(-time) with this maximal number of Killing vectors is called *maximally symmetric*.

An example of a metric with the maximal number of Killing vectors is, none too surprisingly, n-dimensional Minkowski space, where $n(n+1)/2$ agrees with the dimension of the Poincaré group, the group of transformations that leave the Minkowski metric invariant.

Other examples of spaces that are maximally symmetric spaces are provided by spheres with their standard metric (e.g. we already know that the 2-sphere has $3 = 2(2+1)/2$ linearly independent Killing vectors, given explicitly in (9.55)). We will show below that spheres and their negative curvature hyperbolic counterparts are the unique non-trivial maximally symmetric spaces (with a corresponding statement for maximally symmetric space-times, which we will study in detail in section 39).

We will now see how the data $K^\mu(x_0)$ and $\nabla_\mu K_\nu(x_0)$ are related to translations and rotations:

- We define a *homogeneous space* to be such that it has infinitesimal isometries that carry any given point x_0 into any other point in its immediate neighbourhood (this could be stated in more fancy terms!). Thus the metric must admit Killing vectors that, at any given point, can take all possible values. Thus we require the existence of Killing vectors for arbitrary $K_\mu(x_0)$. This means that the n-dimensional space admits the maximal number n of *translational Killing vectors*.

- We define a space to be *isotropic at a point* x_0 if it has isometries that leave the given point x_0 fixed and such that they can rotate any vector at x_0 into any other vector at x_0. Therefore the metric must admit Killing vectors such that $K_\mu(x_0) = 0$ but such that $\nabla_\mu K_\nu(x_0)$ is an arbitrary anti-symmetric matrix (for instance to be thought of as an element of the Lie algebra of $SO(n)$). This means that the n-dimensional space admits the maximal number $n(n-1)/2$ of *rotational Killing vectors*.

- Finally, we define a *maximally symmetric space* to be a space with a metric with the maximal number $n(n+1)/2$ of Killing vectors.

For example, as mentioned before, the 2-sphere is maximally symmetric, with 3 linearly independent Killing vectors, given explicitly e.g. in (9.55). The decomposition of these 3 Killing vectors into 1 rotational and 2 translational Killing vectors depends on the point on the 2-sphere, the rotational Killing vector always being associated with the rotations around the axis through that point, and the translational Killing vectors being formed by the remaining 2 linearly independent combinations of Killing vectors. The decomposition given in (9.55) is adapted to rotations around the north (or south) pole, with $V_{(3)} = \partial_\phi$ the corresponding rotational Killing vector. Note that this Killing vector acts as a rotation at / around the poles but that it acts as a translation away from the poles (where some other linear combination of the 3 Killing vectors would be the rotational Killing vector). We will come back to this in slightly more general terms below.

Some simple and fairly obvious consequences of these definitions are the following:

1. A homogeneous and isotropic space is maximally symmetric.
2. A space that is isotropic for all x is also homogeneous.
3. (1) and (2) now imply that a space which is isotropic around every point is maximally symmetric.
4. Finally one also has the converse, namely that a maximally symmetric space is homogeneous and isotropic.

Property (2) is a consequence of the fact that constant linear combinations of Killing vectors are again Killing vectors and that, as mentioned above in the context of the 2-sphere, away from the origin of the rotation a rotation acts just like a translation. Technically, the difference between two rotational Killing vectors at x and $x + dx$ can be shown to be a translational Killing vector. To see this (roughly), consider 2 Killing vectors K and L describing rotations about a point x_0 and a point $x_0 + dx$ respectively, i.e.

$$K^\mu(x_0) = 0 \quad , \quad L^\mu(x_0 + dx) = 0 \ . \tag{14.5}$$

and, in particular,

$$(\nabla_\mu L_\nu)(x_0 + dx) \neq 0 \ . \tag{14.6}$$

Now consider the difference

$$M^\mu(x) = K^\mu(x) - L^\mu(x) \tag{14.7}$$

which is still a Killing vector. At $x = x_0$ one has

$$M^\mu(x_0) = K^\mu(x_0) - L^\mu(x_0) = -L^\mu(x_0 + dx - dx) \ . \tag{14.8}$$

Now expanding $L^\mu(x)$ around the point $x+dx$, one has (in an inertial coordinate system at x_0, say)

$$M_\mu(x_0) = dx^\lambda \nabla_\lambda L_\mu(x_0 + dx) \neq 0 \tag{14.9}$$

while its matrix of covariant derivatives ∇M vanishes there due to the crucial identity $\nabla\nabla L \sim L$ (14.2). Thus M defines a translational Killing vector at x_0.

In practice the characterisation of a maximally symmetric space which is easiest to use is (3) because it requires consideration of only one type of symmetries, namely rotational symmetries.

14.2 Curvature Tensor of a Maximally Symmetric Space

On the basis of these simple considerations we can already determine the form of the Riemann curvature tensor of a maximally symmetric space. We will see that maximally symmetric spaces are spaces of constant curvature in the sense that the Riemann curvature tensor is simply and purely algebraically related to the metric by

$$R_{ijkl} = k(g_{ik}g_{jl} - g_{il}g_{jk}) \tag{14.10}$$

for some constant k.

This result could be obtained by making systematic use of the higher order integrability conditions for the existence of a maximal number of Killing vectors. The argument given below is less covariant but more elementary.

Assume for starters that the space is isotropic at x_0 and choose a Riemann normal coordinate system centered at x_0. Thus the metric at x_0 is $g_{ij}(x_0) = \eta_{ij}$ where we may just as well be completely general and assume that

$$\eta_{ij} = \mathrm{diag}(\underbrace{-1,\ldots,-1}_{p \text{ times}}\underbrace{+1,\ldots,+1}_{q \text{ times}}) \quad , \tag{14.11}$$

where $p + q = n$ and we only assume $n > 2$.

If the metric is supposed to be isotropic at x_0 then, in particular, the curvature tensor at the origin must be invariant under Lorentz rotations. Now we know (i.e. you should know from your Special Relativity course) that the only invariants of the Lorentz group are the Minkowski metric and products thereof, and the totally anti-symmetric Levi-Civita tensor. Thus the Riemann curvature tensor has to be of the form

$$R_{ijkl}(x_0) = a\eta_{ij}\eta_{kl} + b\eta_{ik}\eta_{jl} + c\eta_{il}\eta_{jk} + d\epsilon_{ijkl} \quad , \tag{14.12}$$

where the last term is only possible for $D = 4$. The symmetries of the Riemann tensor imply that $a = d = b + c = 0$, and hence we are left with

$$R_{ijkl}(x_0) = b(\eta_{ik}\eta_{jl} - \eta_{il}\eta_{jk}) \quad , \tag{14.13}$$

Thus in an arbitrary coordinate system we will have

$$R_{ijkl}(x_0) = b(g_{ik}(x_0)g_{jl}(x_0) - g_{il}(x_0)g_{jk}(x_0)) \ \ , \tag{14.14}$$

If we now assume that the space is isotropic around every point, then we can deduce that

$$R_{ijkl}(x) = b(x)(g_{ik}(x)g_{jl}(x) - g_{il}(x)g_{jk}(x)) \tag{14.15}$$

for some function $b(x)$. Therefore the Ricci tensor and the Ricci scalar are

$$\begin{aligned} R_{ij}(x) &= (n-1)b(x)g_{ij} \\ R(x) &= n(n-1)b(x) \ \ . \end{aligned} \tag{14.16}$$

and the Riemann curvature tensor can also be written as

$$R_{ijkl} = \frac{R}{n(n-1)}(g_{ik}g_{jl} - g_{il}g_{jk}) \ \ , \tag{14.17}$$

while the Einstein tensor is

$$G_{ij} = b[(n-1)(1-n/2)]g_{ij} \ \ . \tag{14.18}$$

For $n > 2$ the contracted Bianchi identity $\nabla^i G_{ij} = 0$ now implies that $b(x)$ has to be a constant, and we have thus established (14.10). Note that we also have

$$R_{ij} = k(n-1)g_{ij} \ \ , \tag{14.19}$$

so that a maximally symmetric space(-time) is automatically a solution to the vacuum Einstein equations with a cosmological constant. In the physically relevant case $p = 1$ these are known as de Sitter or anti de Sitter space-times. We will come back to them in detail later on, in section 39.

14.3 Maximally Symmetric Metrics I: Solving the Constant Curvature Conditions

We are interested not just in the curvature tensor of a maximally symmetric space but in the metric itself. I will give you two derivations of the metric of a maximally symmetric space, one by directly solving the differential equation

$$R_{ij} = k(n-1)g_{ij} \tag{14.20}$$

for the metric g_{ij}, the other by a direct geometrical construction of the metric which makes the isometries of the metric manifest.

As a maximally symmetric space is in particular spherically symmetric, we can write its metric in the form

$$ds^2 = B(r)dr^2 + r^2 d\Omega^2_{(n-1)} \ \ , \tag{14.21}$$

where $d\Omega^2_{(n-1)} = d\theta^2 + \ldots$ is the volume-element for the $(n-1)$-dimensional sphere or its counterpart in other signatures. For concreteness, we now fix on $n = 3$, but the argument given below goes through in general.

It is straighforward to calculate the components of the Ricci tensor of this metric. This can be viewed as a special case of the calculations leading to the Schwarzschild metric in section 24, setting the function called $A(r)$ there to zero (of course before having divided by it anywhere ...).

The only independent components are R_{rr} and $R_{\theta\theta}$,

$$\begin{aligned} R_{rr} &= \frac{1}{r}\frac{B'}{B} \\ R_{\theta\theta} &= -\frac{1}{B} + 1 + \frac{rB'}{2B^2} \quad . \end{aligned} \tag{14.22}$$

We now want to solve the equations

$$\begin{aligned} R_{rr} &= 2kg_{rr} = 2kB(r) \\ R_{\theta\theta} &= 2kg_{\theta\theta} = 2kr^2 \quad . \end{aligned} \tag{14.23}$$

From the first equation we obtain

$$B' = 2krB^2 \quad , \tag{14.24}$$

and from the second equation we deduce

$$\begin{aligned} 2kr^2 &= -\frac{1}{B} + 1 + \frac{rB'}{2B^2} \\ &= -\frac{1}{B} + 1 + \frac{2kr^2B^2}{2B^2} \\ &= -\frac{1}{B} + 1 + kr^2 \quad . \end{aligned} \tag{14.25}$$

This is an algebraic equation for B solved by

$$B = \frac{1}{1 - kr^2} \tag{14.26}$$

(and this also solves the first equation). Therefore we have determined the metric of a a maximally symmetric space to be

$$ds^2 = \frac{dr^2}{1 - kr^2} + r^2 d\Omega^2_{(n-1)} \quad . \tag{14.27}$$

Clearly, for $k = 0$ this is just the flat metric on $\mathbb{R}^n$. For $k = 1$, this should also look familiar as the standard metric on the sphere. If not, don't worry, we will be more explicit about this below.

We will also rederive these metrics in the next section in a way that makes the isometries of the metric manifest (and which thus also excludes the possibility, not logically ruled out by the arguments given so far, that the metrics we have found here for $k \neq 0$ are spherically symmetric and have constant Ricci curvature but are not actually maximally symmetric).

<u>REMARKS:</u>

1. First of all let us note that for $k \neq 0$ essentially only the sign of k matters as $|k|$ only affects the overall size of the space and nothing else (and can therefore be absorbed in the scale factor $a(t)$ of the metric (33.1) that will be the starting point for our investigations of cosmology). To see this note that a metric of the form (14.27), but with k replaced by k/L^2,

$$ds^2 = \frac{dr^2}{1 - kr^2/L^2} + r^2 d\Omega^2 \quad , \tag{14.28}$$

can, by introducing $\tilde{r} = r/L$, be put into the form

$$ds^2 = \frac{dr^2}{1 - kr^2/L^2} + r^2 d\Omega^2 = L^2 \left(\frac{d\tilde{r}^2}{1 - k\tilde{r}^2} + \tilde{r}^2 d\Omega^2 \right) \quad . \tag{14.29}$$

We now see explicitly that a rescaling of k by a constant factor is equivalent to an overall rescaling of the metric, and thus we will just need to consider the cases $k = 0, \pm 1$. However, occasionally it will also be convenient to think of k as a continuous parameter, the 3 geometries then being distinguished by $k < 0, k = 0$ and $k > 0$ respectively.

2. For $k = +1$, we have

$$ds^2 = \frac{dr^2}{1 - r^2} + r^2 d\Omega^2_{(n-1)} \quad . \tag{14.30}$$

Thus, obviously the range of r is restricted to $r \leq 1$ and by the change of variables $r = \sin\psi$, the metric can be put into the standard form of the metric on S^n in polar coordinates,

$$ds^2 \equiv d\Omega^2_n = d\psi^2 + \sin^2\psi d\Omega^2_{n-1} \quad . \tag{14.31}$$

This makes it clear that the singularity at $r = 1$ is just a coordinate singularity.

3. For $k = -1$, on the other hand, we have

$$ds^2 = \frac{dr^2}{1 + r^2} + r^2 d\Omega^2_{n-1} \quad . \tag{14.32}$$

Thus the range of r is $0 \leq r < \infty$, and we can use the change of variables $r = \sinh\psi$ to write the metric as

$$ds^2 \equiv d\tilde{\Omega}^2_n = d\psi^2 + \sinh^2\psi d\Omega^2_{n-1} \quad . \tag{14.33}$$

This is the standard metric of a hyperboloid H^n in polar coordinates, and I have introduced the notation $d\tilde{\Omega}^2_n$ for the line-element on the "unit" n-hyperboloid in analogy with the standard notation $d\Omega^2_n$ for the line-element on the unit n-sphere.

4. Thus, collectively we can write the three metrics as

$$ds^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega^2_{(n-1)} = d\psi^2 + g_k(\psi)^2 d\Omega^2_{n-1} \tag{14.34}$$

where

$$g_k(\psi) = \begin{cases} \psi & k = 0 \\ \sin\psi & k = +1 \\ \sinh\psi & k = -1 \end{cases} \tag{14.35}$$

5. Finally, by making the change of variables

$$r = \bar{r}(1 + k\bar{r}^2/4)^{-1} \quad , \tag{14.36}$$

one can put the metric into the isotropic form

$$ds^2 = (1 + k\bar{r}^2/4)^{-2}(d\bar{r}^2 + \bar{r}^2 d\Omega^2_{(n-1)}) = (1 + k\vec{x}^2/4)^{-2}\, d\vec{x}^2 \quad . \tag{14.37}$$

Note that this differs by the *conformal factor* $(1 + k\bar{r}^2/4)^{-2} > 0$ from the flat metric. One says that such a metric is *conformally flat.* Thus what we have shown is that every maximally symmetric space is conformally flat. Conformally flat, on the other hand, does not by any means imply maximally symmetric (the conformal factor could be any function of the radial and angular variables).

Note also that the metric in this form is just the 3- (or n-) dimensional generalisation of the 2-dimensional constant curvature metric on the 2-sphere in stereographic coordinates (11.62) (for $k = +1$) or of the Poincaré disc metric of H^2 (11.66) (for $k = -1$).

14.4 Maximally Symmetric Metrics II: Embeddings

Recall that the standard metric on the n-sphere can be obtained by restricting the flat metric on an ambient $\mathbb{R}^{n+1}$ to the sphere. We will generalise this construction a bit to allow for $k < 0$ and other signatures as well.

Consider a flat auxiliary vector space V of dimension $(n+1)$ with metric

$$ds^2 = d\vec{x}^2 + \frac{1}{k}dz^2 \quad , \tag{14.38}$$

where $\vec{x} = (x^1, \ldots, x^n)$ and $d\vec{x}^2 = \eta_{ij}dx^i dx^j$. Thus the metric on V has signature $(p, q+1)$ for k positive and $(p+1, q)$ for k negative. The group $G = SO(p, q+1)$ or $G = SO(p+1, q)$ has a natural action on V by isometries of the metric. The full isometry group of V is the semi-direct product of this group with the Abelian group of translations (just as in the case of the Euclidean or Poincaré group).

Now consider in V the hypersurface Σ defined by

$$k\vec{x}^2 + z^2 = 1 \quad . \tag{14.39}$$

This equation breaks all the translational isometries, but by the very definition of the group G it leaves this equation, and therefore the hypersurface Σ, invariant. It follows that G will act by isometries on Σ with its induced metric. Since $\dim G = n(n+1)/2$, the n-dimensional space has $n(n+1)/2$ Killing vectors and is therefore maximally symmetric.

<u>Remarks:</u>

1. In fact, G acts transitively on Σ (thus Σ is homogeneous) and the stabiliser at a given point is isomorphic to $H = SO(p,q)$ (so Σ is isotropic), and therefore Σ can also be described as the homogeneous space

$$\begin{aligned} \Sigma_{k>0} &= SO(p,q+1)/SO(p,q) \\ \Sigma_{k<0} &= SO(p+1,q)/SO(p,q) \ . \end{aligned} \tag{14.40}$$

2. The Killing vectors of the induced metric are simply the restriction to Σ of the standard generators of G on the vector space V.

3. For Euclidean signature, these spaces are spheres for $k > 0$ and hyperboloids for $k < 0$, and in other signatures they are the corresponding generalisations. In particular, for $(p,q) = (1, n-1)$ we obtain *de Sitter space-time* for $k = 1$ and *anti-de Sitter space-time* for $k = -1$. We will discuss their embeddings, and coordinate systems for them, in much more detail in section 39.

It just remains to determine explicitly this induced metric. For this we start with the defining relation of Σ and differentiate it to find that on Σ one has

$$dz = -\frac{k\vec{x}.d\vec{x}}{z} \ , \tag{14.41}$$

so that

$$dz^2 = \frac{k^2(\vec{x}.d\vec{x})^2}{1-k\vec{x}^2} \ . \tag{14.42}$$

Thus the metric (14.38) restricted to Σ is

$$\begin{aligned} ds^2|_\Sigma &= d\vec{x}^2 + \frac{1}{k}dz^2|_\Sigma \\ &= d\vec{x}^2 + \frac{k(\vec{x}.d\vec{x})^2}{1-k\vec{x}^2} \ . \end{aligned} \tag{14.43}$$

Passing from Cartesian coordinates $\vec{x}$ to spherical coordinates (r,θ,ϕ), with

$$r^2 = \vec{x}^2 = \eta_{ij}x^ix^j \quad , \quad rdr = \vec{x}.d\vec{x} \quad , \quad dr^2 = (\vec{x}.d\vec{x})^2/\vec{x}^2 \ , \tag{14.44}$$

this metric can also be written as

$$ds^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega^2_{(n-1)} \ . \tag{14.45}$$

This is precisely the metric (14.27) we obtained in the previous section.

15 Hypersurfaces I: Basics

Hypersurfaces play important roles in general relativity, appearing in many different contexts, e.g. in the form of hypersurfaces of constant time (for some choice of time coordinate), or as boundaries of space-time regions over which one would like to integrate some quantity, etc.

In this section I will describe some of the basic aspects of what is known as the intrinsic geometry of such hypersurfaces. The geometry of surfaces is of course a classical subject of geometry, the study by Gauss of 2-dimensional surfaces embedded in $\mathbb{R}^3$ and his *Theorema Egregium* regarding the intrinsic nature of the curvature of a surface marking the birth of differential geometry, and as such is described in many places. We will just barely scratch the surface of this subject and concentrate on those aspects that are of evident (rather than just potential) relevance for general relativity.[32]

Strictly speaking very little of this is needed or used in the elementary applications of general relativity in the later parts of these notes, and therefore this section could also be skipped at first. However, this is a subject which is interesting in its own right and which also leads to an improved understanding of the things that we have done so far regarding tensors and tensor calculus.

Moreover, some results of this section, and its accompanying sections 16 and 17, come in handy e.g. when one needs to integrate some quantity (like a component of a conserved current) over a hypersurface, say. Moreover, some basic familiarity with this subject is required to better understand certain slightly (but not terribly) advanced aspects of general relativity like the Hamiltonian formulation of general relativity (section 21, this also requires a knowledge of the extrinsic geometry of hypersurfaces to be discussed in section 18) or the event horizon of the Schwarzschild black hole geometry (which turns out to be a null hypersurface with certain special features to be discussed in more detail in section 32).

15.1 Basic Definitions: Embeddings and Embedded Hypersurfaces

We start by defining what we mean (at least roughly speaking) by a hypersurface and an embedding or an embedded hypersurface.

A *hypersurface* $\Sigma = \Sigma_{(n)}$ is an n-dimensional subspace (submanifold) Σ of a $D = n+1$-dimensional space(-time) (manifold) $M = M_{(n+1)}$, $\Sigma \subset M$ (one also says that Σ has *codimension* 1).

[32] For an excellent introduction to the intrinsic and extrinsic geometry of hypersurfaces geared towards applications to general relativity, see section 3 of E. Poisson, *A Relativist's Toolkit: the Mathematics of Black Hole Mechanics*. A compact summary of the relevant results can also be found in Appendices C-E of S. Carroll, *Spacetime and Geometry*.

Ther are two distinct ways of describing and thinking about hypersurfaces.

1. Embeddings

   On the one hand, one can describe a hypersurface in terms of an *embedding*

   $$\Phi: \quad \Sigma = \Sigma_{(n)} \hookrightarrow M = M_{(n+1)} \tag{15.1}$$

   of Σ into M, specified by the map Φ (which will need to satisfy some appropriate regularity conditions - we will come back to this below).

2. Embedded Hypersurfaces

   On the other hand one can think of a hypersurface concretely as a subspace of M, i.e. as an (already) *embedded hypersurface*

   $$\Sigma = \Sigma_{(n)} \subset M = M_{(n+1)} \quad , \tag{15.2}$$

   specified e.g. by

   $$\Sigma = \{x \in M : \; S(x) = 0\} \quad . \tag{15.3}$$

   for some real-valued function S on M.

The 1st description may look a bit abstract, in particular since it seems to grant some autonomy and independent existence to Σ outside the space-time. However, if one equips Σ with coordinates y^a, say, and M is described by coordinates x^α, then such an embedding Φ is given very concretely by specifying the point in M with coordinates x^α that corresponds to a point in Σ with coordinates y^a. Thus an embedding is given by the functions or parametric equations

$$\Phi: \quad x^\alpha = x^\alpha(y^a) \quad . \tag{15.4}$$

Typically in general relativity, at least as far as its more elementary aspects are concerned, hypersurfaces naturally arise as concretely embedded subspaces of space-time (without an independent existence outside of the space-time), for example in the guise of hypersurfaces of constant time $t = t_0$ for some time coordinate t, or as slices of constant r for some radial coordinate r etc.

Nevertheless, for certain purposes it is useful even then to also have the 1st description at one's disposal, in particular when it comes to questions of relating tensors on M to tensors on Σ, determining induced metrics and volume elements on Σ etc. All of this is more transparent when expressed in terms of local coordinates on Σ and M and the relations among them. These are precisely the data $x^\alpha(y^a)$ locally defining an embedding.

<u>Examples:</u>

1. As the first and most basic example, let us consider a spacelike hypersurface Σ of constant time in Minkowski space M, the latter equipped with standard inertial coordinates $x^\alpha = (t, x^k)$.

- In the 1st description one has in mind that one is given the space $\Sigma = \mathbb{R}^3$ with Cartesian coordinates y^k, and that one embeds it into Minkowski space e.g. via the relations $x^\alpha = x^\alpha(y)$ given explicitly by

$$t(y) = t_0 \quad , \quad x^k(y) = y^k \ . \tag{15.5}$$

- In the 2nd description, one defines the same spacelike hypersurface by the equation

$$S(t, x^k) = t - t_0 = 0 \ . \tag{15.6}$$

2. The second example example is the standard 2-sphere S^2 of radius r_0 in $\mathbb{R}^3$. This can be described

   - either in terms of an embedding $x^\alpha(y^a)$, where $x^\alpha = (x^1, x^2, x^3)$ are Cartesian coordinates on $\mathbb{R}^3$ and $y^a = (\theta, \phi)$ are coordinates on S^2, e.g.

$$\begin{aligned} x^\alpha(y^a): \quad & x^1(\theta,\phi) = r_0 \sin\theta \cos\phi \\ & x^1(\theta,\phi) = r_0 \sin\theta \sin\phi \\ & x^3(\theta,\phi) = r_0 \cos\theta \end{aligned} \tag{15.7}$$

   - or by the equation

$$S(x^\alpha) = (x^1)^2 + (x^2)^2 + (x^3)^2 - (r_0)^2 = 0 \ . \tag{15.8}$$

   If one works in spherical coordinates $x^\alpha = (r, \theta_x, \phi_x)$ on $\mathbb{R}^3$ from the outset, then both the parametric and the embedding description simplify accordingly, the former taking the form

$$x^1(y) = r_0 \quad , \quad x^2(y) = y^1 \quad , \quad x^3(y) = y^2 \ , \tag{15.9}$$

   or

$$r = r_0 \quad , \quad \theta_x = \theta \quad , \quad \phi_x = \phi \ , \tag{15.10}$$

   and the latter being the obvious

$$S(x^\alpha) = S(r) = r - r_0 = 0 \ . \tag{15.11}$$

   This shows that it is probably a good idea to try to introduce and use coordinates on the ambient space-time that are somehow adapted to the hypersurface one is interested in.

3. As the third example consider the future lightcone of a point in Minkowski space M. Without loss of generality we can choose that point to be the origin of the coordinate system. Using spherical coordinates $(x^0 = t, x^1 = r, x^2 = \theta, x^3 = \phi)$ on M, we can describe the future lightcone Σ

- either by introducing coordinates $(y^1 = v, y^2, y^3)$ on Σ and specifying the embedding as

$$t(v, y^k) = v \quad , \quad r(v, y^k) = v \quad , \quad x^k(v, y^k) = y^k \tag{15.12}$$

- or by requiring the constraint

$$S(x^\alpha) = x^0 - x^1 = t - r = 0 \ . \tag{15.13}$$

As we will see later, this is an example of a null (or lightlike) surface.

It should be clear from these examples that the description of a hypersurface as an embedded surface $S = 0$ typically looks a bit simpler or more usable but that, depending on what one wants to do, one or the other description may be more convenient, and that it is useful to be able to pass back and forth betwen them.

REMARKS:

1. A simple and simple-minded way of seeing the relation betwen the two descriptions and passing from one to the other, generalising the above embedding of the sphere in terms of spherical coordinates on the ambient space, is to use $S(x^\alpha)$ as a new coordinate, at least in a neighbourhood of the hypersurface Σ, i.e. to trade any one of the coordinates $S(x)$ depends on for S. Calling the new coordinates (S, x^a), where the x^a are arbitrary independent coordinates, one may as well use the x^a as coordinates on the surface Σ defined by $S = 0$. Then the parametric description $x^\alpha(y)$ of the surface $S = 0$ can be chosen to be

$$S(y) = 0 \quad , \quad x^a(y) = y^a \ , \tag{15.14}$$

or $x^a = f^a(y)$ for some functions f^a (this just amounting to a coordinate transformation $y^a \to f^a(y)$ of the coordinates y^a on Σ).

It is frequently convenient to introduce such a coordinate system adapted to Σ at least at intermediate stages of a calculation, and we will occasionally make use of this.

2. One important and recurrent theme is the relation between tensors on a hypersurface Σ and tensors on the ambient space(-time) M, i.e. the relation provided by the embedding of Σ into M between

   - Σ-tensors: objects which transform like tensors under transformations of the coordinates y^a on Σ and are scalars (invariant) under transformations of the coordinates x^α on M, and
   - M-tensors: objects which transform (as usual) like tensors under transformations of the coordinates x^α on M and are scalars (invariant) under transformations of the coordinates y^a on Σ.

3. The Σ-tensor of principal interest is, as in the case of the ambient space M, the metric tensor $h_{ab}(y)$ of Σ. In general the metric tensor (and its associated curvature tensor, discussed at length in sections 8 and 11) provide a complete local characterisation of the *intrinsic geometry* of a space (or space-time), i.e. the properties that can be deduced by measuring lengths, areas, volumes, angles, performing parallel transport etc in that space.

4. In the case at hand, when we do not equip Σ with any independent a priori metric but we embed Σ into M, the latter equipped with a metric $g_{\alpha\beta}(x)$, the metric on Σ will be the *induced metric*, i.e. the metric induced on Σ by the ambient metric $g_{\alpha\beta}(x)$ (in a way to be described below), and it is this metric that describes the intrinsic geometry of the hypersurface Σ.

5. The reason for insisting on the word "intrinsic" in this context is that when it comes to embedded hypersurfaces there is another aspect of the geometry of Σ that goes beyond its purely intrinsic geometry, namely how it is embedded into the ambient space M, i.e. how it bends inside M. A brief discussion of some aspects of this so-called *extrinsic geometry* of Σ appears in section 18. In this section we will focus on the intrinsic geometry of a hypersurface.

The study of the relation between M-tensors and Σ-tensors has a somewhat different flavour for embeddings Φ and embedded hypersurfaces $\{S(x) = 0\}$, and we will consider both points of view in turn in the following.

15.2 Embeddings: Tangent and Normal Vectors and the Induced Metric

In this section we will look at some aspects of the geometry of hypersurfaces from the point of view of embeddings Φ, i.e. in terms of the parametric description $x^\alpha(y^a)$ of a hypersurface Σ.

First of all, let me start by giving a slightly more precise characterisation of what is meant (or deserves to be called) an embedding. Clearly we want to impose some regularity conditions on Φ as for example the map which sends all of Σ to a single point $x \in M$ might be entertaining to contemplate but does not quite capture what one has in mind when one thinks of hypersurfaces.

In practice the conditions we will use are

- that Φ is injective (or one-to-one), i.e. that distinct points in Σ are mapped to distinct points in M,

- and that the Jacobian of Φ, the $(n+1) \times n$ matrix

$$E_a^\alpha = \frac{\partial x^\alpha}{\partial y^a} \quad , \tag{15.15}$$

has maximal rank n.

REMARKS:

1. Strictly speaking, such a map is called an *(injective) immersion*, while an embedding has to satisfy a slightly stronger topological condition, but since we are not concerned with global issues, and since I have not even tried to define what a manifold is (beyond the remarks in section 5.11), it would be ridiculous to worry about such things here and this is more than good enough.

   Actually (and as an aside (of an aside)), if one is just working locally, the first condition is superfluous, i.e. roughly speaking when the Jacobian is non-degenerate, the map Φ can at most lead to a discrete identification of points and is thus locally invertible on its image (some version of the implicit function theorem). However, since (local) injectivity is something we want and will use, we may as well list it explicitly, regardless of whether or not one can prove a theorem that shows that it is implied by some other condition.

2. For our purposes the most important consequence of this definition is that it implies that the images in M of the tangent vector fields ∂_{y^a} to Σ, the vector fields

$$\partial_{y^a} \mapsto E_a = E_a^\alpha \partial_\alpha \equiv \frac{\partial x^\alpha}{\partial y^a} \partial_\alpha \tag{15.16}$$

   (which are tangent to the image $\Phi(\Sigma)$ of Σ in M) are linearly independent. Here "tangent" means that they are tangent to some curve in $\Phi(\Sigma)$, which is evidently the case since one can take the required curve to be the image under Φ of a suitable curve in Σ.

3. We have thus been able to *push forward* the ∂_{y^a} from Σ to M. Such a *push-forward* operation induced by a map Φ is usually denoted by Φ_*, so that we can also write the above as

$$\Phi_*(\partial_{y^a}) = E_a = E_a^\alpha \partial_\alpha \quad . \tag{15.17}$$

4. Since we have not equipped Σ with any other structure than the coordinates y^a, the ∂_{y^a} are the only objects we will be pushing forward to Σ. In fact, as we will see below, it is not even meaningful to try to push forward the differentials dy^a.

Since the E_a^α are linearly independent tangent vectors to (the image of) Σ in M, normal vectors to Σ, i.e. vectors ξ^α orthogonal to Σ, are characterised by

$$g_{\alpha\beta} E_a^\alpha \xi^\beta = E_a^\alpha \xi_\alpha = 0 \quad . \tag{15.18}$$

REMARKS:

1. Note that in order to define normal vectors to Σ we used the metric $g_{\alpha\beta}$ on M. Without a metric on M (or without specifying a metric on M) one can only define the normal covectors ξ_α, characterised by $\xi_\alpha E^\alpha_a = 0$.

2. If ξ^α is a normal vector field, i.e. a vector field on M normal to Σ at points of Σ, then so is $f\xi^\alpha$ for any scalar f (non-zero on Σ).

3. The normal vector is only defined somewhat implicitly through (15.18). We will see below, in section 15.4, that it has a much more concrete description in the case of embedded hypersurfaces.

4. If the normal vector is everywhere timelike on Σ (this statement is independent of the choice of f), then the E_a and therefore all tangent vectors to Σ are spacelike, and then it is reasonable to call Σ a spacelike hypersurface. Assuming that the character of ξ^α (i.e. whether it is timelike, spacelike or null) is constant over Σ, this terminology generalises to

$$\Sigma \text{ is called } \begin{cases} \text{spacelike} & \text{if } \xi^\alpha\xi_\alpha < 0 \\ \text{timelike} & \text{if } \xi^\alpha\xi_\alpha > 0 \\ \text{lightlike or null} & \text{if } \xi^\alpha\xi_\alpha = 0 \end{cases} \tag{15.19}$$

5. I will mostly use the term "null surface" for a surface with a null or lightlike normal vector. The null case is somewhat special and peculiar, and we will occasionally have to treat it separately from the timelike and spacelike case in the following.

6. When Σ is not null, the freedom in the choice of f can be used to normalise the normal vector to unit length ± 1. This normalisation condition determines the normalised normal vector N^α uniquely up to a choice of sign, one possibility being

$$N^\alpha = \frac{\xi^\alpha}{|\xi^\alpha\xi_\alpha|^{1/2}} \quad \Rightarrow \quad N^\alpha N_\alpha = \epsilon = \begin{cases} -1 & \text{if } \Sigma \text{ is spacelike} \\ +1 & \text{if } \Sigma \text{ is timelike} \end{cases} \tag{15.20}$$

One common convention for fixing the sign ambiguity in the case of an embedded hypersurface $\Sigma = \{S(x) = 0\}$ will be mentioned in section 15.4.

One of the main advantages of the parametric (embedding) description of a hypersurface is that it is utterly straightforward to determine the *induced metric* $h_{ab}(y)$ on Σ, i.e. the metric on Σ induced by a metric $g_{\alpha\beta}(x)$ on M. This is simply obtained by restricting the metric to Σ (better, to its image $\Phi(\Sigma)$ in M), and also restricting the displacements dx^α to displacements in (the image of) Σ,

$$\begin{aligned} ds^2|_\Sigma &= g_{\alpha\beta}(x)dx^\alpha dx^\beta|_\Sigma \\ &= g_{\alpha\beta}\frac{\partial x^\alpha}{\partial y^a}\frac{\partial x^\beta}{\partial y^b}dy^a dy^b \equiv h_{ab}(y)dy^a dy^b \quad . \end{aligned} \tag{15.21}$$

Thus the induced metric is

$$h_{ab}(y) = g_{\alpha\beta}(x(y)) \frac{\partial x^\alpha}{\partial y^a}(y) \frac{\partial x^\beta}{\partial y^b}(y) \quad . \tag{15.22}$$

In terms of the tangent vectors E_a (15.16) the induced metric can be written as (and determined from)

$$h_{ab} = g_{\alpha\beta} \, E^\alpha_a E^\beta_b \quad . \tag{15.23}$$

<u>Remarks:</u>

1. While $g_{\alpha\beta}$ is a $(0,2)$-tensor under space-time coordinate transformations (and evidently a scalar under transformations of the coordinates y^a on Σ), h_{ab} is now a $(0,2)$-tensor on Σ, i.e. under coordinate transformations of the y^a, while it has become a *scalar* under space-time coordinate transformation (as is evident from the space-time contractions in (15.23)).

2. We see that here we have been able to *pull back* (restrict) a tensor on M to a tensor on Σ, an operation usually denoted by Φ^*, so that one also frequently writes this as
$$h_{ab} = (\Phi^* g)_{ab} = g_{\alpha\beta} \, E^\alpha_a E^\beta_b \quad . \tag{15.24}$$

3. This induced metric is non-degenerate when Σ is spacelike or timelike, but turns out to be degenerate in the null case. Intuitively this degeneracy in the null case is reasonably obvious (once one has come to terms with the uniniutive properties of null vectors), since a null vector is normal to itself, and therefore a normal vector that is null is also tangent to the surface. The induced metric is then necessarily degenerate in that null tangent direction. This will be discussed in more detail in section 17.2.

This is really all we need and will make use of in the following, while for the restriction of other tensor fields from M to Σ we will principally use the formulation of embedded hypersurfaces rather than that of embeddings of hypersurfaces.

However, in order to better understand why e.g. the operation of pulling back a metric, described above, works so simply, and if or how this can be extended to other tensor fields, it is useful, even though not strictly necessary, and certainly not indispensable for the following, to look at this from a slighly more general perspective (and we will do this in section 15.3 below).

15.3 Embeddings and Pull-Backs

Given an embedding $\Phi : \Sigma \to M$, *pull-back* refers to the operation of restricting (pulling back) tensors on M to tensors on Σ. The simplest prototype of this kind of pull-back operation is the restriction or pull-back of a function (scalar) f on M,

$$f: \quad M \to \mathbb{R} \ , \tag{15.25}$$

from M to a function on Σ. Thinking temporarily about embedded hypersurfaces $\Sigma \subset M$ this ought to be straightforward as the restriction to Σ clearly defines a scalar on Σ,

$$f|_\Sigma: \quad \Sigma \to \mathbb{R} \ . \tag{15.26}$$

In terms of the embedding Φ this can be phrased as the statement that the embedding map Φ can be used to pull back the function f on M to a function $\Phi^* f$ on Σ defined by

$$\begin{aligned} \Phi^* f: \quad & \Sigma \to \mathbb{R} \\ & (\Phi^* f)(y) = f(\Phi(y)) \ . \end{aligned} \tag{15.27}$$

Now let us move on from scalars to vectors and covectors. Thinking of covectors as linear functions on vectors, it is clear that upon restiction a covector field on M to Σ one obtains a covector field on Σ since its action on any vector at $x \in \Sigma \subset M$ is well-defined, therefore in particular its action on vectors tangent to Σ (which is all that is required to make it a well-defined covector on Σ). In equations this amounts to the statement that if U_α is a covector field on M, then it can be pulled back to a covector field u_a on Σ via

$$u_a = (\Phi^* U)_a = \frac{\partial x^\alpha}{\partial y^a} U_\alpha = E^\alpha_a U_\alpha \ . \tag{15.28}$$

This is indeed (rather evidently now) a covector field on Σ, i.e. transforms as such (while it has become a scalar under coordinate transformations in M). This construction can also be understood in terms of the differentials dx^α and the restriction of the generally covariant object $U_\alpha dx^\alpha$. Just as in our discussion of the induced metric, one can simply restrict the dx^α to Σ to obtain

$$U_\alpha dx^\alpha|_\Sigma = U_\alpha E^\alpha_a dy^a \equiv u_a dy^a \ . \tag{15.29}$$

In the same way one can pull back higher rank covariant tensor fields $U_{\alpha\ldots\beta}$ on M to Σ,

$$(\Phi^* U)_{a\ldots b} = E^\alpha_a \ldots E^\beta_b \, U_{\alpha\ldots\beta} \ . \tag{15.30}$$

A special case of this is the pull-back of the (covariant components of the) metric (15.24).

Characteristic features of hypersurfaces, and what one can and cannot do on them, arise from the fact that the Jacobian E^α_a of Φ is not a square matrix and is therefore not invertible even when it has maximal rank (as we assumed). We had used this before to push forward vectors from Σ to M (the map $\partial_{y^a} \to E_a$ in (15.17)) and we have now been able to use it to pull back covectors from M to Σ. However, because of the non-invertibility of the Jacobian, neither can we use it on the nose to push forward covectors on Σ, or their basis dy^a (I will not dwell on this, though), nor can we use it (all by itself) to restrict (pull back) vectors on M to vectors on Σ.

Indeed, given a vector field $V^\alpha(x)$ on M its restriction to Σ or $\Phi(\Sigma)$ is not all by itself a vector field there because it need not be tangent to Σ (or $\Phi(\Sigma)$). This can be rectified by projecting out the components normal to Σ but this requires a metric, whereas the pull-back of covariant tensors did not require this. I consider this projection procedure to be somewhat simpler and more transparent from the "embedded hypersurface" point of view, and we will discuss this in section 16.1.

I want to conclude this section with some (even less indispensable) remarks on the generality of the pull-back procedure and the difference between covariant and contravariant tensor fields with respect to this operation:

1. Note that the pull-back of a scalar, as defined in (15.27), would be well-defined and unambiguous even if Φ were not injective, as the pull-back $\Phi^*(f)$ would then just happen to assign to two points y_1 and y_2 with $\Phi(y_1) = \Phi(y_2)$ the same value of the function,

   $$\Phi(y_1) = \Phi(y_2) \quad \Rightarrow \quad (\Phi^* f)(y_1) = (\Phi^* f)(y_2) \ , \tag{15.31}$$

   something that is unproblematic.

   Thus, more generally whenever one has some (suitably differentiable, say) map

   $$F: \quad N \to M \tag{15.32}$$

   between two spaces N and M, functions on M can always be pulled back to functions on N via

   $$F^* f(n) = f(F(n)) \tag{15.33}$$

   for $m \in M, n \in N$.

2. On the other hand, in general functions cannot be pushed forward from N to M, not even onto the image $F(N) \subset M$ of N: if F is not injective, $m = F(n_1) = F(n_2)$, say, which point $n_k \in N$ should one choose to assign a value of the (would-be) pushed forward function $F_* f$ to $m \in M$: $(F_* f)(m) =?$.

3. More generally, covariant tensors can always be pulled back under arbitrary (differentiable) maps by precisely the same procedure and formulae (15.30) as in the case of embeddings. On the other hand, we already saw above that even with an embedding one cannot simply pull back vector fields (or other contravariant tensor fields).

4. This highlights a crucial distinction between covariant and contravariant tensors, the former behaving in a natural (functorial) way under maps between spaces and the composition of maps.

The ability to pull back covariant tensors endows these tensors with a crucial operation that is not available to the contravariant ones. *It is difficult to overemphasize the importance of this advantage.*[33]

5. This is also one aspect of the naturality of the calculus of differential forms, based on totally anti-symmetric covariant tensors, briefly mentioned in sections 4.6, 4.8 and 5.5.

6. This crucial distinction between covariant and contravariant tensors did not appear in our general discussion of tensors in section 4.3, because we were dealing with coordinate transformations $x^\alpha(y^\mu)$ on M. These can be thought of as (local) diffeomorphisms)

$$\Phi : \quad M \to M \tag{15.34}$$

or

$$\Phi : \quad U \subset M \to \Phi(U) \subset M \quad , \tag{15.35}$$

i.e. suitably differentiable (smooth) and (locally) invertible maps. In that case, the push-forward is as well-defined as the pull-back since one can set $\Phi_* = (\Phi^{-1})^*$, and therefore both covariant and contravariant tensors could be transformed back and forth between the coordinate systems x^α and y^μ.

7. In the case at hand, where $\Phi : \Sigma \to M$ is locally given by $x^\alpha(y^a)$, a priori the relation between the x^α and the y^a is evidently less democratic. The simple rule of thumb regarding what one can and cannot do is:

   - What you can do with E^α_a by contracting indices you are allowed to do.
   - If what you want to do would require the inverse of that matrix, or at least something with the opposite index structure, you cannot do it (or at least not without using some additional structure like a metric).

   Nevertheless, for embeddings into spaces equipped with a metric the crucial distinction between pull-backs and push-forwards and between covariant and contravariant tensors is blurred by two facts, namely

   - by the assumption that (locally) Φ is injective, i.e. invertible on its image,
   - and crucially by the fact that with the additional structure of a metric on M we can in any case freely convert contravariant into covariant tensors and vice-versa.

Thus in the following we can and will proceed without worrying too much about these matters, and perhaps pragmatically speaking the only benefit of having suffered through this section is that you may have gained a better understanding of why we can get away

[33] T. Frankel, *The Geometry of Physics* (2nd edition), section 2.7a.

with this in the case at hand, i.e. for embeddings into space-times equipped with a metric.

15.4 Embedded Hypersurfaces and Normal Vectors

In the following, in order not to have to introduce separate coordinates on Σ from the outset, for the most part we will use the 2nd description of a hypersurface, i.e. we will work with embedded hypersurfaces defined by (15.3)

$$\Sigma = \{x \in M : \; S(x) = 0\} \tag{15.36}$$

for some function $S(x)$ on M, rather than with embeddings (and we will see later how e.g. the induced metric can be described and recovered from this point of view).

Implicitly the characterisation (15.36) of Σ implies not only that $S(x) = 0$ on Σ but that $\{S(x) = 0\}$ actually defines a codimension 1 hypersurface, i.e. that $S(x)$ is not zero when one moves off Σ. We will furthermore choose the defining function $S(x)$ in such a way that it has a 1st order zero on S. This is not necessary in order to define Σ, but it avoids unnecessary complications (why would one want to define a horizontal plane in $\mathbb{R}^3$ (with coordinates (x, y, z), say) by $z^2 = 0$ rather than by $z = 0$?).

Then at any point $x \in \Sigma$ the gradient $\partial_\alpha S(x)$ is not zero on Σ,

$$(\partial_\alpha S)|_{S=0} \neq 0 \quad . \tag{15.37}$$

One advantage of this description of Σ and choice of S is that one can now at once, and very concretely, describe the *normal vectors* to the hypersurface Σ.

Indeed, if the hypersurface is described by $S(x^\alpha) = 0$, then by definition S does not vary along directions in (tangent to) Σ. Thus a vector field V^α tangent to Σ at points of Σ is such that it satisfies

$$x \in \Sigma \quad , \quad V^\alpha(x) \;\text{ tangent to }\; \Sigma \;\text{ at }\; x \quad \Rightarrow \quad V^\alpha(x)\partial_\alpha S(x) = 0 \quad . \tag{15.38}$$

This means that any such tangent vector is orthogonal to the gradient vector $g^{\alpha\beta}\partial_\beta S$ which is non-zero on Σ by assumption. Thus on Σ this gradient vector field is normal to Σ (and actually normal to the family of hypersurfaces Σ defined by $S(x) = \text{const}$).

As in section 15.2,

$$\Sigma \;\text{ is called }\; \begin{cases} \text{spacelike} & \text{if } g^{\alpha\beta}\,\partial_\alpha S\,\partial_\beta S < 0 \\ \text{timelike} & \text{if } g^{\alpha\beta}\,\partial_\alpha S\,\partial_\beta S > 0 \\ \text{null} & \text{if } g^{\alpha\beta}\,\partial_\alpha S\,\partial_\beta S = 0 \end{cases} \tag{15.39}$$

everywhere on Σ. If one introduces $S(x)$ as a new coordinate, as described in section 15.1, then by the usual tensorial transformation rules the norm of the gradient of S is simply the component

$$g^{\alpha\beta}\partial_\alpha S\,\partial_\beta S = g^{SS} \tag{15.40}$$

of the inverse metric. In particular this shows that in these coordinates the locus where a member of the family of surfaces Σ_S of constant S becomes null is determined by

$$\Sigma_S \quad \text{is null} \quad \Leftrightarrow \quad g^{SS} = 0 \ . \tag{15.41}$$

Evidently, with $g^{\alpha\beta}\partial_\beta S$ also any vector field of the form $\xi^\alpha = f\, g^{\alpha\beta}\partial_\beta S$ for some scalar $f(x)$ (non-zero on Σ) is normal to Σ,

$$\xi^\alpha = f\, g^{\alpha\beta}\partial_\beta S \quad \Rightarrow \quad g_{\alpha\beta}\xi^\alpha V^\beta = 0 \quad \forall \ V \ \text{ tangent to } \ \Sigma \ , \tag{15.42}$$

and in the case that Σ is spacelike or timelike this freedom in the rescaling of the normal vector can be used to normalise it in such a way that $N_\alpha = f\partial_\alpha S$ has unit length $\epsilon = \pm 1$. This determines N^α uniquely up to a choice of sign. Explicitly, the choice

$$N_\alpha = \epsilon \frac{\partial_\alpha S}{|g^{\alpha\beta}\partial_\alpha S\partial_\beta S|^{1/2}} \tag{15.43}$$

is such that

$$N^\alpha N_\alpha = \epsilon = \begin{cases} -1 & \text{if } \Sigma \text{ is spacelike} \\ +1 & \text{if } \Sigma \text{ is timelike} \end{cases} \tag{15.44}$$

and such that N^α points in the direction of increasing S,

$$N^\alpha \partial_\alpha S = |g^{\alpha\beta}\partial_\alpha S\partial_\beta S|^{1/2} > 0 \ . \tag{15.45}$$

This is a common but by no means mandatory or universal sign convention.

REMARKS:

1. For spacelike hypersurfaces of constant time, say, given in suitable coordinates $x^\alpha = (t, x^k)$ by $S(t, x^k) = t - t_0$, this convention is such that N^α is future pointing (but it would be past-oriented if one chose S to decrease towards the future, e.g. $S(t, x^k) = t_0 - t$).

2. In the null case there is in general no such preferred choice of normal vector, because any normal vector ξ^α satisfies $\xi^\alpha \xi_\alpha = 0$. We will return to that case in section 17.1.

3. As noted in section 15.2, if the hypersurface is given in parametric form $x^\alpha = x^\alpha(y^a)$, normal vectors ξ^α are characterised by the condition

$$\xi_\alpha \frac{\partial x^\alpha}{\partial y^a} = 0 \quad \Leftrightarrow \quad \xi_\alpha E^\alpha_a = 0 \ . \tag{15.46}$$

   Thus the defining function S is related to the parametric description by the condition that its gradient covector field $\partial_\alpha S$ is in the kernel of the Jacobian,

$$E^\alpha_a \ \partial_\alpha S = 0 \ . \tag{15.47}$$

A priori, given a hypersurface $\Sigma \subset M$, the normal vector field is only defined on Σ, not on all of M and not even in a neighbourhood of Σ. It is frequently desirable in the timelike or spacelike case, however, to have the normalised normal vector N^α defined at least in a neighbourhood of Σ. There are two standard ways to achieve this:

- If, as in the situation considered here, the hypersurface Σ is specified by $S(x^\alpha) = 0$, and $S(x^\alpha) = c$ defines a family of hypersurfaces Σ_c around Σ, then the normal vector field N^α is automatically defined in a neighbourhood of Σ.
- If only Σ is given, and thus N^α is initially only defined on Σ, there is a natural way to extend it to a neighbourhood of Σ when Σ is not null.

  Indeed, noting that $N^\alpha N_\alpha = \epsilon$ is precisely the correct normalisation condition for the tangent vector to an affinely parametrised timelike or spacelike geodesic, one way to extend N^α off Σ is to consider the geodesics in M which emanate from Σ with initial velocity (tangent vector) $N^\alpha(x)$ for $x \in \Sigma$, and to define $N^\alpha(x)$ in a neighbourhood of Σ (chosen sufficiently small so that geodesics do not intersect) to be the tangent vector field to this family (congruence) of geodesics.

Either way, the normal vector now satisfies $N_\alpha N^\alpha = \epsilon$ in a neighbourhood of Σ.

This prescription does not work for null hypersurfaces because for a null hypersurface a normal vector has the, for a null vector tpyical counter-intuitive, property that it is also tangent to Σ and thus generates null curves in Σ rather than away from Σ - these turn out to be geodesics, a fact which is interesting in its own right and which will be established and explored in section 17.2.

15.5 Hypersurface Orthogonality and Frobenius Integrability

The unnormalised gradient normal vector field $g^{\alpha\beta}\partial_\beta S$ to a hypersurface Σ defined by $S(x) = 0$ (or $S(x) = \text{const.}$) satisfies the equation

$$\nabla_\alpha \partial_\beta S - \nabla_\beta \partial_\alpha S = \partial_\alpha \partial_\beta S - \partial_\beta \partial_\alpha S = 0 \quad . \tag{15.48}$$

Conversely if one has a vector field that satisfies

$$\nabla_\alpha \xi_\beta - \nabla_\beta \xi_\alpha = 0 \quad \Leftrightarrow \quad \partial_\alpha \xi_\beta - \partial_\beta \xi_\alpha = 0 \quad , \tag{15.49}$$

then this is the necessary integrability condition for ξ_α to be of the form $\partial_\alpha S$ for some function S. This condition is metric-independent, as it should be. It is well known from standard (vector) calculus that locally this condition is also sufficient (if the curl of a vector field is zero then locally it can be written as a gradient vector field etc.), i.e. one has

$$\partial_\alpha \xi_\beta - \partial_\beta \xi_\alpha = 0 \quad \Rightarrow \quad \text{(locally)} \ \exists S : \quad \xi_\alpha = \partial_\alpha S \quad . \tag{15.50}$$

A general normal vector field $\xi_\alpha = f\partial_\alpha S$ to a hypersurface Σ, in particular usually also the normalised normal vector N^α (when it exists, i.e. when Σ is not null), will not satisfy (15.49). However, it satisfies a generalised equation of this type. Namely, it follows from

$$\xi_\alpha = f\,\partial_\alpha S \quad \Rightarrow \quad \nabla_\alpha \xi_\beta - \nabla_\beta \xi_\alpha = \nabla_\alpha f\,\partial_\beta S - \nabla_\beta f\,\partial_\alpha S \tag{15.51}$$

that ξ_α satisfies

$$\nabla_{[\alpha}\xi_{\beta]} = (\nabla_{[\alpha}\log f)\xi_{\beta]} \tag{15.52}$$

While this is true, in this form it is not a particularly useful characaterisation of hypersurface orthogonality because given ξ it may not be straightforward to see if such a function f exists or not. A more useful condition is the integrability condition implied by this, namely

$$\nabla_{[\alpha}\xi_{\beta]} = (\nabla_{[\alpha}\log f)\xi_{\beta]} \quad \Rightarrow \quad \xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0 \;\; , \tag{15.53}$$

which itself is a trivial consequence of $\xi_{[\alpha}\xi_{\beta]} = 0$. This is the necessary integrability condition for a vector field to be of the form $\xi_\alpha = f\partial_\alpha S$ for some functions f and S, and the advantage of this condition is that it depends only on ξ (and can thus be checked if one is just given ξ).

In strict analogy with the above story for gradient vectors, this condition is also sufficient to establish that locally ξ_α can be written as $\xi_\alpha = f\partial_\alpha S$ for some functions f and S,

$$\xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0 \quad \Rightarrow \quad \text{(locally)}\;\; \exists S, f: \quad \xi_\alpha = f\,\partial_\alpha S \;\; . \tag{15.54}$$

Since $\xi_\alpha = f\partial_\alpha S$ is precisely the statement that ξ^α is orthogonal to the family of hypersurfaces $S(x) = \text{const}$, the condition

$$\xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0 \quad \Leftrightarrow \quad \xi_{[\alpha}\partial_\beta\xi_{\gamma]} = 0 \tag{15.55}$$

is known as the *hypersurface orthogonality condition* and a vector field that satisfies (15.55) is called *hypersurface orthogonal.*

Remarks:

1. The assertion (15.54) is known as *Frobenius' theorem*, and the hypersurface orthogonality condition is therefore also known as the *Frobenius integrability condition.*

2. If we do not just have $\xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0$ but actually $\nabla_{[\beta}\,\xi_{\gamma]} = 0$, then, as we saw before, we can draw the stronger conclusion

$$\nabla_{[\beta}\,\xi_{\gamma]} = 0 \quad \Rightarrow \quad \text{(locally)}\;\; \exists S: \quad \xi_\alpha = \partial_\alpha S \;\; , \tag{15.56}$$

which is a fortiori orthogonal to the hypersurfaces of constant S.

16 Hypersurfaces II: Intrinsic Geometry of non-Null Hypersurfaces

16.1 Projectors for non-Null Hypersurfaces and the Induced Metric

In the case of spacelike or timelike hypersurfaces Σ, with the normalised normal vectors N^α at our disposal we can now construct the induced metric from the metric $g_{\alpha\beta}$ on the ambient space M. More generally, we will construct projectors that allow us to project tensors on M restricted to Σ onto directions (co-)tangent to Σ.

Thus, with $N^\alpha N_\alpha = \epsilon = \pm 1$, consider the tensor $h_{\alpha\beta}$ defined on (or in a neighbourhood of) Σ by

$$h_{\alpha\beta} = g_{\alpha\beta} - \epsilon N_\alpha N_\beta \quad . \tag{16.1}$$

This tensor has the following characteristic properties:

1. It is orthogonal to N^α,

$$N^\alpha h_{\alpha\beta} = 0 \quad , \quad h_{\alpha\beta} N^\beta = 0 \quad . \tag{16.2}$$

   Indeed,

$$h_{\alpha\beta} N^\beta = g_{\alpha\beta} N^\beta - \epsilon N_\alpha N_\beta N^\beta = N_\alpha - \epsilon^2 N_\alpha = 0 \quad ; \tag{16.3}$$

2. For vectors V^α orthogonal to N^α, i.e. tangent to Σ, the scalar product with respect to $h_{\alpha\beta}$ is identical to that with respect to $g_{\alpha\beta}$.

$$V^\beta N_\beta = 0 \quad \Rightarrow \quad h_{\alpha\beta} V^\beta = g_{\alpha\beta} V^\beta \quad . \tag{16.4}$$

These two properties together imply that essentially $h_{\alpha\beta}$ (restricted to Σ) is the metric induced on Σ by $g_{\alpha\beta}$. The precise relation to the induced metric (15.23)

$$h_{ab} = g_{\alpha\beta}\, E^\alpha_a E^\beta_b \quad . \tag{16.5}$$

provided by the parametric (embedding) description of a hypersurface is given by

$$\begin{aligned} h_{ab} &= g_{\alpha\beta}\, E^\alpha_a E^\beta_b \\ &= (h_{\alpha\beta} + \epsilon N_\alpha N_\beta) E^\alpha_a E^\beta_b \\ &= h_{\alpha\beta} E^\alpha_a E^\beta_b \quad , \end{aligned} \tag{16.6}$$

where the 2nd equality follows from the fact that N^α is orthogonal to the E^α_a.

Remarks:

1. Thus, given the covariant tensor $h_{\alpha\beta}$ orthogonal to N^α, we can equivalently think of it as the tensor h_{ab} on Σ. The difference between the two are essentially only that

- $h_{\alpha\beta}(x)$ as a matrix is degenerate (as it has the null vector N^α), while h_{ab} is non-degenerate;
- $h_{\alpha\beta}$ is an M-tensor of type $(0,2)$ and a Σ-scalar while h_{ab} is an M-scalar and a Σ-tensor of type $(0,2)$.

2. In particular, while it makes sense to write $h^{\alpha\beta}$, as usual simply defined by

$$h^{\alpha\beta} = g^{\alpha\gamma}g^{\beta\delta}h_{\gamma\delta} \quad , \tag{16.7}$$

this $h^{\alpha\beta}$ is *not* the inverse of the induced metric $h_{\alpha\beta}$ (indeed, as we just noted, $h_{\alpha\beta}$ does not even have an inverse). Rather, one finds

$$h^{\alpha\beta}h_{\beta\gamma} = \delta^\alpha_\gamma - \epsilon N^\alpha N_\gamma \quad , \tag{16.8}$$

so it is only the inverse on the orthogonal complement to N^α, as expected. Note also that (16.8) implies

$$g^{\alpha\beta}h_{\alpha\beta} = h^{\alpha\beta}h_{\alpha\beta} = 3 \quad , \tag{16.9}$$

as behoves a 3-dimensional metric. Below we will reinterpret these equations in terms of projectors into the directions orthogonal to N^α.

3. We see in this example (and we will see and make use of this more generally below) that on covariant tensors that are orthogonal to N^α in the sense that any contraction with N^α is zero, we can convert space-time indices to hypersurface indices using the E^α_a, i.e. we can convert such tensors into tensors on Σ. For such tangential space-time tensors this conversion does not lose any information (i.e. one is not throwing away any components).

4. Dual to the relation (16.6) one has

$$h^{\alpha\beta} = h^{ab}E^\alpha_a E^\beta_b \tag{16.10}$$

which is manifestly orthogonal to N_α and (up to the conversion of indices) acts in the same way on tangent covectors as h^{ab}.

5. Using this expression for $h^{\alpha\beta}$, we can write (and interpret) the defining relation $h_{\alpha\beta} = g_{\alpha\beta} - \epsilon N_\alpha N_\beta$ for $h_{\alpha\beta}$ as a completeness relation for the linearly independent vectors N^α and E^α_a, namely

$$g^{\alpha\beta} = h^{ab}E^\alpha_a E^\beta_b + \epsilon N^\alpha N^\beta \quad . \tag{16.11}$$

6. In the terminology of section 15.3, h_{ab} is the pull-back of $g_{\alpha\beta}$ or $h_{\alpha\beta}$ to Σ, while $h^{\alpha\beta}$ is the push-forward of h^{ab} from Σ to M.

The tensor $h_{\alpha\beta}$ also provides us with the projectors allowing us to project a general tensor onto its tangential components to Σ. Indeed, first of all we can reinterpret the result (16.8) as the statement that the tensors

$$h^{\alpha}_{\ \beta} = g^{\alpha\gamma}h_{\gamma\beta} = \delta^{\alpha}_{\ \beta} - \epsilon N^{\alpha}N_{\beta} \tag{16.12}$$

are projection operators,

$$h^{\alpha}_{\ \beta}h^{\beta}_{\ \gamma} = h^{\alpha}_{\ \gamma} \ \ . \tag{16.13}$$

More precisely, as a consequence of the properties of $h_{\alpha\beta}$ already established, in particular the orthogonality (16.2), they are projection operators onto vectors tangent to Σ:

$$\begin{aligned} V^{\alpha}N_{\alpha} = 0 \quad &\Rightarrow \quad h^{\alpha}_{\ \beta}V^{\beta} = V^{\alpha} \\ V^{\alpha} \sim N^{\alpha} \quad &\Rightarrow \quad h^{\alpha}_{\ \beta}V^{\beta} = 0 \ \ , \end{aligned} \tag{16.14}$$

while it follows from (16.12) that $\epsilon N^{\alpha}N_{\beta}$ is a projector onto the orthogonal complement, namely the normal direction,

$$\begin{aligned} V^{\alpha}N_{\alpha} = 0 \quad &\Rightarrow \quad \epsilon N^{\alpha}N_{\beta}V^{\beta} = 0 \\ V^{\alpha} = fN^{\alpha} \quad &\Rightarrow \quad \epsilon N^{\alpha}N_{\beta}V^{\beta} = fN^{\alpha}(\epsilon N_{\beta}N^{\beta}) = V^{\alpha} \ \ . \end{aligned} \tag{16.15}$$

These projectors now allow one to map / project an arbitrary covariant or contravariant space-time tensor field onto its components (co-)tangent to Σ:

- E.g. for a vector V^{α} one has

  $$V^{\alpha} \mapsto v^{\alpha} = h^{\alpha}_{\ \beta}V^{\beta} \quad , \quad v^{\alpha}N_{\alpha} = 0 \ \ . \tag{16.16}$$

  Such a v^{α} which is tangent to Σ must be a linear combination of the E^{α}_{a}, i.e. it is related to some vector v^{a} on Σ by

  $$v^{\alpha} = E^{\alpha}_{a}v^{a} \ \ . \tag{16.17}$$

- For a covariant 2-tensor, say, one has

  $$B_{\alpha\beta} \mapsto b_{\alpha\beta} \equiv h_{\alpha}^{\ \gamma}h_{\beta}^{\ \delta}B_{\gamma\delta} \ \ , \tag{16.18}$$

  where $b_{\alpha\beta}$ satisfies

  $$\begin{aligned} V^{\alpha}N_{\alpha} = W^{\alpha}N_{\alpha} = 0 \quad &\Rightarrow \quad b_{\alpha\beta}V^{\alpha}W^{\beta} = B_{\alpha\beta}V^{\alpha}W^{\beta} \\ V^{\alpha} \sim N^{\alpha} \quad &\Rightarrow \quad b_{\alpha\beta}V^{\beta} = 0 \ \ . \end{aligned} \tag{16.19}$$

  A covariant tensor which is tangential in this sense can equivalently be regarded as a covariant tensor on Σ via

  $$b_{ab} = E^{\alpha}_{a}E^{\beta}_{b}b_{\alpha\beta} \ \ . \tag{16.20}$$

We see that the projection procedure is quite straightforward and simple in terms of the normal vector provided by the defining function S of an embedded hypersurface. Using (16.10), we can also write and interpret this projection in terms of the embedding data $x^\alpha(y^a)$, in particular the E^α_a and the induced metric h_{ab}. Let us take a look at that in the case of a vector field. Then we can write the projection (16.16) as

$$v^\alpha = h^\alpha_{\ \beta} V^\beta = h^{\alpha\beta} g_{\beta\gamma} V^\gamma = E^\alpha_a h^{ab} E^\beta_b g_{\beta\gamma} V^\gamma \quad . \tag{16.21}$$

Taking this apart, we see that from the embedding point of view the projection procedure (which is a single step procedure when expressed in terms of $h^\alpha_{\ \beta}$) consists of the following sequence of steps:

- use the space-time metric $g_{\beta\gamma}$ to convert the vector field V^γ into the covector field V_β,
$$V_\beta = g_{\beta\gamma} V^\gamma \tag{16.22}$$
- use E^β_b to pull back V_β to a covector field v_b on Σ,
$$v_b = E^\beta_b V_\beta \tag{16.23}$$
- use the inverse h^{ab} of the induced metric to turn this into a vector field v^a on Σ,
$$v^a = h^{ab} v_b \tag{16.24}$$
- Finally use E^α_a to push this forward to a tangent vector field v^α on the image $\Phi(\Sigma) \subset M$,
$$v^\alpha = E^\alpha_a v^a \quad . \tag{16.25}$$

This is a perfectly logical sequence of operations, but you may now understand why I said in section 15.3 that "I consider this projection procedure to be somewhat simpler and more transparent from the "embedded hypersurface" point of view".

16.2 Intrinsic = Projected Covariant Differentiation

Given the induced metric h_{ab} on Σ, one has the associated canonical Levi-Civita covariant derivative (i.e. the unique torsion-free metric-compatible connection) at one's disposal to define covariant derivatives of Σ-tensor fields. Let us temporarily denote this *intrinsic covariant derivative* by $\nabla^{(h)}$, so that e.g.

$$\nabla^{(h)}_a v^b = \partial_a v^b + \Gamma^{(h)b}{}_{ac} v^c \quad , \quad \nabla^{(h)}_a v_b = \partial_a v_b - \Gamma^{(h)c}{}_{ab} v_c \ , \tag{16.26}$$

where

$$\Gamma^{(h)b}{}_{ac} = \tfrac{1}{2} h^{bd} (\partial_c h_{ad} + \partial_a h_{cd} - \partial_d h_{ac}) \quad . \tag{16.27}$$

On the other hand, given a space-time vector field V^α that is tangent to Σ (on Σ), i.e.

$$v^\alpha \equiv h^\alpha_\beta V^\beta = V^\alpha \quad \Leftrightarrow \quad V^\alpha = E^\alpha_a v^a \quad , \tag{16.28}$$

we can define its *projected covariant derivative* along Σ by taking its covariant derivative and then projecting it to Σ. Let us denote this covariant derivative by $\bar{\nabla}$, so that e.g.

$$\bar{\nabla}_\alpha v_\beta = h_\alpha^\gamma h_\beta^\delta \nabla_\gamma v_\delta \quad . \tag{16.29}$$

Since this is now a projected tensor, it can be pulled back without loss of information to Σ, i.e. pragmatically speaking we can convert its indices using E^α_a,

$$\bar{\nabla}_\alpha v_\beta \quad \longrightarrow \quad \Phi^*(\bar{\nabla} v)_{ab} = E^\alpha_a E^\beta_b \bar{\nabla}_\alpha v_\beta \quad . \tag{16.30}$$

Given that we now appear to have two natural notions of differentiation of Σ-tensor fields, the obvious question that arises at this point is what is the relation between the two, and the (reassuring, and perhaps not too surprising) answer is that they are equal,

$$E^\alpha_a E^\beta_b \bar{\nabla}_\alpha v_\beta = \nabla^{(h)}_a v_b \quad . \tag{16.31}$$

The quickest way to see this is to prove that the projected covariant derivative is symmetric (torsion-free, covariant derivatives commute on scalars) and compatible with the induced metric. The first property is obvious since

$$[\bar{\nabla}_\alpha, \bar{\nabla}_\beta] f = h_\alpha^\gamma h_\beta^\delta [\nabla_\gamma, \nabla_\delta] f = 0 \quad , \tag{16.32}$$

and the second property follows from

$$h_{\alpha\beta} = g_{\alpha\beta} - \epsilon N_\alpha N_\beta \quad \Rightarrow \quad \nabla_\alpha h_{\beta\gamma} = -\epsilon((\nabla_\alpha N_\beta) N_\gamma + N_\beta(\nabla_\alpha N_\gamma)) \tag{16.33}$$

since this expression vanishes after projection into the directions orthogonal to N_α.

Thus for projected tensors the projected covariant derivative is equal to the intrinsic covariant derivative (up to pull-back), and the obvious next questions are e.g. "what are the normal components of the covariant derivative of a projected tensor?" or "what are the projections of the covariant derivative of a non-tangential tensor, i.e. a tensor with a normal component"? These are legitimate and interesting questions. However, they go beyond the intrinsic geometry of hypersurfaces and bring us into the realm of extrinsic geometry, a subject that will be addressed (briefly) in section 18.

16.3 Integration on non-Null Hypersurfaces and the Gauss Theorem

Let Σ be a non-null hypersurface, with local coordinates y^a, and h_{ab} a metric on Σ, e.g. the metric induced by a metric $g_{\alpha\beta}$ on the ambient space-time M. Then $\sqrt{h} d^n y$, with

$$h := |\det h_{ab}| \tag{16.34}$$

is an invariant volume element on Σ and integration of Σ-scalars f can be defined by

$$\int_\Sigma f := \int \sqrt{h} d^n y \; f(y) \quad . \tag{16.35}$$

Integrals over hypersurfaces arise in particular from applications of the Gauss theorem (or Gauss-Stokes theorem) which allows one to express the volume integral over some space-time region $\mathcal{V}$ of a covariant divergence as an integral over the boundary hypersurface

$$\Sigma = \partial\mathcal{V} \tag{16.36}$$

of that space-time region. This is usually written as something like

$$\int_\mathcal{V} \sqrt{g} d^D x \; \nabla_\alpha J^\alpha = \oint_{\partial\mathcal{V}} d\sigma_\alpha J^\alpha \quad , \tag{16.37}$$

where $d\sigma_\alpha$ is some oriented surface (volume/area/...) element. The proof of this identity can be reduced to the proof of the corresponding statement in standard multivariable calculus in Euclidean space by making use of the fact, already noted in (5.61), that

$$\int_\mathcal{V} \sqrt{g} d^D x \; \nabla_\alpha J^\alpha = \int_\mathcal{V} d^D x \; \partial_\alpha(\sqrt{g} J^\alpha) \quad . \tag{16.38}$$

is an ordinary total derivative. Assuming momentarily that we are working in adapted coordinates $x^\alpha = (S, x^i)$ in which $\Sigma_S = \partial\mathcal{V}$ is a surface $S(x) = c$ of constant S, we can write this somewhat more explicitly as

$$\begin{aligned} \int_\mathcal{V} \sqrt{g} d^D x \; \nabla_\alpha J^\alpha &= \int_\mathcal{V} d^D x \; (\partial_S(\sqrt{g} J^S) + \ldots) \\ &= \int_{\Sigma_S} d^n x \; \sqrt{g} \; g^{S\alpha} J_\alpha \quad . \end{aligned} \tag{16.39}$$

It thus remains to understand the relation between the surface element

$$d\sigma^\alpha = d^n x \; \sqrt{g} \; g^{S\alpha} \tag{16.40}$$

appearing in this integral, and the intrinsic invariant volume element $\sqrt{h} d^n y$. To that end, we note that in the adapated coordinates (S, x^i) one has (with the sign convention (15.43))

$$N_\alpha \sim \partial_\alpha S \quad , \quad N^\alpha N_\alpha = \epsilon \quad \Rightarrow \quad N^\alpha = \epsilon g^{\alpha S} / \sqrt{\epsilon g^{SS}} \quad . \tag{16.41}$$

Moreover, by the usual cofactor / minor formula for the components of the inverse metric (5.81), one has

$$g^{SS} = \frac{\det h_{ik}}{\det g_{\alpha\beta}} = -\det(h_{ik})/g \quad , \tag{16.42}$$

where $h_{ik} = g_{ik}$ refers to the (ik)-components of the induced metric $h_{\alpha\beta}$ in the adapted coordinates (S, x^i),

$$g_{\alpha\beta} = h_{\alpha\beta} + \epsilon N_\alpha N_\beta \quad \Rightarrow \quad g_{ik} = h_{ik} \quad . \tag{16.43}$$

Therefore we can write the factor $\sqrt{g}g^{S\alpha}$ appearing in (16.39) as

$$\begin{aligned}\sqrt{g}\; g^{S\alpha} &= \sqrt{g}\;\epsilon N^\alpha \sqrt{|g^{SS}|}\\ &= \epsilon\sqrt{|\det(h_{ik})|}N^\alpha \quad .\end{aligned} \tag{16.44}$$

Finally, noting that

$$h_{ab} = E^i_a E^k_b h_{ik} \quad \Rightarrow \quad \sqrt{|\det(h_{ik})|}d^n x = \sqrt{|\det(h_{ab})|}d^n y \tag{16.45}$$

(the tangent E_a having no normal S-component in these coordinates), we conclude that

$$d\sigma^\alpha = d^n x\;\sqrt{g}\; g^{S\alpha} = \epsilon\sqrt{h}d^n y N^\alpha \quad , \tag{16.46}$$

so that we can also write the Gauss theorem in the convenient and transparent form

$$\int_{\mathcal{V}} \sqrt{g}d^D x\;\nabla_\alpha J^\alpha = \epsilon\int_{\Sigma=\partial\mathcal{V}} d^n y\;\sqrt{h}\;N_\alpha J^\alpha \equiv \int_\Sigma d\sigma_\alpha J^\alpha \quad . \tag{16.47}$$

A standard application of this formula is to conserved charges associated to (covariantly) conserved currents, $\nabla_\alpha J^\alpha = 0$, discussed in section 6.8. Indeed, let us consider a space-time volume $\mathcal{V}$ bounded by two spacelike hypersurfaces

$$\partial\mathcal{V} = \{\Sigma_1\} \cup \{-\Sigma_0\} \tag{16.48}$$

(the minus sign indicating that we equip Σ_0 with the opposite orientation to that induced by $\mathcal{V}$ so that e.g. both surfaces Σ_k have future-pointing normal vectors). Then one finds that

$$\begin{aligned}Q_1 - Q_0 &= \int_{\Sigma_1} d\sigma_\alpha J^\alpha - \int_{\Sigma_0} d\sigma_\alpha J^\alpha\\ &= \int_{\mathcal{V}} \sqrt{g}d^D x\;\nabla_\alpha J^\alpha = 0 \quad ,\end{aligned} \tag{16.49}$$

so that (under suitable asymptotic conditions) covariantly conserved currents will lead to conserved charges. Analogously, and somewhat more generally, this shows that if one has a family Σ_c of hypersurfaces, sweeping out a space-time volume $\mathcal{V} = \cup_c \Sigma_c$, the integral

$$Q_c = \int_{\Sigma_c} d\sigma_\alpha J^\alpha \tag{16.50}$$

is independent of c, i.e. the charge in invariant under deformations of the hypersurface.

16.4 Spacelike Hypersurfaces and Stationary vs Static Metrics

One common instance where the issue of hypersurface orthogonality discussed in section 15.5 plays a crucial role is in the distinction between what are known as *stationary metrics* (or space-times, or gravitational fields) versus *static metrics* (or space-times, or gravitational fields). Both terms refer to gravitational fields that are in a suitable sense time-independent, but "static" is a stronger condition than "stationary".

I used the word "static" in connection with the metric (3.22),

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2 d\Omega^2 \ , \tag{16.51}$$

while in the discussion of the Newtonian limit of the geodesic equation in section 3.3 I used the term "stationary" to refer to the condition (3.45) that the coefficients of the metric be time-independent. In general, we will define a metric to be stationary if it has a time-translation invariance, in the sense that one can find coordinates $x^\alpha = (t, x^k)$, say, such that $\xi = \partial_t$ is timelike and that none of the coefficients $g_{\alpha\beta}$ of the metric depend on t,

$$\text{Stationary Metric:} \quad \partial_t g_{\alpha\beta} = 0 \ . \tag{16.52}$$

Thus the general form of a stationary metric, without assuming the existence of any further symmetries, is just

$$\text{Stationary Metric:} \quad ds^2 = g_{\alpha\beta}(x^k)dx^\alpha dx^\beta \qquad (x^\alpha = (t, x^k)) \ . \tag{16.53}$$

This can be stated in a geometrically more invariant way as the condition that the metric admits a timelike Killing vector ξ (cf. the discussion in sections 3.2 and 9.5). Locally, one can then always introduce coordinates such that the Killing vector has the form $\xi = \partial_t$ (see the discussion after (9.59)), so that in these coordinates the symmetry is that of t-translation invariance, as in (16.52). For present purposes this locally equivalent characterisation of the existence of a time-translation symmetry is good enough.

It will nevertheless be useful (even for present purposes) to be able to write the condition (16.52) in a somewhat more covariant form. To that end, note that for a vector field of the form $\xi = \partial_t$ one has (repeating the calculation leading to (5.71))

$$\begin{aligned} \xi = \partial_t \quad &\Rightarrow \quad \nabla_\alpha \xi^\beta = \Gamma^\beta{}_{\alpha t} \\ &\Rightarrow \quad \nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha = \partial_t g_{\alpha\beta} \end{aligned} \tag{16.54}$$

Thus we find that the fact that the metric is t-translation invariant can be characterised covariantly as the statement that $\xi = \partial_t$ satisfies

$$\partial_t g_{\alpha\beta} = 0 \quad \Leftrightarrow \quad \nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha = 0 \ . \tag{16.55}$$

Unsurprisingly, this is the expression we had already found in (5.72).

The metric (16.51) of course has the property that all the metric coefficients are t-independent so it is certainly stationary, but it also has the further property that $\xi = \partial_t$ is hypersurface orthogonal. Indeed, in this case $\xi = \partial_t$ is evidently normal to the constant time hypersurfaces $t - t_0 = 0$.

This need not be the case, however. To set the stage, consider an arbitrary metric written in coordinates $x^\alpha = (t, x^k)$ and let ξ be the vector $\xi = \partial_t$, i.e. $\xi^\alpha = \delta^\alpha_t$. Then its metric dual covector has the covariant components

$$\xi = \partial_t \quad \Rightarrow \quad \xi_\alpha = g_{\alpha\beta}\xi^\beta = g_{\alpha t} \ . \tag{16.56}$$

In particular, ξ_t is also the norm-squared of ξ,

$$\xi_t = g_{tt} = g_{\alpha\beta}\xi^\alpha\xi^\beta = \xi^\alpha\xi_\alpha \quad , \tag{16.57}$$

as can also be seen directly from (16.56),

$$\xi^\alpha\xi_\alpha = \xi^\alpha g_{\alpha t} = \begin{cases} \xi_t & \text{by "pulling down the index"} \\ g_{tt} & \text{by using } \xi^\alpha = \delta^\alpha_t \end{cases} \tag{16.58}$$

It will also be convenient to write (16.56) in the form

$$\xi_\alpha dx^\alpha = g_{\alpha t}dx^\alpha = g_{tt}dt + g_{tk}dx^k = (g_{tt}\partial_\alpha t + g_{tk}\partial_\alpha x^k)dx^\alpha \quad . \tag{16.59}$$

If the metric is such that $g_{tk} = 0$, then clearly

$$g_{tk} = 0 \quad \Rightarrow \quad \xi_\alpha = g_{tt}\partial_\alpha t \tag{16.60}$$

which is of the characteristic form of a hypersurface orthogonal vector field

$$g_{tt}\partial_\alpha t \equiv f\ \partial_\alpha S \quad , \tag{16.61}$$

so that ∂_t is orthogonal to the surfaces of constant t, $S(x^\alpha) = t - t_0$.

In general, a *static metric* is by definition a metric which is stationary and which is such that the vector field $\xi = \partial_t$ generating the time-translation symmetry is hypersurface orthogonal,

$$\text{Static Metric:} \quad \partial_t g_{\alpha\beta} = 0 \quad \text{and} \quad \partial_t \ \text{hypersurface orthogonal} \quad . \tag{16.62}$$

As we have just seen, this will be the case e.g. if $g_{tk} = 0$,

$$\partial_t g_{\alpha\beta} = 0 \quad \text{and} \quad g_{tk} = 0 \quad \Rightarrow \quad \text{static} \quad . \tag{16.63}$$

The converse to this is also true, i.e. given a stationary metric such that $\xi = \partial_t$ is hypersurface orthogonal, one can find a coordinate transformation (really just $t \to T$), such that $\partial_t = \partial_T$, $\partial_T g_{\alpha\beta} = 0$, and such that $g_{Tk} = 0$ so that ∂_t is manifestly orthognoal to the surfaces of constant T.

I will give two proofs of this, one using the "integrated" version $\xi_\alpha = f\partial_\alpha S$ of the hypersurface orthogonality condition, and the other using the integrability condition (15.55) for hypersurface orthogonality, in conjunction with the covariant characterisation (16.55) of a stationary metric.

1. The first proof is somewhat pedestrian and not particularly elegant but has the virtue that it is clear from the beginning where one wants to go and what one is doing to get there. We begin with the hypersurface orthogonality condition in the form

$$\xi_\alpha = g_{\alpha t} \stackrel{!}{=} f\partial_\alpha S \quad . \tag{16.64}$$

In particular,

$$\xi^\alpha \xi_\alpha = \xi_t = g_{tt} = f \partial_t S \neq 0 \quad . \tag{16.65}$$

The idea will be to change variables from t to $T = S(x^\alpha)$ (because then ξ is orthogonal to the surfaces of constant T), but before doing this we will need a preliminary result following from the assumption that the metric is stationary. Namely, (16.64) and stationarity evidently imply

$$\partial_t g_{\alpha\beta} = 0 \quad \Rightarrow \quad \partial_t (f \partial_\alpha S) = 0 \quad \text{and} \quad \partial_t (\partial_k S / \partial_t S) = 0 \quad . \tag{16.66}$$

Now, since $\partial_t S \neq 0$, S has to depend on t, but this dependence needs to drop out of the ratio $\partial_k S / \partial_t S$. This implies that, as far as its t-dependence is concerned, S is a linear function of t with constant coefficients,

$$S(t, x^k) = at + b + s(x^k) \quad , \tag{16.67}$$

and this in turn implies that f is t-independent,

$$\partial_t (f \partial_t S) = 0 \quad \Rightarrow \quad \partial_t f = 0 \quad . \tag{16.68}$$

Without loss of generality we can assume that $b = 0$ (either because ξ only depends on $\partial_\alpha S$, or by absorbing it into $s(x^k)$), and that $a = 1$ (by absorbing the constant a into f, say). Thus we can assume that $S(x^\alpha)$ has the form

$$S(t, x^k) = t + s(x^k) \quad \Rightarrow \quad \partial_t S(t, x^k) = 1 \quad , \quad \partial_k S(t, x^k) = \partial_k s(x^k) \quad . \tag{16.69}$$

Using $g_{\alpha t} = f \partial_\alpha S$ and $\partial_t S = 1$, one has

$$f dS = f dt + f \partial_k s dx^k = g_{tt} dt + g_{tk} dx^k \quad , \tag{16.70}$$

and

$$g_{tk} / g_{tt} = \partial_k S \quad . \tag{16.71}$$

Therefore we can write the metric as

$$\begin{aligned} ds^2 &= g_{tt} dt^2 + 2 g_{tk} dt dx^k + g_{ik} dx^i dx^k \\ &= g_{tt} (dt + (g_{tk}/g_{tt}) dx^k)^2 + (g_{ik} - g_{ti} g_{tk} / g_{tt}) dx^i dx^k \\ &= f (dS)^2 + (g_{ik} - f \partial_i s \partial_k s) dx^i dx^k \quad . \end{aligned} \tag{16.72}$$

This again strongly suggests that the right thing to do is to introduce a new time-coordinate T through

$$T = S(t, x^k) = t + s(x^k) \quad , \tag{16.73}$$

with

$$\partial_t S = 1 \quad \Rightarrow \quad \partial_T = \partial_t \quad . \tag{16.74}$$

Then the metric is

$$ds^2 = f(x^k)dT^2 + (g_{ik}(x^k) - f(x^k)\partial_i s(x^k)\partial_k s(x^k))dx^i dx^k \qquad (16.75)$$

with

$$\partial_T g_{\alpha\beta} = \partial_t g_{\alpha\beta} = 0 \quad , \quad g_{kT} = 0 \ . \qquad (16.76)$$

and $\partial_t = \partial_T$ is manifestly orthogonal to the surfaces of constant $T = S$.

2. For the second proof, we start with the integrability condition (15.55). Using the fact that $\nabla_\alpha \xi_\beta$ is anti-symmetric, we can write it in a way which makes its anti-symmetry in the indices (α, β) manifest,

$$\xi_\alpha \nabla_\beta \xi_\gamma - \xi_\beta \nabla_\alpha \xi_\gamma + \tfrac{1}{2}\xi_\gamma(\nabla_\alpha \xi_\beta - \nabla_\beta \xi_\alpha) = 0 \ . \qquad (16.77)$$

Contracting this expression with ξ^γ, and using the abbreviation $\xi^2 = \xi^\gamma \xi_\gamma$ for the norm of ξ, one deduces

$$\xi_\alpha \nabla_\beta(\xi^2) - \xi_\beta \nabla_\alpha(\xi^2) + \xi^2(\nabla_\alpha \xi_\beta - \nabla_\beta \xi_\alpha) = 0 \ , \qquad (16.78)$$

which can also be written as

$$\nabla_\alpha(\xi_\beta/\xi^2) - \nabla_\beta(\xi_\alpha/\xi^2) = 0 \ . \qquad (16.79)$$

Thus ξ^α/ξ^2 is (locally) a gradient vector, i.e.

$$\xi_\alpha = \xi^2 \partial_\alpha S \qquad (16.80)$$

for some function S. This is the integrated version $\xi_\alpha = f\partial_\alpha S$ of the hypersurface orthogonality condition, with the additional information that $f = \xi^2$, so that in particular it is independent of t. The proof could now follow the lines of the first argument, but *variatio delectat*, and we will proceed in a slightly different way.

First of all, we note that $\xi_t = \xi^2$ (16.57) implies

$$\xi_t = \xi^2 \quad \Rightarrow \quad \partial_t S = 1 \ , \qquad (16.81)$$

a condition which thus arises here seemingly in a somewhat different way than before. We can then deduce that S is of the form

$$S(t, x^k) = t + s(x^k) \ , \qquad (16.82)$$

so that

$$\xi_k = g_{kt} = \xi^2 \partial_k S = g_{tt}\partial_k s \ . \qquad (16.83)$$

We now change variables from (t, x^k) to (T, x^K) with $T = S(t, x^k)$ and $x^K = x^k$, or

$$t(T, x^K) = T - s(x^K) \quad , \quad x^k(T, x^K) = x^K \ . \qquad (16.84)$$

Then we can deduce that $\partial_T = \partial_t$ and that in the new coordinates the off-diagonal component of the metric g_{KT} is

$$\begin{aligned} g_{KT} &= \frac{\partial x^\alpha}{\partial x^K}\frac{\partial x^\beta}{\partial T}g_{\alpha\beta} = \frac{\partial x^\alpha}{\partial x^K}g_{\alpha t} \\ &= g_{kt} - g_{tt}\partial_k s = 0 \ \ . \end{aligned} \tag{16.85}$$

Either way we have shown that the general form of a static metric, without assuming the existence of any further symmetries, can be chosen to be of the block-diagonal form

$$\text{Static Metric:} \quad ds^2 = g_{tt}(x^k)dt^2 + g_{ik}(x^k)dx^i dx^k \qquad (x^\alpha = (t, x^k)) \ \ . \tag{16.86}$$

This is known as the standard form of a static metric.

<u>REMARKS:</u>

1. We will see in section 24.2 that a stationary and spherically symmetric metric is automatically static. This follows easily from the fact that for a stationary metric, and in spherical symmetry, in coordinates (t, r, θ, ϕ) suitable for expressing both these facts, the only allowed off-diagonal g_{tk}-term of the metric is $C(r) = g_{tr}(r)$, so that the (t, r)-part of the metric takes the form

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + 2C(r)dt\, dr \ \ . \tag{16.87}$$

   Then $C(r)$ can be eliminated by a coordinate transformation $T(t, r) = t + \psi(r)$, and $\partial_t = \partial_T$ is thus orthogonal to the surfaces of constant T.

2. Evidently the *ultrastatic* metrics (2.34),

$$ds^2 = -dt^2 + \tilde{g}_{ij}(x)dx^i dx^j \ \ , \tag{16.88}$$

   whose geodesics were discussed in section 2.9, are the special case of static metrics for which the norm of $\xi = \partial_t$ is constant.

3. We see from a comparison of (16.53) and (16.86) that an equivalent way of characterising static metrics is that they are invariant under time translations (stationary) and invariant under *time reflections* $t \to -t$ or $t \to c - t$

4. Even though the discrete time-reflection invariance is by no means implied by the continuous time-translation invariance, it may at first be difficult to imagine a situation that does not change in time (the stationarity condition) but that is nevertheless not invariant under time-reflections. Intuitively, a stationary non-static situation can arise when one has e.g. something like a stationary (unchanging) stream or flow in one direction. Time-reversal means inverting the direction of the flow, and even though this is again a stationary situation the difference between the two is of physical significance, and can be detected e.g. by throwing a (test-)twig into the stream and observing its motion.

5. The prime example of a stationary but not static metric is precisely of this kind. This is the *Kerr metric* describing the gravitational field outside a rotating star (or black hole), briefly mentioned in section 30.1. This solution is stationary and axially symmetric (around the axis of rotation), but it turns out that the space-time is distorted in the direction of the rotation. In suitable coordinates (t, r, θ, ϕ) this manifests itself in the fact that the metric coefficients are independent of t (stationarity) and ϕ (axial symmetry), but do depend not only on r but also on θ. Thus, in agrement with the remark above, the metric is not spherically symmetric.

   Under $t \to -t$, the sense of rotation is changed and the corresponding metric cannot be invariant under this operation because the gravitational field is now distorted in the opposite angular direction. In fact, in these coordinates the metric turns out to have a non-vanishing $g_{t\phi}(r, \theta)$, and $g_{t\phi} \to -g_{t\phi}$ under $t \to -t$.

17 Hypersurfaces III: Intrinsic Geometry of Null Hypersurfaces

17.1 Null Hypersurfaces

We now look at null hypersurfaces, which we will denote by $\Sigma = \mathcal{N}$. Null hypersurfaces have some special and peculiar properties, as a consequence of the fact that for a null hypersurface the normal vectors are orthogonal to themselves, $\xi^\alpha \xi_\alpha = 0$, and are therefore not only normal to the hypersurface but simultaneously also tangent to the hypersurface,

$$\xi^\alpha \text{ normal to } \mathcal{N} \text{ and } \xi^\alpha \xi_\alpha = 0 \quad \Rightarrow \quad \xi^\alpha \text{ tangent to } \mathcal{N} \ . \tag{17.1}$$

The scalar product between a null and a timelike vector is always non-zero, because it picks out the time-component of the the null vector. Thus we also learn that a tangent vector to a null hypersurface cannot be timelike and is thus either null or spacelike.

One consequence of this is also a converse to the above statement, namely that a null tangent vector to a null hypersurface $\mathcal{N}$ is also normal to $\mathcal{N}$,

$$\eta^\alpha \text{ tangent to } \mathcal{N} \text{ and } \eta^\alpha \eta_\alpha = 0 \quad \Rightarrow \quad \eta^\alpha \text{ normal to } \mathcal{N} \ , \tag{17.2}$$

or

$$\eta^\alpha \text{ tangent to } \mathcal{N} \text{ and } \eta^\alpha \eta_\alpha = 0 \quad \Rightarrow \quad \eta^\alpha \sim \xi^\alpha \ . \tag{17.3}$$

Intuitively, this is clear, because if η^α were null, tangent and not proportional to the normal ξ^α, there would be 2 linearly independent null vectors tangent to $\mathcal{N}$, but for that to be possible $\mathcal{N}$ would need to be timelike.

Formally, one can prove this by expanding η^α as

$$\eta^\alpha \text{ tangent to } \mathcal{N} \quad \Rightarrow \quad \eta^\alpha = f\xi^\alpha + s^\alpha \tag{17.4}$$

for some function f and spacelike vector s^α, and noting that the 2 requirements

$$\eta^\alpha \text{ tangent} \quad \Rightarrow \quad \eta^\alpha \xi_\alpha = s^\alpha \xi_\alpha \stackrel{!}{=} 0 \tag{17.5}$$

and

$$\eta^\alpha \text{ null} \quad \Rightarrow \quad \eta^\alpha \eta_\alpha = 2f s^\alpha \xi_\alpha + s^\alpha s_\alpha \stackrel{!}{=} 0 \tag{17.6}$$

imply $s^\alpha s_\alpha = 0$ and hence (because s^α is spacelike), $s^\alpha = 0$, so that $\eta^\alpha = f\xi^\alpha$, as claimed.

Since in general for a null hypersurface one has $\xi^\alpha \xi_\alpha = 0$ for any normal vector ξ^α, we cannot normalise it as in the spacelike or timelike case, However, given the defining function S, a convenient and natural choice for a normal vector is

$$\ell_\alpha = -\partial_\alpha S \ , \tag{17.7}$$

where the sign has been chosen in such a way that ℓ^α is future-oriented for a function S that increases towards the future (for an illustration of this see the examples below where $S = t - x$ or $S = t - r$ have this property). All other normal vectors are then of the form

$$\xi^\alpha = f\,\ell^\alpha \tag{17.8}$$

for some function f nowhere vanishing on $\mathcal{N}$.

<u>EXAMPLES:</u>

1. To see an illustration of these facts in the simplest case, consider (1+1)-dimensional Minkowski space in lightcone coordinates $u = t - x, v = (t + x)/2$, say, so that the line element takes the form

   $$ds^2 = -dt^2 + dx^2 = -2du\,dv \ . \tag{17.9}$$

   One could have also made a more symmetric choice for (u, v), of course, this is irrelevant for what follows and the present asymmetric choice just serves to avoid some other factors of 2 further down. The signs of the lightcone coordinates have been chosen in such a way that the null vector fields ∂_u and ∂_v are future-oriented (i.e. u and v grow with increasing t).

   Now consider the family of hypersurfaces (straight lines) defined by

   $$S(t, x) = t - x = u = \text{const.} \tag{17.10}$$

   On the one hand, the complementary null coordinate v provides a good coordinate on each null line $u =$ const. On the other hand, a normal vector to this hypersurface is

   $$\ell^\alpha = -g^{\alpha\beta}\partial_\beta S = -g^{\alpha\beta}\partial_\beta u = -g^{\alpha u} \tag{17.11}$$

   i.e.

   $$\ell = \ell^\alpha\partial_\alpha = -g^{\alpha u}\partial_\alpha = +\partial_v \ . \tag{17.12}$$

   We see that

   - the normal vector $\ell = \partial_v$ is evidently also tangent to the hypersurface, as it points in the direction of varying v, i.e. along the lines of constant u;
   - due to the choice of sign the normal vector $\ell = +\partial_v$ is future oriented.

2. If we add further spatial directions (y, z), then the null hypersurface (hyperplane in this case) $u =$ const. would be parametrised by the null coordinate v and the spatial coordinates (y, z). A slightly more interesting higher-dimensional example of a null surface is provided by introducing spherical coordinates for 4-dimensional Minkowski space,

   $$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2 \ , \tag{17.13}$$

and replacing $S(t,x) = t - x$ by its 4-dimensional radial counterpart

$$S(t,r,\theta,\phi) = t - r \quad . \tag{17.14}$$

We can also introduce the coordinates $u = t - r, v = (t + r)/2$, in terms of which the metric takes the form

$$ds^2 = -2dudv + (v - u/2)^2 d\Omega^2 \quad . \tag{17.15}$$

Then $S = 0$ defines the future lightcone of the origin (see also example 3 in section 15.1), and it is clearly a null surface. Indeed, the normal vector $\ell_\alpha = -\partial_\alpha S$ has components

$$(\ell_\alpha) = (\ell_t = -1, \ell_r = +1, \ell_\theta = 0, \ell_\phi = 0) \quad \Rightarrow \quad \ell^\alpha \ell_\alpha = 0 \quad . \tag{17.16}$$

Moreover, as in the previous example $\ell = \ell^\alpha \partial_\alpha = +\partial_v$, so that ℓ is again future pointing. A point on the lightcone is then specified by the spatial coordinates (θ, ϕ) and the parameter v along the null lines with $u = 0$ and constant (θ, ϕ).

17.2 Null Hypersurfaces and their Null Geodesic Generators

Since ℓ^α is tangent to the hypersurface or family of hypersurfaces $\mathcal{N}$, the integral curves $x^\alpha(\lambda)$ of ℓ^α, characterised by

$$\frac{dx^\alpha(\lambda)}{d\lambda} = \ell^\alpha(x(\lambda)) \quad , \quad x^\alpha(0) \in \mathcal{N} \quad , \tag{17.17}$$

lie entirely in the null hypersurface $\mathcal{N}$. These curves turn out to be null geodesics, although not necessarily affinely parametrised,

$$\ell^\beta \nabla_\beta \ell^\alpha \sim \ell^\alpha \quad , \tag{17.18}$$

and the same thing is true for any other choice of normal vector ξ^α,

$$\xi^\beta \nabla_\beta \xi^\alpha \sim \xi^\alpha \quad . \tag{17.19}$$

I will give 3 proofs of this fact, in increasing order of generality,

1. first for the canonical choice of normal vector $\ell_\alpha = -\partial_\alpha S$,
2. then for any normal vector $\xi^\alpha = f\ell^\alpha$,
3. and then for any null vector field ξ^α satisfying the hypersurface orthogonality condition (15.55).

Proof:

1. Let $\ell_\alpha = -\partial_\alpha S$. Then

$$\ell^\beta \nabla_\beta \ell_\alpha = \ell^\beta \nabla_\alpha \ell_\beta = \tfrac{1}{2}\nabla_\alpha(\ell^\beta \ell_\beta) \ . \tag{17.20}$$

Since $\ell^\beta \ell_\beta = 0$ everywhere on $\mathcal{N}$, $\ell^\beta \ell_\beta$ is constant along directions tangent to $\mathcal{N}$. Now there are two possibilities:

- $\nabla_\alpha(\ell^\beta \ell_\beta) = 0$ on $\mathcal{N}$
  In this case clearly ℓ^α is not only geodesic but even affinely parametrised,

  $$\ell^\beta \nabla_\beta \ell^\alpha = 0 \ . \tag{17.21}$$

- $\nabla_\alpha(\ell^\beta \ell_\beta) \neq 0$ on $\mathcal{N}$
  In that case $\nabla_\alpha(\ell^\beta \ell_\beta)$ is normal to $\mathcal{N}$. Since it is normal to $\mathcal{N}$, it is necessarily proportional to the normal vector ℓ_α,

  $$\nabla_\alpha(\ell^\beta \ell_\beta) \sim \ell_\alpha \ , \tag{17.22}$$

  and thus one deduces (17.18).

Either way, we have shown that

$$\ell^\beta \nabla_\beta \ell^\alpha(x) = \kappa_\ell(x)\ell^\alpha \tag{17.23}$$

for some scalar function $\kappa_\ell(x)$ measuring the *inaffinity* (lack of affinity) of the family of geodesics (geodesic congruence) defined by the normal vector field ℓ (as in (2.136) for a single geodesic curve).

REMARKS:

(a) The situation in the 1st case arises if $S(x) = c$ defines a family of null hypersurfaces, i.e. not just the surface $\mathcal{N}$ defined by $S(x) = 0$ is null but also the surfaces $S(x) = c$ for c in some interval around 0, because then $\ell_\alpha \ell^\alpha = 0$ not just on the surface $S(x) = 0$ but in a neighbourhood of $\mathcal{N}$.

(b) In the 2nd case one could have chosen

$$S = \ell^\beta \ell_\beta \tag{17.24}$$

as a defining function of the surface $\mathcal{N}$. This is natural if one is initially given a vector field and looks at the locus where this vector field becomes null. A prominent example of this we will come back to later is the *Killing Horizon* of a black hole (section 32.5).

2. Now let $\xi^\alpha = f\ell^\alpha$ be any normal vector to $\mathcal{N}$. To establish (17.19), it is sufficient to note that

$$\begin{aligned}\xi^\beta\nabla_\beta\xi^\alpha &= f\ell^\beta\nabla_\beta(f\ell^\alpha)\\ &= f\ell^\beta(\nabla_\beta f)\ell^\alpha + f^2\ell^\beta\nabla_\beta\ell^\alpha\\ &= (f\ell^\beta\partial_\beta f + f^2\kappa_\ell)\ell^\alpha\\ &= (\ell^\beta\partial_\beta f + f\kappa_\ell)\xi^\alpha \equiv \kappa_\xi\,\xi^\alpha\quad .\end{aligned} \tag{17.25}$$

Thus we have shown that ξ^α is geodesic iff ℓ^α is geodesic and that the inaffinities of ℓ and $\xi = f\ell$ are related by

$$\xi^\alpha = f\;\ell^\alpha \quad\Rightarrow\quad \kappa_\xi = \ell^\alpha\partial_\alpha f + f\kappa_\ell\quad . \tag{17.26}$$

In particular, if $f(x)$ is a solution of the differential equation $\ell^\alpha\partial_\alpha\log f + \kappa_\ell = 0$ (along any orbit of ℓ this reduces to the differential equation in (2.137)), the normal vector field $\xi^\alpha = f\ell^\alpha$ is affinely parametrised,

$$\ell^\alpha(x)\partial_\alpha\log f(x) = -\kappa_\ell(x) \quad\Rightarrow\quad \xi^\beta\nabla_\beta\xi^\alpha = 0\quad . \tag{17.27}$$

3. Let us assume that we are given a hypersurface orthogonal vector field ξ^α. Explicitly, we can write the condition (15.55) as

$$\xi_\alpha(\nabla_\beta\xi_\gamma - \nabla_\gamma\xi_\beta) + \xi_\beta(\nabla_\gamma\xi_\alpha - \nabla_\alpha\xi_\gamma) + \xi_\gamma(\nabla_\alpha\xi_\beta - \nabla_\beta\xi_\alpha) = 0\quad . \tag{17.28}$$

Contracting this with ξ^α and assuming that $\xi^\alpha\xi_\alpha = 0$ on some hypersurface $\mathcal{N}$, one finds that on $\mathcal{N}$ one has the condition

$$\tfrac{1}{2}(\xi_\beta\nabla_\gamma(\xi^\alpha\xi_\alpha) - \xi_\gamma\nabla_\beta(\xi^\alpha\xi_\alpha)) + (\xi_\beta\xi^\alpha\nabla_\alpha\xi_\gamma - \xi_\gamma\xi^\alpha\nabla_\alpha\xi_\beta) = 0\quad , \tag{17.29}$$

containing two kinds of terms. The 1st term is of the familiar type already dealt with above. Either $\xi^\alpha\xi_\alpha = 0$ also off the surface, or $\nabla_\gamma(\xi^\alpha\xi_\alpha) \sim \xi_\gamma$. Either way, the 1st term is zero. We are thus left with the condition

$$\xi_\beta\xi^\alpha\nabla_\alpha\xi_\gamma = \xi_\gamma\xi^\alpha\nabla_\alpha\xi_\beta\quad . \tag{17.30}$$

In general, if one has two vector fields V and W satisfying $V_\alpha W_\beta = V_\beta W_\alpha$, and neither of them is identically zero, it follows that V and W are proportional,

$$V_\alpha W_\beta = V_\beta W_\alpha \quad\Rightarrow\quad W_\alpha = f\;V_\alpha \tag{17.31}$$

for some scalar f. Here is a low-brow proof of this statement:

- The condition is empty for $\alpha = \beta$; thus fix $\alpha \neq \beta$ (and let us choose $\alpha = 1$, say):
- If $V_1 = 0$, it follows that $V_\beta W_1 = 0$ for all β, and therefore $W_1 = 0$ as well.
- If $V_1 \neq 0$, one can write $W_\beta = (W_1/V_1)V_\beta$ for all β. Thus $W_1 \neq 0$, and therefore $W_\beta = fV_\beta$ for some non-zero scalar f, as claimed

In the case at hand, this implies

$$(17.30) \quad \Rightarrow \quad \xi^\alpha \nabla_\alpha \xi_\gamma \sim \xi_\gamma \ , \tag{17.32}$$

which is precisely the statement we set out to prove, namely that on $\mathcal{N}$ the null normal vector field is (possibly non-affinely) geodesic.

Since any point on $\mathcal{N}$ lies on one of these null geodesics, one says that the null surface is generated by these null geodesics. The null geodesics, in turn, are known as the *null generators of* $\mathcal{N}$.

<u>REMARKS:</u>

1. Returning to the examples discussed at the beginning of this section, in both cases the normal and geodesic null vector field $\ell = \partial_v$ is actually affinely parametrised,

$$\ell = \partial_v \quad \Rightarrow \quad \ell^\beta \nabla_\beta \ell^\alpha = \Gamma^\alpha_{vv} = 0 \ , \tag{17.33}$$

   i.e. the null coordinate v is an affine parameter along these (right-moving respectively radial outgoing null geodesics), and is thus e.g. a natural coordinate to use on $\mathcal{N}$. The reason one finds affinely parametrised geodesics in this case is that $S(u) = u = c$ defines a family of null hypersurfaces.

2. In this general context of null surfaces, the inaffinity κ_ξ associated with a particular choice $\xi^\alpha = f\ell^\alpha$ of normal vector field has no particular significance since, as we have seen, it can be changed at will by changing f.

3. However, these geodesics and their associated inaffinity acquire a particular importance when the normal vector field in question cannot be rescaled in an arbitrary way by a scalar f.

   This arises for example when one has a Killing vector ξ that becomes normal to some null hypersurface, and is thus in particular null there (this hypersurface is then called a *Killing horizon*). Since $f(x)\xi$ will then not be a Killing vector unless $f(x) = a$ is constant, the ambiguity in the inaffinity is greatly reduced, to multiplication of κ_ξ by a constant,

$$\xi \to a\xi \quad \Rightarrow \quad \kappa_\xi \to a\kappa_\xi \ . \tag{17.34}$$

   This remaining ambiguity can e.g. be fixed completely by demanding that $\xi^\alpha \xi_\alpha \to -1$ asymptotically if ξ is asymptotically timelike. In this case κ_ξ, known as the *surface gravity* associated with the Killing horizon, carries intrinsic information about the space-time itself and plays an important role in the study of black holes - an illustration of this in the simplest possible case of the Schwarzschild metric is given in section 27.10, and a more general discussion of Killing horizons and surface gravity can be found in sections 32.5 and 32.6.

17.3 Adapted Coordinates and Induced Metric for Null Hypersurfaces

Since in general the null geodesics which are the generators of a null surface are naturally associated with the null surface, it is also convenient to adapt the coordinates y^a on $\mathcal{N}$ to ℓ^α by choosing the coordinates to be

$$y^a = (v = \lambda, y^k) \tag{17.35}$$

where λ is the (not necessarily affine) parameter along the null geodesics and y^k are spatial coordinates labelling the individual null geodesics. In particular, therefore, the y^k are constant along the null geodesics and can be constructed e.g. from the constants of motion or the constants of integration of the null geodesic equation.

In these coordinates, the tangent vectors E_a (15.16) to the null surface are

$$E_v^\alpha = \frac{\partial x^\alpha}{\partial \lambda} = \ell^\alpha \quad , \quad E_k^\alpha = \frac{\partial x^\alpha}{\partial y^k} \tag{17.36}$$

and therefore the induced metric (15.23) has the components

$$h_{vv} = g_{\alpha\beta}\ell^\alpha\ell^\beta = 0 \quad , \quad h_{vk} = g_{\alpha\beta}\ell^\alpha E_k^\beta = 0 \quad , \quad h_{km} \equiv s_{km} = g_{\alpha\beta}E_k^\alpha E_m^\beta \quad , \tag{17.37}$$

where $h_{vk} = 0$ follows because by construction the E_k^α are tangent to the surface while by definition ℓ^α is normal to the surface and therefore in particular normal to the tangent vectors E_k^α.

Thus the metric is clearly degenerate (a characteristic feature of null surfaces) and the line element takes the form

$$ds^2|_{\mathcal{N}} = s_{km}dy^k dy^m = g_{\alpha\beta}E_k^\alpha E_m^\beta dy^k dy^m \quad . \tag{17.38}$$

Note that this form of the metric is independent of whether one chooses λ to be the original (perhaps non-affine) parameter or the affine parameter, as this just amounts to changing $\ell^\alpha \to \xi^\alpha = f\ell^\alpha$ for a suitable choice of f, so that one still has $h_{vv} = h_{vk} = 0$.

Returning to the example of the future lightcone in Minkowski space (example 3 in section 15.1 and example 2 in section 17.1), we find that in the coordinates (v, θ, ϕ) the metric induced on the lightcone by the ambient Minkowski metric is the degenerate metric with line element

$$ds^2 = v^2(d\theta^2 + \sin^2\theta d\phi^2) \ , \tag{17.39}$$

as could also have been deduced directly by restricting the metric (17.15) to $u = t - r = 0$,

$$\left(-2dudv + (v - u/2)^2 d\Omega^2\right)|_{u=0} = v^2 d\Omega^2 \quad . \tag{17.40}$$

17.4 Projectors for Null Hypersurfaces

As in the case of non-null hypersurfaces, one can also study the induced metric from the point of view of embedded hypersurfaces and projection operators. However, the construction is somewhat different in this case because the tangent directions E_a^α and the normal direction ℓ^α are not independent. It is clear that, in order to e.g. have a completeness relation akin to (16.11), we should adjoin to the spatial tangent directions E_k^α and the null tangent direction ℓ^α to the surface another linearly independent vector which can conveniently be chosen to be a null vector n^α on $\mathcal{N}$, but of course not tangent to $\mathcal{N}$, such that

$$n_\alpha n^\alpha = 0 \quad , \quad n_\alpha E_k^\alpha = 0 \quad , \quad n_\alpha \ell^\alpha \neq 0 \ . \tag{17.41}$$

We can always rescale n^α in such a way that $n_\alpha \ell^\alpha = -1$, and this is a convenient choice we will adapt in the following (the minus sign having been chosen such that n^α is future directed iff ℓ^α is future directed). Thus, given a choice of spatial basis vectors E_k^α, the set-up is the set of vectors $\{\ell^\alpha, n^\alpha, E_k^\alpha\}$ satisfying the relations

$$n_\alpha n^\alpha = \ell_\alpha \ell^\alpha = 0 \quad , \quad n_\alpha E_k^\alpha = \ell_\alpha E_k^\alpha = 0 \quad , \quad n_\alpha \ell^\alpha = -1 \ . \tag{17.42}$$

Given ℓ^α and a choice of E_k^α (up to purely spatial coordinate transformations of the y^k), the complementary null vector n^α is uniquely determined by these conditions. The freedom $\ell^\alpha \to \xi^\alpha$ in the choice of normal vector to multiply it by a non-zero function amounts to a boost in the (ℓ, n)-plane,

$$\ell^\beta \to \mathrm{e}^{+\alpha(x)} \ell^\beta \quad \Rightarrow \quad n^\beta \to \mathrm{e}^{-\alpha(x)} n^\beta \ . \tag{17.43}$$

With these vectors we can define the tensor $s_{\alpha\beta}$ on $\mathcal{N}$ by

$$g_{\alpha\beta} = s_{\alpha\beta} - (\ell_\alpha n_\beta + \ell_\beta n_\alpha) \ , \tag{17.44}$$

or

$$s_{\alpha\beta} = g_{\alpha\beta} + (\ell_\alpha n_\beta + \ell_\beta n_\alpha) \ . \tag{17.45}$$

Note that $s_{\alpha\beta}$ is invariant under the boost (17.43). This tensor has the properties

$$s_{\alpha\beta}\ell^\beta = s_{\alpha\beta} n^\beta = 0 \tag{17.46}$$

and

$$V^\alpha \ell_\alpha = V^\alpha n_\alpha = 0 \quad \Rightarrow \quad g_{\alpha\beta} V^\beta = s_{\alpha\beta} V^\beta \ . \tag{17.47}$$

It thus defines the induced metric in the directions orthogonal to ℓ^α and n^α, and is thus the induced (degenerate, spatial) metric on the surface $\mathcal{N}$, the properties

$$s^{\alpha\beta} = E_k^\alpha E_m^\beta s^{km} \quad , \quad s_{km} = E_k^\alpha E_m^\beta s_{\alpha\beta} \tag{17.48}$$

being the analogues of (16.10) and (16.6) respectively. One also has

$$g^{\alpha\beta} = s^{km} E_k^\alpha E_m^\beta - (\ell^\alpha n^\beta + \ell^\beta n^\alpha) \ , \tag{17.49}$$

which is the null analogue of the completeness relation (16.11). The properties (17.46) and (17.47) also imply

$$g^{\alpha\beta}s_{\alpha\beta} = s^{\alpha\beta}s_{\alpha\beta} = s^{\alpha}_{\ \alpha} = n-1 \quad . \tag{17.50}$$

We can now introduce the projectors

$$s^{\alpha}_{\ \beta} = \delta^{\alpha}_{\ \beta} + (\ell^{\alpha}n_{\beta} + n^{\alpha}\ell_{\beta}) \quad , \quad s^{\alpha}_{\ \gamma}s^{\gamma}_{\ \beta} = s^{\alpha}_{\ \beta} \tag{17.51}$$

onto the transverse space. They satisfy, in particular,

$$s^{\alpha}_{\ \beta}\ell^{\beta} = s^{\alpha}_{\ \beta}n^{\beta} = 0 \quad , \tag{17.52}$$

and are therefore such that if e.g. V^{α} is a space-time vector, $v^{\alpha} \equiv s^{\alpha}_{\ \beta}V^{\beta}$ is a spatial vector tangent to $\mathcal{N}$, i.e. orthogonal to ℓ^{α} and n^{α},

$$v^{\alpha} = s^{\alpha}_{\ \beta}V^{\beta} \quad \Rightarrow \quad v^{\alpha}\ell_{\alpha} = v^{\alpha}n_{\alpha} = 0 \quad . \tag{17.53}$$

Note that $v^{\alpha} = E^{\alpha}_{k}v^{k}$ is a purely spatial vector so that, in particular,

$$v \neq 0 \quad \Rightarrow \quad v^{\alpha}v_{\alpha} > 0 \tag{17.54}$$

and likewise for other tensors.

A variant of the set-up in this section (in particular the auxiliary complementary null vector n^{α} and the corresponding projectors) appeared in the discussion of the Raychaudhuri equation for null geodesic congruences in section 12.4.

As an aside (a useful aside, though), note that this entire set-up can be phrased in a somewhat more satisfactory manner in terms of an orthonormal basis or frame E_a (section 4.8) rather than in terms of the basis E_k associated to the choice of coordinates y^k (similar remarks apply to the timelike case). Namely, by introducing suitable linear combinations

$$E^{\alpha}_{a} = E^{k}_{a}E^{\alpha}_{k} \tag{17.55}$$

of the E^{α}_{k} which diagonalise the spatial metric s_{km},

$$E^{k}_{a}E^{m}_{b}s_{km} = \delta_{ab} \quad \Rightarrow \quad g_{\alpha\beta}E^{\alpha}_{a}E^{\beta}_{b} = \delta_{ab} \quad , \tag{17.56}$$

one obtains the (pseudo-)orthonormal basis (12.83)

$$\{E_A\} = \{E_+ = \ell, E_- = n, E_a\} : \quad g_{\alpha\beta}E^{\alpha}_{A}E^{\beta}_{B} = \eta_{AB} \tag{17.57}$$

with

$$\eta_{++} = \eta_{--} = 0 \quad , \quad \eta_{+-} = -1 \quad , \quad \eta_{a+} = \eta_{a-} = 0 \quad , \quad \eta_{ab} = \delta_{ab} \quad . \tag{17.58}$$

Then the boost (17.43) really is a Lorentz boost (in the tangent space), and the ambiguity in the identification of a choice of null vector n and complementary spatial directions can be identified as the possibility to perform a *null Lorentz rotation* (12.81) around ℓ,

$$\ell \to \ell \quad , \quad n \to n + \beta^a E_a + \tfrac{1}{2}\beta^2 \ell \quad , \quad E_a \to E_a + \beta_a \ell \;\; , \tag{17.59}$$

where $\beta^2 = \delta_{ab}\beta^a\beta^b$. Note that this transformation leaves invariant ℓ^α and the orthonormal frame counterpart

$$n_\alpha n^\alpha = \ell_\alpha \ell^\alpha = 0 \quad , \quad n_\alpha E^\alpha_a = \ell_\alpha E^\alpha_a = 0 \quad , \quad n_\alpha \ell^\alpha = -1 \tag{17.60}$$

of the conditions (17.42).

In Minkowski space, with $\ell = \partial_v$, $n = \partial_u$ and $E_a = \partial_{z^a}$, say, so that the metric has the standard lightcone form

$$ds^2 = -2dudv + d\vec{z}^2 = -2dudv + \delta_{ab}dz^a dz^b \;\; , \tag{17.61}$$

this null Lorentz rotation is generated by the Lorentz transformation

$$(u', v', z'^a) = (u, v + \beta_a z^a + \tfrac{1}{2}\beta^2 u, z^a + \beta^a u) \;\; , \tag{17.62}$$

$$-2du'dv' + \delta_{ab}dz'^a dz'^b = -2dudv + \delta_{ab}dz^a dz^b \;\; . \tag{17.63}$$

18 Hypersurfaces IV: Extrinsic Geometry of non-Null Hypersurfaces

In this section we will briefly touch upon some aspects of *extrinsic geometry*, more specifically of the extrinsic geometry of non-null hypersurfaces. One can also develop the extrinsic geometry of null hypersurfaces and of surfaces of higher codimension, but we will not do this here.

18.1 Introduction: Intrinsic vs Extrinsic Geometry

As mentioned before, the (local) *intrinsic geometry* of a space, i.e. the properties that can be deduced by measuring lengths, areas, volumes, angles, performing parallel transport etc in that space, is completely described by the metric and objects that can be derived from it, like the Riemann curvature tensor. In particular, the intrinsic geometry of a hypersurface Σ, is completely described by its metric, e.g. by the metric induced on it by a metric on the ambient embedding space M.

However, for hypersurfaces there is another aspect of the geometry of Σ beyond its purely intrinsic geometry, namely how it is embedded into the ambient space M, i.e. how it bends inside M. As one needs to be able to move off Σ to even detect that there is such an embedding, this aspect of the geometry is something that cannot be captured by intrinsic measurements on Σ alone, and is therefore known as the *extrinsic geometry* of Σ.

Before developing this, let us look at some simple examples of embedded hypersurfaces:

1. Cylinder $C \subset \mathbb{R}^3$

   For example, a cylinder
$$C = \mathbb{R} \times S^1 \tag{18.1}$$
   with circumference $2\pi L$ along the circle can be obtained by "rolling up" $\mathbb{R}^2$, i.e. by performing the periodic identification
$$(x^1, x^2) \sim (x^1, x^2 + 2\pi L) \ . \tag{18.2}$$
   In this way it clearly inherits the flat metric from $\mathbb{R}^2$. This flat metric is the induced metric on the cylinder when one embeds it in the standard way into $\mathbb{R}^3$. Indeed, introducing cylindrical coordinates (r, ϕ, z) in $\mathbb{R}^3$, the metric on $\mathbb{R}^3$ takes the form
$$ds^2 = dr^2 + r^2 d\phi^2 + dz^2 \ . \tag{18.3}$$
   Identifying the points
$$(r, \phi, z) \sim (r, \phi + 2\pi, z) \tag{18.4}$$

and restricting to constant $r = L$, one obtains a cylinder with circumference $2\pi L$ and induced metric

$$ds^2|_{r=L} = L^2 d\phi^2 + dz^2 \quad . \tag{18.5}$$

Since the components of this metric are constant, the Christoffel symbols and the curvature tensor are zero. Thus, the intrinsic curvature of the cylinder is zero, it is flat (and locally looks just like Euclidean space). In particular, parallel transport is rather obviously path independent. The fact that it looks curved to an outside observer is therefore not something that can be detected by somebody performing local measurements on the cylinder.

2. Torus $T^2 \subset \mathbb{R}^3$ and $T^2 \subset \mathbb{R}^4$

Let us now consider the 2-torus T^2. If one visualises it in the standard way as (the surface of a doughnut) embedded in $\mathbb{R}^3$, then it inherits a non-flat metric from the ambient flat metric on $\mathbb{R}^3$. To see this explicitly, place the torus around the origin of the above cylindrical coordinates, i.e. such that its "center" is at $r = z = 0$ and such that it is invariant under rotations in ϕ around the z-axis. Fixing ϕ, the cross-section of the torus is a circle of radius L_2, say, centered at a distance $r = L_1 > L_2$ from the orgin at the point $(r = L_1, z = 0)$. Thus the points on this circle, and therefore, by including ϕ, the points on the T^2 are described by the equation

$$S(r, z, \phi) = z^2 + (r - L_1)^2 - L_2^2 = 0 \quad . \tag{18.6}$$

(Check: $r = L_1 \Rightarrow z = \pm L_2, r = L_1 \pm L_2 \Rightarrow z = 0$). Then, eliminating z, say, the induced metric is

$$(dr^2 + r^2 d\phi^2 + dz^2)|_{S=0} = \frac{L_2^2}{L_2^2 - (r - L_1)^2} dr^2 + r^2 d\phi^2 \quad , \tag{18.7}$$

with $L_1 - L_2 \leq r \leq L_1 + L_2$ (and the standard range for ϕ). This 2-dimensional metric is not flat (in fact you can check that e.g. its scalar curvature is $R = g_{rr,r}/r(g_{rr})^2$), and in this case both the intrinsic and the extrinsic geometry of the torus are non-trivial.

If, on the other hand, one thinks of the 2-torus T^2 simply as a doubly periodically identified $\mathbb{R}^2$, with periods $2\pi L_1$ and $2\pi L_2$, say, then it certainly inherits the flat metric from $\mathbb{R}^2$,

$$dx^2 + dy^2 \quad \rightarrow \quad (L_1)^2 (d\phi_1)^2 + (L_2)^2 (d\phi_2)^2 \qquad (\phi_k \sim \phi_k + 2\pi) \quad . \tag{18.8}$$

It is thus intrinsically flat, but at the moment we have not embedded this flat torus into any higher-dimensional space. It is not possible to embed T^2 into $\mathbb{R}^3$ in such a way that the induced metric is this flat metric, but it is easy to see that it is possible to achieve this via an embedding into $\mathbb{R}^4$.

Indeed let us introduce polar coordinates (r_1, ϕ_1) in the (12)-plane, and (r_2, ϕ_2) in the (34)-plane, so that the Euclidean metric on $\mathbb{R}^4$ has the form

$$\begin{aligned} ds^2 &= (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2 \\ &= (dr_1)^2 + (r_1)^2(d\phi_1)^2 + (dr_2)^2 + (r_2)^2(d\phi_2)^2 \ . \end{aligned} \tag{18.9}$$

We now identify the points

$$(r_1, \phi_1, r_2, \phi_2) \sim (r_1, \phi_1 + 2\pi, r_2, \phi_2) \sim (r_1, \phi_1, r_2, \phi_2 + 2\pi) \ . \tag{18.10}$$

Then all the lines of constant (r_1, r_2, ϕ_2) and of constant (r_1, r_2, ϕ_1) are circles in orthogonal (12)- and (34)-planes in $\mathbb{R}^4$. Thus the surfaces of constant r_1 and r_2 are tori, and choosing $r_1 = L_1$ and $r_2 = L_2$, one finds that the metric induced on this torus by the ambient flat metric on $\mathbb{R}^2$ is precisely the above flat metric (18.8),

$$ds^2|_{r_k=L_k} = (L_1)^2(d\phi_1)^2 + (L_2)^2(d\phi_2)^2 \ . \tag{18.11}$$

In a parametric description $x^\alpha(\phi_1, \phi_2)$, with respect to the Cartesian coordinates x^α on $\mathbb{R}^4$, this *Clifford embedding* of T^2 into $\mathbb{R}^4$ is given by

$$(x^\alpha(\phi_1, \phi_2)) = (L_1 \cos\phi_1, L_1 \sin\phi_1, L_2 \cos\phi_2, L_2 \sin\phi_2) \ . \tag{18.12}$$

In particular, while these flat tori cannot be realised as codimension-1 hypersurface of $\mathbb{R}^3$, this parametrisation shows that they *can* be realised as codimension-1 hypersurfaces of the 3-sphere

$$(x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 = (L_1)^2 + (L_2)^2 \ , \tag{18.13}$$

with the standard non-flat induced metric on S^3 in turn inducing the flat metric on the embedded T^2.

3. Circle $S^1 \subset \mathbb{R}^2$

   In order to understand how to quantify that both for the flat cylinder and the flat torus the extrinsic geometry is non-trivial it is sufficient to look at the simplest possible lower-dimensional counterpart of this example, namely a 1-dimensional closed space (a loop S^1) embedded into $\mathbb{R}^2$ as a circle of constant radius L.

   This one-dimensional space is evidently intrinsically flat (because the Riemann tensor vanishes identically in 1 dimension), but equally evidently the circle seems to bend / curve around in 2 dimensions (in order to be able to form a circle in the first place).

   In order to quantify this somewhat, one possible strategy is to determine how the (unit) normal vector $N = \partial_r$ to the circle changes as one moves along the circle. The change in ∂_r in the ambient space is given (in polar coordinates) by

$$\nabla_\phi \partial_r = \Gamma^\alpha_{\ \phi\beta}(\partial_r)^\beta \partial_\alpha = \Gamma^\phi_{\ \phi r}\partial_\phi \ . \tag{18.14}$$

This vector is already tangent to the circle (had it not been, we could have now projected it back), and the result can be written as

$$\nabla_\phi N_\phi|_{r=L} = -\Gamma^r{}_{\phi\phi}|_{r=L} = L \quad . \tag{18.15}$$

Alternatively, in order to explore how the embedded circle sits inside the ambient geometry, one can study how the induced metric changes in the normal direction,

$$\partial_r g_{\phi\phi}|_{r=L} = 2L \quad . \tag{18.16}$$

These are equivalent characterisations and quantifications of the extrinsic geometry of the circle (and hence also of the cylinder or the flat torus).

18.2 Extrinsic Curvature Tensor

In order to capture this extrinsic aspect of the geometry in general, we are thus led to define the extrinsic curvature of Σ in M either by

$$K^{(1)}_{\alpha\beta} = h_\alpha^\gamma h_\beta^\delta \nabla_\gamma N_\delta \tag{18.17}$$

or (with a judicious and conventional factor of 1/2) by

$$K^{(2)}_{\alpha\beta} = \tfrac{1}{2} h_\alpha^\gamma h_\beta^\delta L_N g_{\gamma\delta} \quad , \tag{18.18}$$

where, as in section 16.1,

$$h^\alpha_\beta = \delta^\alpha_\beta - \epsilon N^\alpha N_\beta \tag{18.19}$$

is the tangential projector, and $L_N g_{\alpha\beta}$ denotes the Lie derivative of the metric $g_{\alpha\beta}$ along the normal direction N.

Cooperatively and conveniently, the two tensors defined in (18.17) and (18.18) turn out to be identical in general. To see this, we first make use of the formula (9.38)

$$L_N g_{\alpha\beta} = \nabla_\alpha N_\beta + \nabla_\beta N_\alpha \tag{18.20}$$

for the Lie derivative of the metric to write $K^{(2)}_{\alpha\beta}$ as

$$K^{(2)}_{\alpha\beta} = \tfrac{1}{2} h_\alpha^\gamma h_\beta^\delta (\nabla_\gamma N_\delta + \nabla_\delta N_\gamma) \quad . \tag{18.21}$$

This already resembles (18.17), apart from the explicit symmetrisation in (18.21). This symmetrisation, however, is not necessary: since N_α is by definition hypersurface orthogonal, its anti-symmetrised derivative satisfies (15.52)

$$\nabla_{[\alpha} N_{\beta]} = V_{[\alpha} N_{\beta]} \tag{18.22}$$

for some (gradient) vector V_α, and therefore the tangential projection of the anti-symmetric part of $\nabla_\alpha N_\beta$ is equal to zero, and we can simplify (18.21) to

$$K^{(2)}_{\alpha\beta} = h_\alpha^\gamma h_\beta^\delta \nabla_\gamma N_\delta = K^{(1)}_{\alpha\beta} \quad . \tag{18.23}$$

We can therefore drop the labels on $K_{\alpha\beta}$ and define the (symmetric) *extrinsic curvature tensor* $K_{\alpha\beta}$ by

$$K_{\alpha\beta} = \tfrac{1}{2}h_\alpha^\gamma h_\beta^\delta L_N g_{\gamma\delta} = h_\alpha^\gamma h_\beta^\delta \nabla_\gamma N_\delta = K_{\beta\alpha} \quad . \tag{18.24}$$

<u>Remarks:</u>

1. Due to the tangential projections, this definition is independent of how the normal vector N^α is extended off the hypersurface Σ. If it is extended in such a way that $N^\alpha N_\alpha = \epsilon$ also off Σ (e.g. if N^α is the normal vector field to a family of hypersurfaces), then $(\nabla_\alpha N_\beta)N^\beta = 0$ and the 2nd projection in the above definition is unnecessary. In that case one finds

$$K_{\alpha\beta} = h_\alpha^\gamma \nabla_\gamma N_\beta = \nabla_\alpha N_\beta - \epsilon N_\alpha a_\beta \tag{18.25}$$

where

$$a_\beta = N^\gamma \nabla_\gamma N_\beta \tag{18.26}$$

is the "acceleration" of N^α, and the 2nd term is simply there to subtract this normal component of the 1st term. In particular if, as suggested in section 15.4, N^α is extended off the hypersurface as an affinely parametrised geodesic vector field, one simply has $K_{\alpha\beta} = \nabla_\alpha N_\beta$.

2. If the surface is given in parametrised form $x^\alpha(y^a)$, then one can equivalently think of the extrinsic curvature tensor as (or define the extrinsic curvature tensor by)

$$K_{ab} = E_a^\alpha E_b^\beta \nabla_\alpha N_\beta = E_a^\alpha E_b^\beta K_{\alpha\beta} \quad . \tag{18.27}$$

If one adopts the first of these as the definition of K_{ab}, then its symmetry follows from

$$E_b^\beta N_\beta = 0 \quad \Rightarrow \quad E_b^\beta \nabla_\alpha N_\beta = -(\nabla_\alpha E_b^\beta) N_\beta \quad . \tag{18.28}$$

Indeed, using this identity, as well as $E_a^\alpha \partial_\alpha = \partial_a$ and the explicit expression for the covariant derivative ∇_α, one has

$$\begin{aligned} E_a^\alpha E_b^\beta \nabla_\alpha N_\beta &= -E_a^\alpha (\nabla_\alpha E_b^\beta) N_\beta \\ &= -(\partial_a E_b^\beta + \Gamma^\beta_{\alpha\gamma} E_a^\alpha E_b^\gamma) N_\beta \end{aligned} \tag{18.29}$$

which is manifestly symmetric in the indices a, b because

$$\partial_a E_b^\beta = \frac{\partial^2 x^\beta}{\partial y^a \partial y^b} \tag{18.30}$$

and because of the symmetry of the Christoffel symbols.

3. The induced metric h_{ab} and the extrinsic curvature tensor K_{ab} (or equivalently $h_{\alpha\beta}$ and $K_{\alpha\beta}$) are also known as the *1st fundamental form* and *2nd fundamental form* of Σ respectively.

4. Writing the hypersurface orthogonal N_α as

$$N_\alpha = f\partial_\alpha S \tag{18.31}$$

(so that it is orthogonal to the surfaces of constant S), one has

$$\nabla_\alpha N_\beta = (\partial_\alpha f/f)N_\beta + f(\partial_\alpha\partial_\beta S - \Gamma^\gamma{}_{\alpha\beta}\partial_\gamma S) \ \ . \tag{18.32}$$

The first term is killed by the tangential projection, and we see that the remaining second term is manifestly symmetric. In adapted coordinates, i.e. choosing S to be one of the coordinates, one evidently has $\partial_\alpha\partial_\beta S = 0$, and therefore

$$K_{ab} = -fE^\alpha_a E^\beta_b \Gamma^\gamma{}_{\alpha\beta}\partial_\gamma S \equiv -f\Gamma^S{}_{ab} \ \ . \tag{18.33}$$

Therefore K_{ab} essentially consists of the normal components of the Christoffel symbol. The minus sign in this equation is due to our choice of sign convention in the definition of $K_{\alpha\beta}$ or K_{ab}, but frequently it is also defined with an additional minus sign, which then results in $K_{ab} \sim +\Gamma^S{}_{ab}$.

5. The extrinsic curvature also depends on a choice of orientation convention for the normal vector (such as "inward pointing" versus "outward pointing" in situations where this makes sense). When one has several boundary components, some of them timelike and some of them spacelike, say, each one with its own extrinsic curvature tensor, sorting out one's signs in extrinsic geometry provides one with a practically unlimited source of entertainment and/or frustration.

6. The trace of the extrinsic curvature tensor is identical to the space-time divergence of the vector field N^α,

$$K := g^{\alpha\beta}K_{\alpha\beta} = h^{\alpha\beta}K_{\alpha\beta} = \nabla_\alpha N^\alpha \ \ . \tag{18.34}$$

In particular if N^α is extended off the hypersurface as a geodesic vector field, $K = \theta$ measures the *expansion* of this geodesic congruence, as defined in section 12.2.

The sign convention adopted here is such that e.g. $K > 0$ for the sphere with its standard metric, with the outward pointing normal vector. This sign agrees with the sign of the Ricci scalar (and this is one of the reasons for adopting this convention). The property $K = \theta > 0$ indicates that the congruence of geodesics piercing the sphere diverges (rather than converges) in the outgoing direction.

7. In the case of the circle $S^1 \subset \mathbb{R}^2$ of radius L discussed above, the only non-zero component of $K_{\alpha\beta}$ is

$$K_{\phi\phi} = L \tag{18.35}$$

and the trace of the extrinsic curvature of a circle of radius L is

$$K = g^{\alpha\beta}K_{\alpha\beta} = \nabla_\alpha N^\alpha|_{r=L} = 1/L \ \ . \tag{18.36}$$

8. More generally, the trace of the extrinsic curvature of the sphere $S^n_R \subset \mathbb{R}^{n+1}$ of radius R, with its standard metric

$$h_{ab}dy^a dy^b = R^2 d\Omega_n^2 \tag{18.37}$$

is

$$K = \tfrac{1}{2}h^{ab}(\partial_r h_{ab})|_{r=R} = \frac{n}{R} \quad . \tag{18.38}$$

9. An elementary property of the extrinsic curvature is that $K_{\alpha\beta} = 0$ if the normal vector happens to also be a Killing vector,

$$\nabla_\alpha N_\beta = -\nabla_\beta N_\alpha \quad \Rightarrow \quad K_{\alpha\beta} = 0 \;\; , \tag{18.39}$$

because then $K_{\alpha\beta}$ would have to be symmetric as well as anti-symmetric. While this (a Killing vector of constant length, hence also geodesic) is an exceedingly rare situation, N_α only needs to be proportional to a Killing vector for this conclusion to hold,

$$N_\alpha = fK_\alpha \quad , \quad \nabla_\alpha K_\beta + \nabla_\beta K_\alpha = 0 \quad \Rightarrow \quad K_{\alpha\beta} = 0 \;\; , \tag{18.40}$$

because the second term in

$$\nabla_\alpha N_\beta = f\nabla_\alpha K_\beta + (\nabla_\alpha f/f)N_\beta \tag{18.41}$$

will not contribute to the tangential projection of $\nabla_\alpha N_\beta$.

Now recall from section 16.4 that a *static space-time* is a space-time with a hypersurface orthogonal timelike Killing vector. Therefore concrete examples of hypersurfaces with vanishing extrinsic curvature tensor are provided by the spacelike hypersurfaces in static space-times orthogonal to the orbits of the timelike Killing vector.

10. Another useful property of $K_{\alpha\beta}$ is that if u^α is the tangent vector field to a congruence of geodesics that are tangent to Σ, that then the diagonal component of $K_{\alpha\beta}$ in the direction of u^α is zero,

$$u^\alpha \nabla_\alpha u^\beta = u^\alpha N_\alpha = 0 \quad \Rightarrow \quad K_{\alpha\beta}u^\alpha u^\beta = 0 \;\; . \tag{18.42}$$

Indeed, if $u^\alpha N_\alpha = 0$ one has $u^\alpha h_\alpha^\gamma = u^\gamma$ and thus

$$K_{\alpha\beta}u^\alpha u^\beta = u^\alpha u^\beta \nabla_\alpha N_\beta = u^\alpha \nabla_\alpha (u^\beta N_\beta) - (u^\alpha \nabla_\alpha u^\beta)N_\beta = 0 \;\; . \tag{18.43}$$

One also sees that the geodesics need not be affinely parametrised for this to hold; if one has $u^\alpha \nabla_\alpha u^\beta \sim u^\beta$, one still has $K_{\alpha\beta}u^\alpha u^\beta = 0$.

Roughly speaking this says that geodesics do not "bend" in the ambient space, and this can be made more precise in the context of the extrinsic geometry of surfaces of higher codimension (like curves). In that case (which we will not develop here), one can define extrinsic curvatures associated with all of the normal directions to the surface ("$K_{\alpha\beta}$ takes values in the normal bundle"), and for a curve the geodesic equation turns out to be the condition that this extrinsic curvature tensor vanishes.

We will briefly return to this issue in the next section.

18.3 Extrinsic Curvature and the Normal Components of the Connection

In section 16.2 we had already seen that the tangential projection of the covariant derivative of a tangential vector field V^α, i.e. $V^\alpha N_\alpha = 0$ or $V^\alpha = E^\alpha_a v^a$, agrees with the intrinsic covariant derivative defined by the induced metric h_{ab}, i.e.

$$\nabla^{(h)}_a v_b = (\bar{\nabla}_\alpha V_\beta) E^\alpha_a E^\beta_b \equiv \bar{\nabla}_a v_b \quad . \tag{18.44}$$

where

$$\bar{\nabla}_\alpha V_\beta = h_\alpha^{\ \gamma} h_\beta^{\ \delta} \nabla_\gamma V_\delta \quad . \tag{18.45}$$

The extrinsic curvature captures other components of the covariant derivative:

1. For example, if the vector V^α is not tangent to Σ, then one can decompose V_β into a tangent and normal part according to

   $$V^\beta = E^\beta_a v^a + \epsilon (N^\gamma V_\gamma) N^\beta \quad . \tag{18.46}$$

   Then there is a 2nd contribution to (18.44), namely

   $$\begin{aligned} E^\alpha_a E^\beta_b \nabla_\alpha V_\beta &= \bar{\nabla}_a v_b + E^\alpha_a E^\beta_b \nabla_\alpha (\epsilon (N^\gamma V_\gamma) N_\beta) \\ &= \bar{\nabla}_a v_b + \epsilon (N^\gamma V_\gamma) E^\alpha_a E^\beta_b \nabla_\alpha N_\beta \\ &= \bar{\nabla}_a v_b + \epsilon (N^\gamma V_\gamma) K_{ab} \quad . \end{aligned} \tag{18.47}$$

2. The extrinsic curvature tensor also enters when one inquires about the normal component of the simply-projected quantity $E^\alpha_a \nabla_\alpha V_\beta$ (with V_β again assumed to be tangential, say, $V_\beta N^\beta = 0$). This normal component is given by the scalar product with N^β, and can be written as

   $$(E^\alpha_a \nabla_\alpha V_\beta) N^\beta = -E^\alpha_a V^\beta \nabla_\alpha N_\beta = -K_{ab} V^b \tag{18.48}$$

   so that one has the decomposition

   $$E^\alpha_a \nabla_\alpha V^\beta = \bar{\nabla}_a v^b E^\beta_b - \epsilon K_{ab} v^b N^\beta \quad . \tag{18.49}$$

   Both (18.47) and (18.49) illustrate that the extrinsic curvature tensor is essentially the same as a Σ-tensorial repackaging of the normal components of the connection (Christoffel symbols $\Gamma^\alpha_{\ \beta\gamma}$), as already anticipated in (18.33).

3. Note that (18.48) implies that for a vector field V^α tangent to Σ, $V^\alpha N_\alpha = 0$ one has
$$K_{\alpha\beta}V^\alpha V^\beta \quad (\text{or } \; K_{ab}v^a v^b) \quad = -(V^\alpha \nabla_\alpha V^\beta)N_\beta \;\; , \tag{18.50}$$
and thus in particular
$$V^\alpha \nabla_\alpha V^\beta = 0 \quad \Rightarrow \quad K_{\alpha\beta}V^\alpha V^\beta = 0 \;\; , \tag{18.51}$$
which is the statement (18.43) already established in the previous section.

4. From (18.49) one can deduce a stronger statement. Namely, contracting with v^a one has
$$V^\alpha \nabla_\alpha V^\beta = (v^a \bar{\nabla}_a v^b)E^\beta_b - \epsilon K_{ab}v^a v^b N^\beta \;\; . \tag{18.52}$$
Thus we see that all geodesics on Σ (with respect to the induced metric h_{ab}) are also geodesics of the embedding space if and only if $K_{\alpha\beta} = 0$. Such hypersurfaces are called *totally geodesic*. In particular, by the comment in the previous section, made in connection with (18.40), constant time surfaces in static space-times are examples of such totally geodesic hypersurfaces.

18.4 Gauss-Codazzi Equations

Similar manipulations to those performed above allow one to obtain the so-called *Gauss-Codazzi equations*, which express certain components of the space-time curvature tensor (restricted to Σ) in terms of the intrinsic and extrinsic curvatures $\bar{R}_{\alpha\beta\gamma\delta}$ and $K_{\alpha\beta}$ (or $\bar{R}_{abcd}$ and K_{ab}) of the hypersurface Σ.

We first consider the space-time Riemann tensor with purely spatial components and its relation to the intrinsic Riemann curvature tensor $\bar{R}_{\alpha\beta\gamma\delta}$ (or $\bar{R}_{abcd}$) of the metric $h_{\alpha\beta}$ (or h_{ab}) on Σ. For example, if V^α is tangent to Σ, then one can define the Riemann curvature tensor of $h_{\alpha\beta}$ by
$$[\bar{\nabla}_\alpha, \bar{\nabla}_\beta]V^\gamma = \bar{R}^\gamma_{\;\;\delta\alpha\beta}V^\delta \tag{18.53}$$
where the term $\bar{\nabla}_\alpha \bar{\nabla}_\beta V^\gamma$ on the left-hand side is the fully projected expression
$$\bar{\nabla}_\alpha \bar{\nabla}_\beta V^\gamma = h^{\alpha'}_\alpha h^{\beta'}_\beta h^\gamma_{\gamma'} \nabla_{\alpha'}(h^\delta_{\beta'} h^{\gamma'}_\epsilon \nabla_\delta V^\epsilon) \;\; . \tag{18.54}$$

Analysing this bit by bit, for example the term $h^{\alpha'}_\alpha h^{\beta'}_\beta \nabla_{\alpha'} h^\delta_{\beta'}$ evaluates to
$$\begin{aligned} h^{\alpha'}_\alpha h^{\beta'}_\beta \nabla_{\alpha'} h^\delta_{\beta'} &= -\epsilon h^{\alpha'}_\alpha h^{\beta'}_\beta \nabla_{\alpha'}(N_{\beta'}N^\delta) \\ &= -\epsilon h^{\alpha'}_\alpha h^{\beta'}_\beta (\nabla_{\alpha'}N_{\beta'})N^\delta \\ &= -\epsilon K_{\alpha\beta}N^\delta \end{aligned} \tag{18.55}$$

and this vanishes after anti-symmetrisation in α and β because $K_{\alpha\beta}$ is symmetric. Another contribution is (by the same calculation as above, and using $N^\alpha V_\alpha = 0$)

$$\begin{aligned} h_\alpha^{\alpha'} h_\gamma^{\gamma'} (\nabla_{\alpha'} h_{\gamma'}^\epsilon) h_\beta^{\beta'} h_{\beta'}^\delta \nabla_\delta V_\epsilon &= -\epsilon K_{\alpha\gamma} N^\epsilon h_\beta^\delta \nabla_\delta V_\epsilon \\ &= +\epsilon K_{\alpha\gamma} V^\epsilon h_\beta^\delta \nabla_\delta N_\epsilon \\ &= +\epsilon K_{\alpha\gamma} K_{\beta\delta} V^\delta \quad . \end{aligned} \tag{18.56}$$

Therefore

$$\bar{R}_{\gamma\delta\alpha\beta} = h_\gamma^{\gamma'} h_\delta^{\delta'} h_\alpha^{\alpha'} h_\beta^{\beta'} R_{\gamma'\delta'\alpha'\beta'} + \epsilon(K_{\gamma\alpha} K_{\beta\delta} - K_{\gamma\beta} K_{\alpha\delta}) \quad . \tag{18.57}$$

This result can also be written in terms of Σ-tensors as

$$\bar{R}_{abcd} = E_a^\alpha E_b^\beta E_c^\gamma E_d^\delta R_{\alpha\beta\gamma\delta} + \epsilon(K_{ac} K_{bd} - K_{ad} K_{bc}) \quad . \tag{18.58}$$

It thus expresses the purely tangential components of the space-time curvature tensor in terms of the intrinsic and and extrinsic curvature tensors of Σ.

It requires significantly less effort to express the component of the space-time Riemann tensor with 1 normal component and 3 tangential components in terms of the extrinsic curvature. Indeed, simply calculating $\bar{\nabla}_{[\gamma} K_{\beta]\alpha}$ one finds on the nose

$$\begin{aligned} \bar{\nabla}_\gamma K_{\beta\alpha} - \bar{\nabla}_\beta K_{\gamma\alpha} &= h_\alpha^{\alpha'} h_\beta^{\beta'} h_\gamma^{\gamma'} (\nabla_{\gamma'} K_{\beta'\alpha'} - \nabla_{\beta'} K_{\gamma'\alpha'}) \\ &= h_\alpha^{\alpha'} h_\beta^{\beta'} h_\gamma^{\gamma'} (\nabla_{\gamma'} \nabla_{\beta'} N_{\alpha'} - \nabla_{\beta'} \nabla_{\gamma'} N_{\alpha'}) \\ &= h_\alpha^{\alpha'} h_\beta^{\beta'} h_\gamma^{\gamma'} (-R^\delta{}_{\alpha'\gamma'\beta'} N_\delta) \end{aligned} \tag{18.59}$$

or

$$R_{\delta\alpha\beta\gamma} N^\delta E_a^\alpha E_b^\beta E_c^\gamma = \bar{\nabla}_c K_{ab} - \bar{\nabla}_b K_{ac} \quad . \tag{18.60}$$

<u>Remarks:</u>

1. We could have set up the calculations in such a way that we obtain directly the Σ-tensorial form (18.58) or (18.60) of the results, by starting with $\bar{\nabla}_a \bar{\nabla}_b v_c$, say, but then we would have had to deal with covariant derivatives of the E_a^α at intermediate stages of the calculation - the derivation given above appears to be somewhat simpler in that respect (but this may be a matter of taste).

2. In particular, for a hypersurface embedded into a flat (Euclidean or Minkowski) space from (18.58) one has the relation

$$\bar{R}_{abcd} = \epsilon(K_{ac} K_{bd} - K_{ad} K_{bc}) \tag{18.61}$$

   between the intrinsic and extrinsic curvature tensors. This can be directly verified e.g. for the sphere $S^n \subset \mathbb{R}^{n+1}$ of radius L, for which one has $\epsilon = +1$ and

$$\bar{R}_{abcd} = L^{-2}(g_{ac} g_{bd} - g_{ad} g_{bc}) \quad , \quad K_{ab} = L^{-1} g_{ab} \quad . \tag{18.62}$$

We also see that if the induced metric on such a hypersurface is flat, then necessarily the extrinsic curvature tensor is also zero. This also substantiates the claim, made in the introduction to this section, section 18.1, that there can be no embedding of the flat torus T^2 into $\mathbb{R}^3$, because such an embedding would have to be both intrinsically and extrinsically flat.

3. Note that the above (purely tangential, or 3 tangential and 1 normal) components of the Riemann tensor could be expressed in terms of $\bar{R}_{abcd}$, K_{ab} and $\bar{\nabla}K_{ab}$, i.e. in terms of the tangential and 1st normal derivatives of the metric. In general the Riemann tensor depends on all the second derivatives of the metric, in particular also on the second normal derivatives of the metric. Thus the remaining components of the Riemann tensor (with two normal and two tangential directions) are more complicated and cannot be expressed solely in terms of the intrinsic and extrinsic curvatures of Σ and their tangential derivatives, and we will not determine their explicit form here.

4. In particular, the space-time Ricci tensor
$$R_{\alpha\beta} = g^{\gamma\delta}R_{\gamma\alpha\delta\beta} = h^{\gamma\delta}R_{\gamma\alpha\delta\beta} + \epsilon N^\gamma N^\delta R_{\gamma\alpha\delta\beta} \tag{18.63}$$
depends explicitly on the components of the Riemann tensor with 2 normal components.

This previous remark notwithstanding, certain components of the Ricci tensor and certain components of the Einstein tensor *can* be expressed entirely in terms of the intrinsic and extrinsic curvature tensors of Σ.

For example, contracting (18.59) with $h^{\alpha\gamma}$ or $g^{\alpha\gamma}$ (this has the same effect on tangential tensors), one finds
$$\bar{\nabla}^\alpha K_{\alpha\beta} - \bar{\nabla}_\beta K^{\alpha}_{\ \alpha} = h^\alpha_\beta R_{\alpha\gamma}N^\gamma \quad \Leftrightarrow \quad R_{\alpha\beta}E^\alpha_a N^\beta = \bar{\nabla}^a K_{ab} - \bar{\nabla}_b K^{a}_{\ a} \ , \tag{18.64}$$
which expresses the mixed normal / tangential components of the space-time Ricci tensor in terms of the (tangential derivatives of the) extrinsic curvature tensor of Σ. Because N^α and E^α_a are orthogonal with respect to the space-time metric, one has the same expression for the mixed components of the Einstein tensor
$$G_{\alpha\beta} = R_{\alpha\beta} - \tfrac{1}{2}g_{\alpha\beta}R \ , \tag{18.65}$$
namely
$$G_{\alpha\beta}E^\alpha_a N^\beta = \bar{\nabla}^a K_{ab} - \bar{\nabla}_b K^{a}_{\ a} \ . \tag{18.66}$$
Moreover, from (18.63) one finds, using the symmetries of the Riemann tensor, that the Ricci scalar and the normal-normal component of the Ricci tensor can be written as
$$\begin{aligned} R &= g^{\alpha\beta}R_{\alpha\beta} = h^{\gamma\delta}h^{\alpha\beta}R_{\gamma\alpha\delta\beta} + 2\epsilon h^{\alpha\beta}N^\gamma N^\delta R_{\gamma\alpha\delta\beta} \\ R_{\alpha\beta}N^\alpha N^\beta &= h^{\alpha\beta}N^\gamma N^\delta R_{\gamma\alpha\delta\beta} \ . \end{aligned} \tag{18.67}$$

It follows that the normal-normal component of the Einstein tensor has the simple form

$$\begin{aligned} G_{\alpha\beta}N^\alpha N^\beta &= R_{\alpha\beta}N^\alpha N^\beta - \tfrac{1}{2}g_{\alpha\beta}N^\alpha N^\beta R \\ &= R_{\alpha\beta}N^\alpha N^\beta - \tfrac{1}{2}\epsilon R \\ &= -\tfrac{1}{2}\epsilon\, h^{\gamma\delta}h^{\alpha\beta}R_{\gamma\alpha\delta\beta} \quad . \end{aligned} \tag{18.68}$$

Using (18.57), this can be written as

$$\begin{aligned} -2\epsilon G_{\alpha\beta}N^\alpha N^\beta &= h^{\alpha\gamma}h^{\beta\delta}(\bar{R}_{\alpha\beta\gamma\delta} - \epsilon(K_{\alpha\gamma}K_{\beta\delta} - K_{\alpha\delta}K_{\beta\gamma})) \\ &= \bar{R} + \epsilon(K_{\alpha\beta}K^{\alpha\beta} - K^2) \quad . \end{aligned} \tag{18.69}$$

Finally, we will also derive a useful expression for the Ricci scalar. First of all, from (18.67) and (18.69) we have

$$\begin{aligned} R &= \bar{R} + \epsilon(K_{\alpha\beta}K^{\alpha\beta} - K^2) + 2\epsilon h^{\alpha\beta}N^\gamma N^\delta R_{\gamma\alpha\delta\beta} \\ &= \bar{R} + \epsilon(K_{\alpha\beta}K^{\alpha\beta} - K^2) + 2\epsilon R_{\alpha\beta}N^\alpha N^\beta \quad . \end{aligned} \tag{18.70}$$

The first two terms in this expression are already of the desired form, depending only on the intrinsic and extrinsic curvature of Σ, while the third is not. However, it turns out that, up to a total derivative, we can trade $R_{\alpha\beta}N^\alpha N^\beta$ for a term depending only on $K_{\alpha\beta}$. Indeed, it is straightforward to establish the identity

$$\begin{aligned} \nabla_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) &= (\nabla_\alpha N_\beta)\nabla^\beta N^\alpha + N^\beta[\nabla_\alpha, \nabla_\beta]N^\alpha - (\nabla_\alpha N^\alpha)\nabla_\beta N^\beta \\ &= R_{\alpha\beta}N^\alpha N^\beta + K_{\alpha\beta}K^{\alpha\beta} - K^2 \quad . \end{aligned} \tag{18.71}$$

The only minor subtlety is to verify that no normal components of $\nabla_\alpha N_\beta$ contribute to $(\nabla_\alpha N_\beta)\nabla^\beta N^\alpha$, so that one indeed has

$$(\nabla_\alpha N_\beta)\nabla^\beta N^\alpha = K_{\alpha\beta}K^{\alpha\beta} \quad , \tag{18.72}$$

and this in turn follows from $N^\beta\nabla_\alpha N_\beta = 0$ etc. With the help of this identity we can eliminate $R_{\alpha\beta}N^\alpha N^\beta$ from (18.70) and write the scalar curvature (now with the opposite sign for the K^2-term) as

$$R = \bar{R} + \epsilon(K^2 - K^{\alpha\beta}K_{\alpha\beta}) + 2\epsilon\nabla_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) \quad . \tag{18.73}$$

These relations play an important role in particular in the Hamiltonian and initial value formulations of the Einstein equations, where the first step is the choice of an initial spacelike hypersurface Σ and an accompanying $4 \to 3+1$ decomposition of the curvature tensor and the Einstein equations. This will be discussed in section 21.

C: Dynamics of the Gravitational Field

19 The Einstein Equations

19.1 Heuristics

We expect the gravitational field equations to be non-linear second order partial differential equations for the metric. If we knew more about the weak field equations of gravity (which should thus be valid near the origin of an inertial coordinate system) we could use the Einstein equivalence principle (or the principle of general covariance) to deduce the equations for strong fields.

However, we do not know a lot about gravity beyond the Newtonian limit of weak time-independent fields and low velocities, simply because gravity is so 'weak'. Hence, we cannot find the gravitational field equations in a completely systematic way and some guesswork will be required.

Nevertheless we will see that with some very natural assumptions (and the benefit of hindsight) we will arrive at an essentially unique set of equations. Further theoretical (and aesthetical) confirmation for these equations will then come from the fact that they turn out to be the Euler-Lagrange equations of the absolutely simplest action principle for the metric imaginable.

Recall that, way back, in section 1.1, we had briefly discussed the possibility of a scalar relativistic theory of gravity described by an equation of the form (1.3)

$$\Delta\phi = 4\pi G_N \rho \quad \longrightarrow \quad \Box\phi = 4\pi G_N \rho \ . \tag{19.1}$$

We had noted there that one way to render this equation (tensorially) consistent is to think of both the left and the right hand side as (00)-components of some tensor, which we expressed in (1.6) as

$$\{\text{Some tensor generalising } \Delta\phi\}_{\alpha\beta} \sim 4\pi G_N T_{\alpha\beta} \ . \tag{19.2}$$

While this appeared to be an exotic proposal back in section 1.1, we now understand that this is exactly what is required, and we have a fairly precise idea of what this tensor on the left-hand side should be.

Indeed, recall from our discussion of the Newtonian limit of the geodesic equation that the weak static field produced by a non-relativistic mass density ρ is

$$g_{00} = -(1+2\phi) \ , \tag{19.3}$$

With the identification

$$T_{00} = \rho \ , \tag{19.4}$$

the Newtonian field equation $\Delta\phi = 4\pi G_N\rho$ can now also be written as

$$\Delta g_{00} = -8\pi G_N T_{00} \ . \tag{19.5}$$

This suggests that the weak-field equations for a general energy-momentum tensor take the form

$$E_{\alpha\beta} = \quad \{\text{Some tensor generalising } (-\Delta g_{00})\}_{\alpha\beta} \quad = 8\pi G_N T_{\alpha\beta} \ , \tag{19.6}$$

where $E_{\alpha\beta}$ is constructed from the metric and its first and second derivatives.

By the Einstein equivalence principle, if this equation is valid for weak fields (i.e. near the origin of an inertial coordinate system) then also the equations which govern gravitational fields of arbitrary strength must be of this form, with $E_{\mu\nu}$ a tensor constructed from the metric and its first and second derivatives.

Another way of anticipating what form the field equations for gravity may take is via an analogy, a comparison of the geodesic deviation equations in Newton's theory and in General Relativity. Recall that in Newton's theory we have

$$\begin{aligned} \frac{d^2}{dt^2}\delta x^i &= -K^i{}_j \delta x^j \\ K^i{}_j &= \partial^i \partial_j \phi \ , \end{aligned} \tag{19.7}$$

whereas in General Relativity we have

$$\begin{aligned} (D_\tau)^2 \delta x^\mu &= -K^\mu_\nu \delta x^\nu \\ K^\mu_\nu &= R^\mu_{\lambda\nu\rho} \dot{x}^\lambda \dot{x}^\rho \ . \end{aligned} \tag{19.8}$$

Now Newton's field equation is

$$\operatorname{Tr} K \equiv \Delta\phi = 4\pi G_N \rho \ , \tag{19.9}$$

while in General Relativity we have

$$\operatorname{Tr} K = R_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \ . \tag{19.10}$$

This suggests that somehow in the gravitational field equations of General Relativity, $\Delta\phi$ should be replaced by the Ricci tensor $R_{\mu\nu}$. Note that, at least roughly, the tensorial structure of this identification is compatible with the relation between ϕ and g_{00} in the Newtonian limit, the relation between ρ and the 0-0 component T_{00} of the energy momentum tensor, and the fact that for small velocities $R_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \sim R_{00}$.

We will now turn to a somewhat more precise argument along these lines which will enable us to determine $E_{\mu\nu}$.

19.2 More Systematic Approach

Let us take stock of what we know about $E_{\mu\nu}$.

1. $E_{\mu\nu}$ is a tensor

2. $E_{\mu\nu}$ has the dimensions of a second derivative. If we assume that no new dimensionful constants enter in $E_{\mu\nu}$ then it has to be a linear combination of terms which are either linear in second derivatives of the metric or quadratic in the first derivatives of the metric. (Later on, we will see that there is the possibility of a zero derivative term, but this requires a new dimensionful constant, the *cosmological constant* Λ. Higher derivative terms or higher non-linearities could in principle appear but would only be relevant at very high energies.)

3. $E_{\mu\nu}$ is symmetric since $T_{\mu\nu}$ is symmetric.

4. Since $T_{\mu\nu}$ is covariantly conserved, the same has to be true for $E_{\mu\nu}$,

$$\nabla_\mu T^{\mu\nu} = 0 \Rightarrow \nabla_\mu E^{\mu\nu} = 0 \quad . \tag{19.11}$$

5. Finally, for a weak static gravitational field and non-relativistic matter we should find

$$E_{00} = -\Delta g_{00} \quad . \tag{19.12}$$

Now it turns out that these conditions (1)-(5) determine $E_{\mu\nu}$ *uniquely*! First of all, (1) and (2) tell us that $E_{\mu\nu}$ has to be a linear combination

$$E_{\mu\nu} = aR_{\mu\nu} + bg_{\mu\nu}R \quad , \tag{19.13}$$

where $R_{\mu\nu}$ is the Ricci tensor and R the Ricci scalar. Then condition (3) is automatically satisfied.

To implement (4), we rewrite the above as a linear combination of the Einstein tensor (8.108) and $g_{\mu\nu}R$,

$$E_{\mu\nu} = aG_{\mu\nu} + cg_{\mu\nu}R \equiv a(R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R) + cg_{\mu\nu}R \tag{19.14}$$

and recall the contracted Bianchi identity (8.106,8.107),

$$\nabla^\mu G_{\mu\nu} = 0 \quad . \tag{19.15}$$

It follows that (4) is satisfied iff $c\nabla_\nu R = c\partial_\nu R = 0$. We therefore have to require either $\nabla_\nu R = 0$ or $c = 0$. That the first possibility is ruled out (inconsistent) can be seen by taking the trace of (19.6),

$$E^\mu_{\ \mu} = (4c - a)R = 8\pi G_N T^\mu_{\ \mu} \quad . \tag{19.16}$$

Thus, R is proportional to $T^\mu_{\ \mu}$ and since this quantity need certainly not be constant for a general matter configuration, we are led to the conclusion that $c = 0$. Thus we find

$$E_{\mu\nu} = aG_{\mu\nu} \quad . \tag{19.17}$$

We can now use condition (5) to determine the constant a.

19.3 Newtonian Weak-Field Limit

By the above considerations we have determined the field equations to be of the form

$$aG_{\mu\nu} = 8\pi G_N T_{\mu\nu} \ , \tag{19.18}$$

with a some, as yet undetermined, constant. We will now consider the Newtonian weak-field limit of this equation. We need to find that G_{00} is proportional to Δg_{00} and we can then use the condition (5) to fix the value of a. The following manipulations are somewhat analogous to those we performed in section 3.3 when considering the Newtonian limit of the geodesic equation. The main difference is that now we are dealing with second derivatives of the metric rather than with just its first derivatives entering in the geodesic equation.

As in section 3.3, let us begin by stating the assumptions that we make when considering the Newtonian limit:

1. Weak Fields

   We take this to mean that there exists a coordinate system $x^\alpha = (x^0, x^i)$ in which the metric takes the form
   $$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \tag{19.19}$$
   with $\eta_{\alpha\beta}$ the standard form of the Minkowski metric and $h_{\alpha\beta}$ and its derivatives small. In practice this means that in the following we will neglect terms that are quadratic or of higher order in $h_{\alpha\beta}$.

2. Time-independent Fields

   We assume that in these coordinates the gravitational field is time-independent, i.e. that one has
   $$\partial_0 g_{\alpha\beta} = \partial_0 h_{\alpha\beta} = 0 \ . \tag{19.20}$$

3. Non-relativistic Matter Source

   This replaces the condition that particles move non-relativistically (with coordinate speeds $v \ll c$), and we take this to mean that the only non-negligible contribution to the energy-momentum tensor $T_{\alpha\beta}$ comes from the energy density $T_{00} = \rho$,
   $$T_{00} = \rho \neq 0 \quad , \quad T_{\alpha\beta} = 0 \quad \text{otherwise} \ . \tag{19.21}$$

So we need to determine

$$G_{00} = R_{00} - \tfrac{1}{2} g_{00} R \ . \tag{19.22}$$

Since the scalar curvature is at least linear in $h_{\alpha\beta}$, to leading order in $h_{\alpha\beta}$ we can replace $g_{00} \to \eta_{00} = -1$ to obtain

$$G_{00} = R_{00} + \tfrac{1}{2} R \ , \tag{19.23}$$

where here and in the remainder of this calculation equality signs signify equalities to leading order in $h_{\alpha\beta}$.

To bootstrap the calculation of G_{00}, we start from

$$T_{ij} = 0 \quad \Rightarrow \quad G_{ij} = 0 \quad \Leftrightarrow \quad R_{ij} = \tfrac{1}{2} g_{ij} R \ , \tag{19.24}$$

and by the same reasoning as above we can write this as

$$R_{ij} = \tfrac{1}{2}\delta_{ij} R \ . \tag{19.25}$$

Therefore, for the scalar curvature we find

$$R = g^{\alpha\beta} R_{\alpha\beta} = \eta^{\alpha\beta} R_{\alpha\beta} = -R_{00} + \delta^{ij} R_{ij} = -R_{00} + \tfrac{3}{2} R \ , \tag{19.26}$$

or

$$R = 2R_{00} \ . \tag{19.27}$$

Thus

$$G_{00} = R_{00} + \tfrac{1}{2} R = 2R_{00} \tag{19.28}$$

and it just remains to calculate this one component of the Ricci tensor. In the weak field limit, R_{00} is given by

$$R_{00} = R^{k}_{\ 0k0} = \delta^{ik} R_{i0k0} \ . \tag{19.29}$$

Moreover, in this limit only the linear (second derivative) part of $R_{\mu\nu\lambda\sigma}$ will contribute, not the terms quadratic in first derivatives. Thus we can use the expression (8.20) for the curvature tensor. Additionally, in the static case we can ignore all time derivatives. Then only one term (the third) of (8.20) contributes and we find

$$R_{i0k0} = -\tfrac{1}{2} g_{00},_{ik} \ , \tag{19.30}$$

and therefore

$$R_{00} = -\tfrac{1}{2}\Delta g_{00} \ . \tag{19.31}$$

Putting everything together, we get

$$G_{00} = 2R_{00} = -\Delta g_{00} \ . \tag{19.32}$$

Thus we obtain the correct functional form of E_{00} and comparison with condition (5) determines $a = +1$ and therefore $E_{\alpha\beta} = G_{\alpha\beta}$. See also section 23.3 for a somewhat more streamlined and covariant derivation of this result from the linearised Einstein equations.

19.4 Einstein Equations

We have finally arrived at the Einstein equations for the gravitational field (metric) of a matter-energy configuration described by the energy-momentum tensor $T_{\mu\nu}$. It is

$$\boxed{R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R = 8\pi G_N\, T_{\mu\nu}} \tag{19.33}$$

These are the equations that replace the Newtonian (Poisson) equation for the gravitational potential.

Another common way of writing the Einstein equations is obtained by taking the trace of (19.33), which yields

$$R - 2R = 8\pi G_N\, T^{\mu}_{\ \mu} \equiv 8\pi G_N\, T \quad , \tag{19.34}$$

and substituting this back into (19.33) to obtain

$$R_{\mu\nu} = 8\pi G_N (T_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} T) \quad . \tag{19.35}$$

In particular, for the vacuum, $T_{\mu\nu} = 0$, the Einstein equations are

$$T_{\mu\nu} = 0 \quad \Rightarrow \quad R_{\mu\nu} = 0 \quad , \tag{19.36}$$

and this condition is equivalent to the vanishing of the Einstein tensor,

$$G_{\mu\nu} = 0 \quad \Leftrightarrow \quad R_{\mu\nu} = 0 \quad . \tag{19.37}$$

A space-time metric satisfying this equation is, for obvious reasons, said to be Ricci flat.

A priori, the Einstein equations constitute 10 coupled non-linear (actually quasi-linear, since they are linear in second derivatives) second order partial differential equations for the metric $g_{\mu\nu}(x)$, which appears both in the Einstein tensor on the left-hand side of these equations as well as usually also on the right-hand side in the matter energy-momentum tensor (we will see below, in section 19.7, that these 10 equations are linked by 4 differential identities, the contracted Bianchi identities).

This is a tremendously complicated set of equations, and trying to learn and say something about general properties of solutions to these equations is very challenging.[34]

[34] The current state of knowledge and understanding of the mathematical structure of the Einstein equations, in particular regarding the properties of the Cauchy (initial value) problem for the Einstein equations, is described in detail in the awe-inspiring œuvre *General Relativity and the Einstein Equations* by Y. Choquet-Bruhat (warning: not for the faint of heart). A readable historical introduction to the Cauchy problem for the Einstein equations is given by her in Y. Choquet-Bruhat, *Beginnings of the Cauchy problem*, `arXiv:1410.3490 [gr-qc]`.

Even the vacuum Einstein equations still constitute a complicated set of non-linear coupled partial differential equations whose general solution is not, and probably will never be, known. Usually one makes some assumptions, in particular regarding the symmetries of the metric, that reduce the number of independent variables from 10 functions $g_{\alpha\beta}(x)$ of 4 variables to a smaller number of functions depending on a smaller number of variables, and which then simplify the equations to the extent that they can be analysed explicitly, either analytically, or at least qualitatively or numerically. How to do this in practice (in the simplest non-trivial situations), will be explained in detail later on in these notes.

<u>Remarks:</u>

1. With c not set equal to one, and with the convention that T_{00} is normalised such that it gives the energy-density rather than the mass-density, one finds that the factor $8\pi G_N$ on the right hand side should be replaced by

$$8\pi G_N \quad \rightarrow \quad \frac{8\pi G_N}{c^4} \quad . \tag{19.38}$$

A note on dimensions: Newton's constant has dimensions (M mass, L length, T time) $[G_N] = \mathsf{M}^{-1}\mathsf{L}^3\mathsf{T}^{-2}$ so that

$$[G_N] = \mathsf{M}^{-1}\mathsf{L}^3\mathsf{T}^{-2} \quad \Rightarrow \quad [G_N/c^4] = \mathsf{L}^{-1}\mathsf{M}^{-1}\mathsf{T}^2 \quad . \tag{19.39}$$

Moreover, an energy density $\rho = \mu c^2$, μ a mass density, has dimensions

$$[\rho] = [\mu c^2] = \mathsf{M}\mathsf{L}^{-3}\mathsf{L}^2\mathsf{T}^{-2} = \mathsf{M}\mathsf{L}^{-1}\mathsf{T}^{-2} \quad . \tag{19.40}$$

Thus

$$[\rho G_N/c^4] = \mathsf{L}^{-2} = [R_{\mu\nu}] \quad , \tag{19.41}$$

as it should be. Frequently, an alternative (and equally reasonable) convention is used in which T_{00} is a mass density, so that then $T_{tt} = c^2 T_{00}$ is the energy density. In that case, the factor on the right-hand side of the Einstein equations is $8\pi G_N/c^2$.

2. The streamlined "derivation" of the Einstein equations given here may give the misleading impression that also for Einstein this was a straighforward affair. Nothing could be further from the truth. Not only do we have the benefit of hindsight. We also have a much more systematic and advanced understanding of Riemannian geometry and tensor calculus than was available to Einstein at the time. This concerns in particular things like the contracted Bianchi identities and their importance for energy-momentum (non-)conservation and for general covariance (to be briefly discussed in section 19.7 below).[35]

[35] For an illuminating brief account of the torturous and convoluted route and crucial final stages that led Einstein (and Grossmann) to the correct field equations, see N. Straumann, *Einstein's 'Zürich Notebook' and his Journey to General Relativity*, `arXiv:1106.0900v1 [physics.hist-ph]`.

3. As an aside, note that the trace (19.34) of the Einstein equations

$$R = -8\pi G_N\, T \quad , \tag{19.42}$$

is a scalar generally covariant differential equation for the metric (but it is of course far from sufficient to determine 10 independent components of the metric up to coordinate transformations). If one assumes, however, that the space-time metric can be parametrised by a single scalar ψ, say (somewhat like in the Newtonian limit), e.g. by stipulating that it only differs from the Minkowski metric by a conformal factor, as in

$$g_{\mu\nu} = \psi^2 \eta_{\mu\nu} \quad , \tag{19.43}$$

then a scalar equation like (19.42) (the numerical constant needs to be adjusted appropriately in order to obtain the correct Newtonian limit) provides a differential equation for ψ and thus a generally covariant scalar theory of gravity. A theory of this kind, a geometrisation and covariantisation of previous scalar theories of gravity, was proposed by Einstein and Fokker in 1913/14, some two years before Einstein arrived at the final (tensorial) form of the field equations. In this theory, there is no coupling of gravity and Maxwell theory (which has a traceless energy-momentum tensor), and null lines are identical to null lines in Minkowski space (because of conformal flatness), so for either of these reasons there is no bending of lightrays by the gravitational field in such a theory.

4. As we saw before, in two and three dimensions, vanishing of the Ricci tensor implies the vanishing of the Riemann tensor. Thus in these cases, space-times are necessarily flat away from where there is matter, i.e. at points at which $T_{\mu\nu}(x) = 0$. Thus there are no true gravitational fields and no gravitational waves.

   In four dimensions, however, the situation is completely different. As we saw, the Ricci tensor has 10 independent components whereas the Riemann tensor has 20. Thus there are 10 components of the Riemann tensor which can curve the vacuum, as e.g. in the field around the sun, and a lot of interesting physics is already contained in the vacuum Einstein equations.

5. If for whatever reason one is interested in studying solutions to the matter + Einstein equations in dimensions other than $D = 3 + 1$, this is straightforward and there are just a a few small points to pay attention to:

   - The Einstein tensor, i.e. the (unique) rank-2 tensor that can be constructed from the Riemann curvature tensor which has vanishing covariant divergence, has the same form in any dimension, $G_{\mu\nu} = R_{\mu\nu} - (1/2)g_{\mu\nu}R$.
   - Likewise, what will appear on the right-hand side of the equations is the appropriate generally covariant energy-momentum tensor.

- However, in the higher-dimensional analogue of the Einstein equations the constant of proportionality between the Einstein and energy-momentum tensors should not be called $8\pi G_N$. After all, this factor was determined from the Newtonian limit of the $(3+1)$-dimensional Einstein equations where e.g. a factor 4π has, via the Poisson equation for a point mass, its origin in the fact that the area of a unit 2-sphere is 4π. Thus we will just call it κ (which is then related in a dimension-dependent way to however one wants to normalise the D-dimensional gravitational coupling constant). Thus precisely as in 4 dimensions one can write the Einstein equations as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu} R = \kappa\, T_{\mu\nu} \quad . \tag{19.44}$$

- If one wants to use the analogue of (19.35), one should pay attention to the fact that it is less symmetric with respect to (19.44) than its 4-dimensional counterpart since in $D = n+1$ dimensions it takes the form

$$R_{\mu\nu} = \kappa (T_{\mu\nu} - \frac{1}{D-2}\, g_{\mu\nu} T^{\lambda}_{\lambda}) \quad . \tag{19.45}$$

19.5 Cosmological Constant

As mentioned before, there is one more term that can be added to the Einstein equations provided that one relaxes the condition (2) that only terms quadratic in derivatives should appear. This term takes the form $\Lambda g_{\mu\nu}$. This is compatible with the condition (4) (the conservation law) provided that Λ is a constant, the *cosmological constant*. It is a dimensionful parameter with dimension $[\Lambda] = \mathsf{L}^{-2}$ one over length squared.

The Einstein equations with a cosmological constant now read

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \quad . \tag{19.46}$$

To be compatible with condition (5) ((1), (3) and (4) are obviously satisfied), Λ has to be quite small (and observationally it is very small indeed).

<u>Remarks:</u>

1. The vacuum Einstein equations with a cosmological constant read

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R = -\Lambda g_{\mu\nu} \quad . \tag{19.47}$$

Taking traces, this implies (and is equivalent to)

$$R_{\mu\nu} = \Lambda g_{\mu\nu} \quad , \tag{19.48}$$

which is the counterpart of the Ricci-flatness condition for vacuum solutions of the Einstein equations without a cosmological constant. In general, solutions to the equation $R_{\mu\nu} = c g_{\mu\nu}$ for some constant c (and either Riemannian or Lorentzian signature) are known as *Einstein manifolds* in the mathematics literature.

2. Λ gives a contribution to the energy-momentum tensor that, in Minkowski space, would be proportional to the Minkowski metric and Lorentz-invariant, thus compatible with the symmetries of the vacuum, and Λ is often said to play the role of a vacuum energy density (more precisely vacuum energy should perhaps be considered as one possible contribution to the cosmological constant - see section 38.4 for further discussion of this issue).

3. Comparing with the energy-momentum tensor of, say, a perfect fluid (see (7.70) in section 7.5 or section 35.2),

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + pg_{\mu\nu} \quad , \tag{19.49}$$

we see that Λ corresponds to the energy density and pressure values

$$\rho_\Lambda = -p_\Lambda = \frac{\Lambda}{8\pi G_N} \quad , \tag{19.50}$$

and to an energy-momentum tensor

$$T^\Lambda_{\mu\nu} = -\rho_\Lambda g_{\mu\nu} \quad . \tag{19.51}$$

Thus, depending on the sign of Λ either the energy density or the pressure is negative,

$$\Lambda < 0 \quad \Rightarrow \quad \rho_\Lambda < 0 \quad , \quad \Lambda > 0 \quad \Rightarrow \quad p_\Lambda < 0 \quad . \tag{19.52}$$

4. The cosmological constant was originally introduced by Einstein because he was unable to find static cosmological solutions without it. We will review this *Einstein Static Universe* in section 37.2. After Hubble's discovery of the expansion of the universe, a static universe fell out of fashion and the cosmological constant was no longer required.

5. However, things are not as simple as that. Just because it is not required does not mean that it is not there. In fact, one of the biggest puzzles in theoretical physics today is why the cosmological constant is so small. According to standard quantum field theory lore, the vacuum energy density should be many many orders of magnitude larger than astrophysical observations allow. Now usually in quantum field theory one does not worry too much about the vacuum energy as one can normal-order it away. However, as we know, gravity is unlike any other theory in that not only energy-differences but absolute energies matter (and cannot just be dropped).

   The question why the observed cosmological constant is so small (and recent astrophysical observations appear to favour a tiny non-zero value) is one aspect of what is known as the *Cosmological Constant Problem*. See section 38.4 for a brief discussion of this profound issue and some references.

6. We will consider the possibility that $\Lambda \neq 0$ only in the sections on cosmology (in all other applications, Λ can indeed be neglected).

The Einstein equations

$$G_{\mu\nu} = 8\pi G_N \, T_{\mu\nu} \tag{19.53}$$

can, taken at face value, be regarded as ten algebraic equations for certain traces of the Riemann tensor $R_{\mu\nu\rho\sigma}$. $R_{\mu\nu\rho\sigma}$ has, as we know, twenty independent components, so how are the other ten determined? The obvious answer, already given above, is of course that we solve the Einstein equations for the metric $g_{\mu\nu}$ and then calculate the Riemann curvature tensor of that metric.

However, this answer leaves something to be desired because it does not really provide an explanation of how the information about these other components is encoded in the Einstein equations. It is interesting to understand this because it is precisely these components of the Riemann tensor wich represent the effects of gravity in vacuum, i.e. where $T_{\mu\nu} = 0$, like tidal forces and gravitational waves.

The more insightful answer is that the information is encoded in the Bianchi identities which serve as propagation equations for the trace-free parts of the Riemann tensor away from the regions where $T_{\mu\nu} \neq 0$.

Let us see how this works. Recall from section 11.4 the decomposition of the Riemann tensor into the traceless Weyl tensor and the trace parts, the Ricci tensor and Ricci scalar,

$$\begin{aligned} C_{\mu\nu\rho\sigma} &= R_{\mu\nu\rho\sigma} \\ &\quad - \frac{1}{D-2}(g_{\mu\rho}R_{\nu\sigma} + R_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}R_{\mu\sigma} - R_{\nu\rho}g_{\mu\sigma}) \\ &\quad + \frac{1}{(D-1)(D-2)}R(g_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}g_{\mu\sigma}) \ . \end{aligned}$$

In the vacuum, $R_{\mu\nu} = 0$, and therefore at points where the energy-momentum tensor vanishes one has

$$T_{\mu\nu}(x) = 0 \quad \Rightarrow \quad R_{\mu\nu\rho\sigma}(x) = C_{\mu\nu\rho\sigma}(x) \ . \tag{19.54}$$

As anticipated, the Weyl tensor thus encodes the information about the gravitational field in vacuum.

The question thus is how $C_{\mu\nu\rho\sigma}$ is determined everywhere in space-time by an energy-momentum tensor which may be localised in some finite region of space-time. To address that question we make use of the relation (11.88) derived in section 11.4,

$$\nabla^\mu C_{\mu\nu\rho\sigma} = (D-3)\,(\nabla_\rho P_{\nu\sigma} - \nabla_\sigma P_{\nu\rho}) \ . \tag{19.55}$$

Here $P_{\mu\nu}$ is the Schouten tensor (11.81),

$$P_{\mu\nu} = \frac{1}{D-2}\left(R_{\mu\nu} - \frac{1}{2(D-1)}g_{\mu\nu}R\right) \ . \tag{19.56}$$

Using the D-dimensional Einstein equations (19.44), (19.45) to replace the Ricci tensor and Ricci scalar by the energy-momentum tensor, one now obtains a propagation equation for the Weyl tensor of the form

$$\nabla^\mu C_{\mu\nu\rho\sigma} = J_{\nu\rho\sigma} \quad , \tag{19.57}$$

where $J_{\nu\rho\sigma}$ depends only on the energy-momentum tensor and its derivatives,

$$J_{\nu\rho\sigma} = \kappa \, \frac{D-3}{D-2} \left[\nabla_\rho T_{\nu\sigma} - \nabla_\sigma T_{\nu\rho} - \frac{1}{D-1} \left[\nabla_\rho T^\lambda_{\ \lambda} g_{\nu\sigma} - \nabla_\sigma T^\lambda_{\ \lambda} g_{\nu\rho} \right] \right] \quad . \tag{19.58}$$

This is the equation which determines the Weyl tensor components in terms of the sources. It is reminiscent of the Maxwell equation

$$\nabla^\mu F_{\mu\nu} = -J_\nu \quad , \tag{19.59}$$

and provides an intuitive (as well as, if required, detailed analytical) understanding of the propagation properties of the gravitational field.

19.7 General Covariance and Significance of the Bianchi Identities

Let us try to understand in a bit more detail, but necessarily at a very superficial and unsophisticated level, the structure of the Einstein equations.

As a first step, let us do something that we should have perhaps done rightaway, namely count the number of dynamical variables and the number of equations we have:

- the dynamical variables are the components $g_{\alpha\beta}(x)$ of the metric, i.e. 10 functions of 4 variables.

- the Ricci or Einstein tensor is symmetric; therefore the Einstein equations consitute a set of ten algebraically independent second order differential equations for the metric $g_{\alpha\beta}$.

At first, this "ten 2nd order equations for ten unknowns" looks exactly right: specifying the values of the metric and its first time-derivative as initial values on some (constant "time") hypersurface, say, this should then uniquely determine the ten components of the metric in some region to the future of that hypersurface.

At second sight, however, this cannot possibly be right and the end of the story and, if true, would actually be a major disaster. After all, the Einstein equations are generally covariant. Thus, given one metric that is a solution to the Einstein equations, one should be able to perform an arbitrary coordinate transformation and still have a (physically equivalent) solution to the Einstein equations. That means that the (ten?) Einstein equations should not determine the ten components of the metric uniquely but only up to arbitrary coordinate transformations, i.e. up to four arbitrary functions of four variables.

Phrased in terms of initial values, one should be able to perform arbitrary time-dependent coordinate transformations on a solution, but if these coordinate transformations happen to be the identity transformation on the initial hypersurface, then these solutions related by (future) coordinate transformations should arise from the same initial data.

Either way we should expect only six independent generally covariant equations for the metric, determining the 10 components of the metric up to 4 arbitary functions. How does that happen? Here we should recall the contracted Bianchi identities. They tell us that

$$\nabla^{\alpha} G_{\alpha\beta} = 0 \quad . \tag{19.60}$$

We see that, even though the ten Einstein equations are algebraically independent, there are actually four differential relations among them, so this is just right.

It is no coincidence, by the way, that the Bianchi identities come to the rescue of general covariance. We will see in section 20.6 that the Bianchi identities can in fact be understood as a consequence of the general covariance of the Einstein equations (and of the corresponding action principle).

The general covariance of the Einstein equations is reflected in the fact that only six of the ten equations are truly dynamical 2nd-order differential equations while four of them constrain the initial values of the fields on some spacelike hypersurface. Indeed, in terms of some choice $(x^{\alpha}) = (t, x^k)$ of time and space coordinates, the Bianchi identities

$$\nabla_{\alpha} G^{\alpha\beta} = \partial_{\alpha} G^{\alpha\beta} + \Gamma^{\alpha}_{\alpha\gamma} G^{\beta\gamma} + \Gamma^{\beta}_{\alpha\gamma} G^{\alpha\gamma} = 0 \tag{19.61}$$

can be written as

$$\partial_t G^{t\beta} = -\partial_k G^{k\beta} - \Gamma^{\alpha}_{\alpha\gamma} G^{\beta\gamma} - \Gamma^{\beta}_{\alpha\gamma} G^{\alpha\gamma} \quad . \tag{19.62}$$

Since the 3 terms on the right-hand side contain at most 2nd time derivatives of the metric, the 4 components $G_{t\beta}$ of the Einstein tensor can contain at most 1st time derivatives of the metric. Thinking of initial data as being given by the metric and its 1st time-derivative on some initial hypersurface, this means that the components $G_{t\beta} = 0$ of the Einstein equations (or their counterpart in the presence of matter) impose constraints on these initial data and do not provide evolution equations for these initial data.

The perhaps more familiar counterpart of these constraints in the case of Maxwell theory is the Gauss Law constraint $\vec{\nabla}.\vec{E} = 0$, which arises as the 0-component of the Maxwell equations $\partial_a F^{ab} = 0$,

$$\partial_a F^{a0} = \partial_k F^{k0} = -\vec{\nabla}.\vec{E} = 0 \quad , \tag{19.63}$$

and which also involves at most 1st time-derivatives of the dynamical field (the gauge field), and thus constitutes a constraint on the initial conditions rather than a true evolution equation.

In this case, the obvious ("contracted Bianchi") identity

$$\partial_a \partial_b F^{ab} = 0 \quad \Leftrightarrow \quad \partial_0(\partial_i F^{i0}) = -\partial_k(\partial_a F^{ak}) \tag{19.64}$$

implies

1. that the 4 Maxwell equations are not independent (as required by gauge invariance as they should only determine the 4 components A_a of the gauge field up to gauge transformations) and

2. that the Gauss Law contraint equation is "propagated", i.e. that by virtue of the true equations of motion it will hold at all times if it holds initially:

$$(\partial_a F^{ak})|_{t=0} = 0 \quad \Rightarrow \quad (\partial_k(\partial_a F^{ak}))_{t=0} = 0 \quad \Rightarrow \quad (\partial_0(\partial_i F^{i0}))_{t=0} = 0 \ \ , \tag{19.65}$$

   and likewise for the higer t-derivatives,

$$(\partial_0 \partial_a F^{ak})|_{t=0} = 0 \quad \Rightarrow \quad (\partial_k \partial_0(\partial_a F^{ak}))_{t=0} = 0 \quad \Rightarrow \quad (\partial_0^2(\partial_i F^{i0}))_{t=0} = 0 \ \ , \tag{19.66}$$

   etc. Thus if the true dynamical equations are satisfied at all times, the constraints will be satisfied at all times provided that they are satisfied initially.

Analogously, for the Einstein equations the contracted Bianchi identity in the form (19.62) implies not only 4 relations among the 10 field equations (as required by general covariance) but also that the constraints of general relativity are again "propagated" in this sense. One simple way to see this (or that this is plausible, at least - in order to prove a theorem one would need te be more precise about the initial value formulation and make sure that it leads to a well-defined time evolution etc.) is to note that by (19.62) $G^{\alpha\beta} = 0$ at $t = t_0$ (thus also $\partial_k G^{\alpha\beta} = 0$ at $t = t_0$) implies

$$G^{\alpha\beta}|_{t=t_0} = 0 \quad \Rightarrow \quad (\partial_t G^{t\beta})|_{t=t_0} = 0 \tag{19.67}$$

(and likewise for higher t-derivatives).

We will discuss this and related issues in some more detail from a slightly different (Hamiltonian) perspective in section 21.

20 Einstein Equations from an Action Principle

20.1 Einstein-Hilbert Action

To increase our confidence that the Einstein equations we have derived above are in fact reasonable and almost certainly correct, we can adopt a more modern point of view. We can ask if the Einstein equations follow from an action principle or, alternatively, what would be a natural action principle for the metric.

After all, for example in the construction of the Standard Model, one also does not start with the equations of motion but one writes down the simplest possible Lagrangian with the desired field content and symmetries.

We will start with the gravitational part, i.e. the Einstein tensor $G_{\alpha\beta}$ of the Einstein equations, and deal with the matter part, the energy-momentum tensor $T_{\alpha\beta}$, later.

By general covariance, an action for the metric $g_{\alpha\beta}$ will have to take the form

$$S = \int \sqrt{g} d^4x \; \Phi(g_{\alpha\beta}) \;\; , \tag{20.1}$$

where Φ is a scalar constructed from the metric. So what is Φ going to be? Clearly, the simplest choice is the Ricci scalar R, and this is also the unique choice if one is looking for a scalar constructed from not higher than second derivatives of the metric. Therefore we postulate the beautifully simple and elegant action

$$\boxed{S_{EH}[g_{\alpha\beta}] = \int \sqrt{g} d^4x \; R} \tag{20.2}$$

known as the Einstein-Hilbert action. It was presented by Hilbert practically on the same day that Einstein presented his final form (19.33) of the gravitational field equations. Discussions regarding who did what first and who deserves credit for what have been a favourite occupation of historians of science ever since. However, Hilbert's work would certainly not have been possible without Einstein's realisation that gravity should be regarded not as a force but as a property of space-time and his physical insight that Riemannian geometry and tensor analysis provide the correct framework for embodying the equivalence principle. Regarding the action principle for general relativity, in his superb scientific biography of Einstein, A. Pais says

> Hilbert was not the first to apply this principle to gravitation. Lorentz had done it before him. So had Einstein, a few weeks earlier. Hilbert was the first, however, to state this principle correctly.[36]

[36] A. Pais, *Subtle is the Lord* (chapter 14.d, which also contains a detailed account of the interaction between Einstein and Hilbert in the crucial November 1915 period).

We will now prove that the Euler-Lagrange equations following from the Einstein-Hilbert Lagrangian indeed give rise to the Einstein tensor and the vacuum Einstein equations. It is truly remarkable, that such a simple Lagrangian is capable of explaining practically all known gravitational, astrophysical and cosmological phenomena.

Before turning to a proof of this statement, I need to make one preliminary remark:

In this discussion we will at first ignore total derivative (or boundary) terms that one picks up from integration by parts of the variations and concentrate on the bulk Euler-Lagrange equations of motion. In standard variational problems one usually justifies this by appealing to the fact that one can e.g. choose the variations of the fields to vanish on the boundary and that therefore such boundary terms are zero. In the case at hand, things are a bit more complicated since the boundary terms that one picks up in the process of performing the variations turns out to depend both on the variation of the field (i.e. the metric) on the boundary and on its normal derivative, and it is not consistent to require both to be zero (i.e. to impose both Dirichlet and Neumann boundary conditions). This whole issue is interesting in its own right and warrants a separate discussion, and therefore we will deal with it afterwards, in sections 20.4 and 20.5.

Returning to the Einstein-Hilbert action, we now need to determine its behaviour under a variation of the metric. Since the Ricci scalar is $R = g^{\alpha\beta}R_{\alpha\beta}$, it turns out to be more convenient to consider variations $\delta g^{\alpha\beta}$ of the inverse metric instead of $\delta g_{\alpha\beta}$. This is of course equivalent, the two variations being related by

$$\delta(g^{\alpha\beta}g_{\beta\gamma}) = \delta(\delta^\alpha_\gamma) = 0 \quad \Rightarrow \quad \delta g^{\alpha\beta} = -g^{\alpha\gamma}(\delta g_{\gamma\delta})g^{\delta\beta} \quad . \tag{20.3}$$

Thus, as a first step we write

$$\begin{aligned} \delta S_{EH} &= \delta \int \sqrt{g} d^4x \; g^{\alpha\beta}R_{\alpha\beta} \\ &= \int d^4x \; \Big((\delta\sqrt{g})g^{\alpha\beta}R_{\alpha\beta} + \sqrt{g}(\delta g^{\alpha\beta})R_{\alpha\beta} + \sqrt{g}g^{\alpha\beta}\delta R_{\alpha\beta} \Big) \quad . \end{aligned} \tag{20.4}$$

Now we make use of the identity (5.77)

$$\delta\sqrt{g} = \tfrac{1}{2}\sqrt{g}g^{\alpha\beta}\delta g_{\alpha\beta} = -\tfrac{1}{2}\sqrt{g}g_{\alpha\beta}\delta g^{\alpha\beta} \quad . \tag{20.5}$$

to deduce

$$\begin{aligned} \delta S_{EH} &= \int \sqrt{g} d^4x \; [(-\tfrac{1}{2}g_{\alpha\beta}R + R_{\alpha\beta})\delta g^{\alpha\beta} + g^{\alpha\beta}\delta R_{\alpha\beta}] \\ &= \int \sqrt{g} d^4x \; (R_{\alpha\beta} - \tfrac{1}{2}g_{\alpha\beta}R)\delta g^{\alpha\beta} + \int \sqrt{g} d^4x \; g^{\alpha\beta}\delta R_{\alpha\beta} \quad . \end{aligned} \tag{20.6}$$

The first term all by itself would already give the Einstein tensor. Thus we need to show that the second term is a total derivative. I do not know of any particularly elegant argument to establish this (in a coordinate basis - written in terms of differential forms this would be completely obvious), so this will require a little bit of work, but it is not difficult.

Postponing the proof of this statement to the next section 20.2, we have established that (ignoring boundary terms) the variation of the Einstein-Hilbert action gives the gravitational part (left hand side) of the Einstein equations,

$$\delta S_{EH}[g_{\alpha\beta}] = \delta \int \sqrt{g} d^4x \; R = \int \sqrt{g} d^4x \; G_{\alpha\beta} \delta g^{\alpha\beta} + \oint \ldots \quad . \tag{20.7}$$

<u>Remarks:</u>

1. If one wants to include the cosmological constant Λ, then the action gets modified to

   $$S_{EH,\Lambda} = \int \sqrt{g} d^4x \; (R - 2\Lambda) \quad . \tag{20.8}$$

   Indeed, the only effect of including Λ is to replace $R \to R - 2\Lambda$ in the Einstein equations, so that

   $$G_{\alpha\beta} = R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R \to R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta}(R - 2\Lambda) = G_{\alpha\beta} + \Lambda g_{\alpha\beta} \quad , \tag{20.9}$$

   which gives rise to the modified Einstein equation (19.46).

2. Of course, once one is working at the level of the action, it is easy to come up with covariant generalisations of the Einstein-Hilbert action, such as

   $$S = \int \sqrt{g} d^4x \; (R + c_1 R^2 + c_2 R_{\alpha\beta} R^{\alpha\beta} + c_3 R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} + c_4 R \Box R + \ldots) \quad , \tag{20.10}$$

   with dimensionful coefficients c_k, but these invariably involve higher-derivative terms and/or higher non-linearities and are therefore irrelevant for low-energy physics and thus the world we live in. Such terms could be relevant for the early universe, however, and are also typically predicted by quantum theories of gravity like string theory.

3. A particular class of such higher-order actions has attracted some attention. As already briefly mentioned at the end of section 8.8, in $D > 4$ space-time dimensions there are other candidate tensors that could replace the Einstein tensor, provided that one is willing to give up linearity of the 2nd derivative terms of the metric. These tensors can be obtained from a variational principle involving very special linear combinations of higher order terms in the action, e.g. the *Gauss-Bonnet* term

   $$L_{GB} = R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} - 4 R^{\alpha\beta} R_{\alpha\beta} + R^2 \quad . \tag{20.11}$$

   In $D = 4$ this term is (locally) a total derivative, and thus does not contribute to the equations of motion. It is non-trivial in $D > 4$, however, but nevertheless (a priori totally non-obviously) leads to equations of motion that are no higher than 2nd order in derivatives.[37]

20.2 Appendix: A Formula for the Variation of the Ricci Tensor

The purpose of this technical appendix to the previous section is to derive a formula for the metric variation of the Ricci tensor which shows that indeed $g^{\alpha\beta}\delta R_{\alpha\beta}$ is a total derivative.

First of all, we need the explicit expression for the Ricci tensor in terms of the Christoffel symbols, which can be obtained by contraction of (8.5),

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\ \mu\nu} - \partial_\nu \Gamma^\lambda_{\ \mu\lambda} + \Gamma^\lambda_{\ \lambda\rho}\Gamma^\rho_{\ \nu\mu} - \Gamma^\lambda_{\ \nu\rho}\Gamma^\rho_{\ \lambda\mu} \ . \tag{20.12}$$

Now we need to calculate the variation of $R_{\mu\nu}$. We will not require the explicit expression in terms of the variations of the metric, but only in terms of the variations $\delta\Gamma^\mu_{\ \nu\lambda}$ induced by the variations of the metric. This simplifies things considerably.

Obviously, $\delta R_{\mu\nu}$ will then be a sum of six terms,

$$\delta R_{\mu\nu} = \partial_\lambda \delta\Gamma^\lambda_{\ \mu\nu} - \partial_\nu \delta\Gamma^\lambda_{\ \mu\lambda} + \delta\Gamma^\lambda_{\ \lambda\rho}\Gamma^\rho_{\ \nu\mu} + \Gamma^\lambda_{\ \lambda\rho}\delta\Gamma^\rho_{\ \nu\mu} - \delta\Gamma^\lambda_{\ \nu\rho}\Gamma^\rho_{\ \lambda\mu} - \Gamma^\lambda_{\ \nu\rho}\delta\Gamma^\rho_{\ \lambda\mu} \ . \tag{20.13}$$

Now the crucial observation is that $\delta\Gamma^\mu_{\ \nu\lambda}$ is a tensor. This follows from the arguments given in section 5.4, but I will repeat it here in the present context. Of course, we know that the Christoffel symbols themselves are not tensors, because of the inhomogeneous (second derivative) term appearing in the transformation rule under coordinate transformations, but this term is independent of the metric. Thus the metric variation of the Christoffel symbols indeed transforms as a tensor.

This can also be confirmed by explicit calculation. Just for the record, I will give an expression for $\delta\Gamma^\mu_{\ \nu\lambda}$ which is easy to remember as it takes exactly the same form as the definition of the Christoffel symbol, only with the metric replaced by the metric variation and the partial derivatives by covariant derivatives, i.e.

$$\delta\Gamma^\mu_{\ \nu\lambda} = \tfrac{1}{2} g^{\mu\rho}(\nabla_\nu \delta g_{\rho\lambda} + \nabla_\lambda \delta g_{\rho\nu} - \nabla_\rho \delta g_{\nu\lambda}) \ . \tag{20.14}$$

It turns out, none too surprisingly, that $\delta R_{\mu\nu}$ can be written rather compactly in terms of covariant derivatives of $\delta\Gamma^\mu_{\ \nu\lambda}$, namely as

$$\delta R_{\mu\nu} = \nabla_\lambda \delta\Gamma^\lambda_{\ \mu\nu} - \nabla_\nu \delta\Gamma^\lambda_{\ \mu\lambda} \ . \tag{20.15}$$

Thus one simply needs to replace the partial derivatives in (20.13) by covariant derivatives and drop the other terms that involve the undifferentiated Christoffel symbols.

[37] For a review of these so-called *Lanczos-Lovelock* models, see e.g. T. Padmanabhan, D. Kothawala, *Lanczos-Lovelock models of gravity*, `arXiv:1302.2151 [gr-qc]`.

In fact, this could not have been otherwise as (20.13) depends on the partial derivatives of the $\delta\Gamma$ but must at the same time be tensorial. The expression (20.15) is the unique possibility that fulfills these requirements. If you don't trust this argument (which essentially amounts to working at the origin of an inertial coordinate system where partial = covariant derivatives), you can also check this in detail (and thus perhaps in this way learn to trust and appreciate the quick argument):

As a first check on (20.15), note that the first term on the right hand side is manifestly symmetric and that the second term is also symmetric because of (5.49) and (6.71). To establish (20.15), one simply has to use the definition of the covariant derivative. The first term is

$$\nabla_\lambda \delta\Gamma^\lambda_{\ \mu\nu} = \partial_\lambda \delta\Gamma^\lambda_{\ \mu\nu} + \Gamma^\lambda_{\ \lambda\rho}\delta\Gamma^\rho_{\ \mu\nu} - \Gamma^\rho_{\ \mu\lambda}\delta\Gamma^\lambda_{\ \rho\nu} - \Gamma^\rho_{\ \nu\lambda}\delta\Gamma^\lambda_{\ \rho\mu} \ , \tag{20.16}$$

which takes care of the first, fourth, fifth and sixth terms of (20.13). The remaining terms are

$$-\partial_\nu \delta\Gamma^\lambda_{\ \mu\lambda} + \delta\Gamma^\lambda_{\ \lambda\rho}\Gamma^\rho_{\ \nu\mu} = -\nabla_\nu \delta\Gamma^\lambda_{\ \mu\lambda} \ , \tag{20.17}$$

which establishes (20.15).

What we really need is the contraction $g^{\mu\nu}\delta R_{\mu\nu}$, which we can now write as

$$\begin{aligned} g^{\mu\nu}\delta R_{\mu\nu} &= \nabla_\lambda (g^{\mu\nu}\delta\Gamma^\lambda_{\ \mu\nu}) - \nabla_\nu (g^{\mu\nu}\delta\Gamma^\lambda_{\ \lambda\mu}) \\ &= \nabla_\lambda \left(g^{\mu\nu}\delta\Gamma^\lambda_{\ \mu\nu} - g^{\mu\lambda}\delta\Gamma^\nu_{\ \nu\mu} \right) \ . \end{aligned} \tag{20.18}$$

This establishes the claim that this term is a total derivative and hence gives rise to a boundary term in the variation of the Einstein-Hilbert action, a boundary term that does, however, require further discussion - see sections 20.4 and 20.5 below.

Using the explicit expression for $\delta\Gamma^\mu_{\ \nu\lambda}$ given above, we see that we can also write (20.18) rather neatly and explicitly as

$$\begin{aligned} g^{\mu\nu}\delta R_{\mu\nu} &= (\nabla^\mu\nabla^\nu - g^{\mu\nu}\Box)\delta g_{\mu\nu} \\ &= (g^{\mu\alpha}g^{\nu\beta} - g^{\mu\nu}g^{\alpha\beta})\nabla_\mu\nabla_\nu\delta g_{\alpha\beta} \\ &= \nabla_\lambda \left((g^{\lambda\alpha}g^{\nu\beta} - g^{\lambda\nu}g^{\alpha\beta})\nabla_\nu\delta g_{\alpha\beta} \right) \ . \end{aligned} \tag{20.19}$$

This result will turn out to be useful on a couple of occasions later on in these notes, e.g. for the discussion of Noether currents associated to general covariance in section 20.6 and for the derivation of the energy-momentum tensor of a non-minimally coupled scalar field in section 22.3.

One can also use the identity (20.19) to rather painlessly determine the metric variation of some more complicated Lagrangians of the type (20.10). Consider e.g. the class of Lagrangians known as "$F(R)$ Lagrangians" where the Lagrangian is (none too surprisingly) some function $F(R)$ of the scalar curvature R,

$$S = \int \sqrt{g} d^4x \ F(R) \tag{20.20}$$

(for no particularly compelling reason, at least as far as I can see ("it can be done" is *not* a compelling reason ...), a lot of work has been dedicated to such Lagrangians in the last ten years, as a quick look at the `arXiv` will reveal).

The metric variation of this action is evidently

$$\delta S = \int \sqrt{g} d^4x \, \sqrt{g} \left(-\frac{1}{2} g_{\mu\nu} F(R) \delta g^{\mu\nu} + F'(R) \delta R \right) \quad . \tag{20.21}$$

Using (20.19) in the form

$$\delta R = R_{\mu\nu} \delta g^{\mu\nu} - (\nabla_\mu \nabla_\nu - g_{\mu\nu} \Box) \delta g^{\mu\nu} \quad , \tag{20.22}$$

and assuming there are no (or ignoring) boundary terms, so that we can integrate by parts the differential operator acting on $\delta g^{\mu\nu}$ and let it act on $F'(R)$ instead, one finds

$$\delta S = \int \sqrt{g} \left(-\frac{1}{2} g_{\mu\nu} F(R) + F'(R) R_{\mu\nu} - (\nabla_\mu \nabla_\nu - g_{\mu\nu} \Box) F'(R) \right) \delta g^{\mu\nu} \quad . \tag{20.23}$$

From this one can immediately read off the vacuum field equations.

One evident consequence of this is that for non-pathological choices of $F(R)$, a solution of the vacuum Einstein equations ($R_{\mu\nu} = 0, R = 0$, such as the Schwarzschild solution) will continue to be a solution of this $F(R)$-gravity theory, so that such proposed modifications of the Einstein-Hilbert action do not immediately run afoul of precision solar system tests of general relativity.

20.3 Matter Action and the Covariant Energy-Momentum Tensor

In order to obtain the non-vacuum Einstein equations, we need to decide what the matter Lagrangian should be. Now there is an obvious choice for this. If we have matter, then in addition to the Einstein equations we also want the equations of motion for the matter fields. Therefore we should add to the Einstein-Hilbert action the standard minimally coupled matter action

$$S_M[\phi, g_{\alpha\beta}] = \int \sqrt{g} d^4x \; L_M(\phi(x), \partial_\lambda \phi(x), \ldots ; g_{\mu\nu}(x), \partial_\lambda g_{\mu\nu}(x), \ldots) \quad , \tag{20.24}$$

ϕ representing any kind of (scalar, vector, tensor, ...) matter field, obtained by suitable covariantisation of the corresponding matter action in Minkowski space via the principle of minimal coupling (section 5). Thus e.g. the matter action for a Klein-Gordon field would be (6.11),

$$S_M[\phi, g_{\alpha\beta}] = \int \sqrt{g} d^4x \left[-\tfrac{1}{2} g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi - \tfrac{1}{2} m^2 \phi^2 \right] \quad , \tag{20.25}$$

and that for Maxwell theory would be (6.51),

$$S[A_\alpha, g_{\alpha\beta}] = -\tfrac{1}{4} \int \sqrt{g} d^4x \; g^{\mu\lambda} g^{\nu\rho} F_{\mu\nu} F_{\lambda\rho} \quad . \tag{20.26}$$

Of course, the variation of the matter action with respect to the matter fields will give rise to the covariant equations of motion of the matter fields. If we now want to derive the coupled gravity-matter equations from a variational principle, then the matter contribution to the gravitational field equations (i.e. the source terms for the gravitational field) will necessarily be given by the metric variation of the matter action. As already discussed in detail in section 7.6, we may as well simply define the *covariant energy-momentum tensor* to be the source of the gravitational field equations (7.105),

$$\delta_{\mathrm{metric}} S_M[\phi, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g} d^4x \, T_{\mu\nu}\delta g^{\mu\nu} \quad \Leftrightarrow \quad T_{\mu\nu} := -\frac{2}{\sqrt{g}}\frac{\delta}{\delta g^{\mu\nu}} S_M[\phi, g_{\alpha\beta}] \ , \tag{20.27}$$

In particular, we had already seen in section 7.6 that this definition reproduces the known results in the case of a scalar field or Maxwell theory, and that in general it automatically gives a symmetric and gauge invariant tensor without the need for some improvement procedure. It is also automatically covariantly conserved on-shell as a consequence of general covariance of the matter action (cf. section 20.6).

Therefore, the complete gravity-matter action for General Relativity is

$$S[g_{\alpha\beta}, \phi] = \frac{1}{16\pi G_N} S_{EH}[g_{\alpha\beta}] + S_M[\phi, g_{\alpha\beta}] \tag{20.28}$$

with

$$\frac{\delta S[g_{\alpha\beta}, \phi]}{\delta g^{\mu\nu}} = 0 \quad \Leftrightarrow \quad G_{\mu\nu} = 8\pi G_N T_{\mu\nu} \ . \tag{20.29}$$

Remarks:

1. If one were to try to deduce the gravitational field equations by starting from a variational principle, i.e. by constructing the simplest generally covariant action for the metric and the matter fields (and this would be the modern approach to the problem, had Einstein not already solved it for us a 100 years ago), then one would also invariably be led to the above action.

   The relative numerical factor $16\pi G_N$ between the two terms would of course then not be fixed a priori, because this approach will not (and cannot possibly be expected to) determine Newton's constant. The prefactor could once again be determined by looking at the Newtonian limit of the resulting equations of motion.

2. As we saw above, a cosmological constant term can be included by adding a constant term to the Einstein-Hilbert Lagrangian. One can equally well add a constant term to the matter Lagrangian instead (and this clearly reveals its interpretation as a constant shift of the energy, e.g. by a vacuum energy contribution, of the matter fields).

As we have seen above, the variation of the Einstein-Hilbert action leads to a boundary term that depends not just on the metric but also on the derivatives of the metric. This is related to the fact that the action itself depends also on the second derivatives of the metric. Indeed, it follows from the explicit expression for the scalar curvature in terms of the Christoffel symbols and the metric, obtained by contracting (8.5),

$$R = g^{\mu\nu}(\partial_\lambda \Gamma^\lambda_{\ \mu\nu} - \partial_\nu \Gamma^\lambda_{\ \mu\lambda} + \Gamma^\lambda_{\ \lambda\rho}\Gamma^\rho_{\ \nu\mu} - \Gamma^\lambda_{\ \nu\rho}\Gamma^\rho_{\ \lambda\mu}) \ , \tag{20.30}$$

that the Einstein-Hilbert action contains terms that are quadratic in first derivatives of the metric, as well as terms that depend linearly on the second derivatives.

In ordinary Lagrangian field theory, such linear second derivative terms can usually be introduced or eliminated (depending on what one wants to achieve) by the addition of suitable boundary terms to the action. As an example, consider the action of a free scalar field (in Minkowski space, say):

- The standard action is
$$S_0[\phi] = -\tfrac{1}{2}\int \partial^\alpha\phi\partial_\alpha\phi \ . \tag{20.31}$$
When one uses this action, the boundary term arising from the variation of the action will depend on $\delta\phi$ but not on its derivatives,
$$\delta S[\phi] = \int (\Box\phi)\delta\phi - \int \partial_\alpha(\delta\phi\partial^\alpha\phi) = \int (\Box\phi)\delta\phi - \oint_\Sigma d\sigma_\alpha(\delta\phi\partial^\alpha\phi) \ , \tag{20.32}$$
where the boundary of the integration region is denoted by Σ. This is thus the appropriate action for
$$\text{Dirichlet boundary conditions:} \quad (\delta\phi)|_\Sigma = 0 \ . \tag{20.33}$$

- One can also consider the action
$$S_1[\phi] = \tfrac{1}{2}\int \phi\Box\phi \ . \tag{20.34}$$
This action differs from S_0 by a total derivative (or boundary) term,
$$S_1[\phi] = S_0[\phi] + \tfrac{1}{2}\int \partial_\alpha(\phi\partial^\alpha\phi) = S_0[\phi] + \tfrac{1}{2}\oint d\sigma_\alpha(\phi\partial^\alpha\phi) \ . \tag{20.35}$$
It will therefore give rise to the same Euler-Lagrange bulk equations of motion. In this case, however, the boundary term arising from the variation of the action will depend both on $\delta\phi$ and its derivatives,
$$\delta S_1[\phi] = \int (\Box\phi)\delta\phi + \tfrac{1}{2}\oint d\sigma_\alpha(\phi\partial^\alpha\delta\phi - \delta\phi\partial^\alpha\phi) \ . \tag{20.36}$$
There are no obvious (or at least no obviously useful) boundary conditions compatible with this form of the action.

- Another option is the action

$$S_2[\phi] = S_0[\phi] + \int \partial_\alpha(\phi\partial^\alpha\phi) = \tfrac{1}{2}\int \partial^\alpha\phi\partial_\alpha\phi + \int \phi\Box\phi \\ = 2S_1[\phi] - S_0[\phi] \ . \tag{20.37}$$

  Its variation is

$$\delta S_2[\phi] = \int (\Box\phi)\delta\phi + \oint d\sigma_\alpha(\phi\partial^\alpha\delta\phi) \ . \tag{20.38}$$

  In this case, the boundary term only depends on the normal derivative of the variation,

$$d\sigma_\alpha(\phi\partial^\alpha\delta\phi) \sim \phi N^\alpha\partial_\alpha\delta\phi \ , \tag{20.39}$$

  and therefore $S_2[\phi]$ is the appropriate choice of action for

$$\text{Neumann boundary conditions:} \quad (N^\alpha\partial_\alpha\delta\phi)|_\Sigma = 0 \ . \tag{20.40}$$

Let us now return to the case at hand, the action for gravity. In this case second derivative terms are required by general covariance of the action (since there is no scalar that can be constructed solely from the metric and its first derivatives). However, the fact that that the second derivatives appear linearly and that $g^{\alpha\beta}\delta R_{\alpha\beta}$ is a total derivative reflect the fact that these second derivatives are spurious in the sense that they can be eliminated by an integration by parts or by adding a suitable boundary or total derivative term, albeit at the expense of general covariance.

In fact, by straightforward manipulation of (20.30), using identities such as

$$\partial_\lambda g^{\mu\nu} = -g^{\mu\alpha}\,\partial_\lambda g_{\alpha\beta}\,g^{\beta\nu} = -(\Gamma^\mu_{\ \beta\lambda}g^{\beta\nu} + \Gamma^\nu_{\ \beta\lambda}g^{\beta\mu}) \tag{20.41}$$

and $\partial_\lambda(\sqrt{g}) = \sqrt{g}\,\Gamma^\mu_{\ \mu\lambda}$, one finds that the Lagrangian density $\sqrt{g}R$ can be written as

$$\sqrt{g}R = 2\sqrt{g}\,g^{\alpha\beta}(\Gamma^\mu_{\ \nu\alpha}\Gamma^\nu_{\ \mu\beta} - \Gamma^\mu_{\ \alpha\beta}\Gamma^\nu_{\ \nu\mu}) + \partial_\lambda(\sqrt{g}\,B^\lambda) \tag{20.42}$$

where

$$B^\lambda = g^{\mu\nu}\Gamma^\lambda_{\mu\nu} - g^{\mu\lambda}\Gamma^\nu_{\ \nu\mu} \ . \tag{20.43}$$

With due care, one can also write the total derivative term as

$$\partial_\lambda(\sqrt{g}\,B^\lambda) = \sqrt{g}\,\nabla_\lambda B^\lambda \ , \tag{20.44}$$

as long as one remembers that B^λ is not a tensor. Either way, we see that instead of the generally covariant Einstein-Hilbert action one can use the non-covariant but quadratic action

$$S_{EH}[g_{\alpha\beta}] \to S_E[g_{\alpha\beta}] = \int \sqrt{g}d^4x\ 2g^{\alpha\beta}(\Gamma^\mu_{\ \nu\alpha}\Gamma^\nu_{\ \mu\beta} - \Gamma^\mu_{\ \alpha\beta}\Gamma^\nu_{\ \nu\mu}) \ . \tag{20.45}$$

This action was originally considered by Einstein himself and is therefore also known as the *Einstein action*.

It will be useful for the following to write this in the form

$$S_E[g_{\alpha\beta}] = S_{EH}[g_{\alpha\beta}] - \int \sqrt{g} d^4x \; \nabla_\lambda B^\lambda = S_{EH}[g_{\alpha\beta}] - \oint d\sigma_\lambda B^\lambda \quad , \tag{20.46}$$

in which the non-covariant terms are now manifestly confined to the boundary of the region of integration. This is now a reasonably respectable action, but there is a more attractive variant of this construction which we will discuss in the next section.

20.5 Gibbons-Hawking-York Boundary Term

In section 20.2 we have seen that the metric variation of the Einstein-Hilbert action has the form

$$\delta S_{EH}[g_{\alpha\beta}] = \int \sqrt{g} d^4x \; G_{\alpha\beta}\delta g^{\alpha\beta} + \int \sqrt{g} d^4x \; \nabla_\lambda (\Delta B)^\lambda \quad , \tag{20.47}$$

where

$$\begin{aligned} (\Delta B)^\lambda &= g^{\mu\nu}\delta\Gamma^\lambda_{\ \mu\nu} - g^{\mu\lambda}\delta\Gamma^\nu_{\ \nu\mu} \\ &= (g^{\lambda\alpha}g^{\nu\beta} - g^{\lambda\nu}g^{\alpha\beta})\nabla_\nu \delta g_{\alpha\beta} \quad . \end{aligned} \tag{20.48}$$

The reason for this notation, and the relation between this object $(\Delta B)^\lambda$ and the quantity B^λ introduced in (20.43) will be explained below. The first (bulk) term gives us as the Euler-Lagrange equations the vacuum Einstein equations $G_{\alpha\beta} = 0$, while the second term is a total derivative.

Thus, when one performs the integral over a space-time region $\mathcal{V}$ bounded by the hypersurface $\partial\mathcal{V} = \Sigma$ (which we shall assume to be spacelike or timelike), upon use of the Gauss integral formula (16.47) one finds that the second (total derivative) term can be written as

$$\begin{aligned} \int_{\mathcal{V}} \sqrt{g} d^4x \; \nabla_\lambda (\Delta B)^\lambda &= \oint_\Sigma d\sigma_\lambda (\Delta B)^\lambda \\ &= \epsilon \oint_\Sigma d^n y \; \sqrt{h} \; N_\lambda (\Delta B)^\lambda \quad , \end{aligned} \tag{20.49}$$

where N_λ is the normal vector to the boundary Σ in $\mathcal{V}$ and h_{ab} is the induced metric on the boundary. Explicitly the boundary integrand is

$$N_\lambda (\Delta B)^\lambda = (N^\rho g^{\mu\nu} - N^\mu g^{\rho\nu})\nabla_\mu \delta g_{\rho\nu} \quad . \tag{20.50}$$

Using the decomposition

$$g^{\mu\nu} = h^{\mu\nu} + \epsilon N^\mu N^\nu \tag{20.51}$$

of the metric on Σ (with $N^\mu N_\mu = \epsilon$), one sees that the terms with 3 N's cancel, and one is left with

$$N_\lambda (\Delta B)^\lambda = N^\rho h^{\mu\nu}\nabla_\mu \delta g_{\rho\nu} - N^\mu h^{\rho\nu}\nabla_\mu \delta g_{\rho\nu} \quad . \tag{20.52}$$

The first term only depends on the variations $\delta g_{\rho\nu}$ on the boundary, and its tangential derivatives $h^{\mu\nu}\nabla_\mu \delta g_{\rho\nu}$. Therefore that term is zero if one imposes standard Dirichlet boundary conditions

$$\delta g_{\alpha\beta}|_\Sigma = 0 \tag{20.53}$$

on the metric at Σ. The second term, on the other hand, depends on $\delta g_{\rho\nu}$ and its normal derivative $N^\mu \nabla_\mu \delta g_{\rho\nu}$, and is therefore non-zero,

$$\delta g_{\alpha\beta}|_\Sigma = 0 \quad \Rightarrow \quad N_\lambda (\Delta B)^\lambda|_\Sigma = -N^\mu h^{\rho\nu} \nabla_\mu \delta g_{\rho\nu}|_\Sigma = -h^{\rho\nu} N^\mu \partial_\mu \delta g_{\rho\nu}|_\Sigma \ . \tag{20.54}$$

Therefore with Dirichlet boundary conditions the variation of the Einstein-Hilbert action gives rise to a non-zero boundary term,

$$\delta S_{EH}[g_{\alpha\beta}] = \int \sqrt{g} d^4x \ G_{\alpha\beta} \delta g^{\alpha\beta} + \epsilon \oint_\Sigma \sqrt{h} d^3y \ (-h^{\rho\nu} N^\mu \partial_\mu \delta g_{\rho\nu}) \ . \tag{20.55}$$

It is therefore not true that the variation of the Einstein-Hilbert action is the Einstein tensor,

$$\frac{\delta}{\delta g^{\alpha\beta}} S_{EH} \neq \sqrt{g} G_{\alpha\beta} \ . \tag{20.56}$$

In fact, the presence of this boundary term means that the functional $S_{EH}[g_{\alpha\beta}]$ is not even differentiable (in the sense of variational calculus).

The way to resolve this issue is, as for the scalar field discussed above, to add a suitable boundary term to the action itself. This will not change the bulk variation, and it turns out that e.g. for Dirichlet boundary conditions the boundary term can be chosen in such a way that its variation cancels the boundary term above.

Actually, we already have one candidate for the boundary term, namely the one relating the Einstein and Einstein-Hilbert actions in (20.46). The variation of the Einstein action is

$$\begin{aligned} \delta S_E[g_{\alpha\beta}] &= \delta S_{EH}[g_{\alpha\beta}] - \delta \oint d\sigma_\lambda B^\lambda \\ &= \int \sqrt{g} d^4x \ G_{\alpha\beta} \delta g^{\alpha\beta} + \oint_\Sigma d^ny \ \sqrt{h} \ \epsilon N_\lambda (\Delta B)^\lambda - \delta (\oint_\Sigma d^ny \ \sqrt{h} \ \epsilon N_\lambda B^\lambda) \ . \end{aligned} \tag{20.57}$$

Calculating the variation in the second term for Dirichlet boundary conditions on Σ, one finds

$$\delta (\oint_\Sigma d^ny \ \sqrt{h} \ \epsilon N_\lambda B^\lambda) = \oint_\Sigma d^ny \ \sqrt{h} \ \epsilon N_\lambda \delta B^\lambda \ , \tag{20.58}$$

with

$$\delta B^\lambda = \delta (g^{\mu\nu} \Gamma^\lambda_{\mu\nu} - g^{\mu\lambda} \Gamma^\nu{}_{\nu\mu}) \ = g^{\mu\nu} \delta \Gamma^\lambda_{\mu\nu} - g^{\mu\lambda} \delta \Gamma^\nu{}_{\nu\mu} = (\Delta B)^\lambda \ . \tag{20.59}$$

Thus for Dirichlet boundary conditions, the variation of B^λ on the boundary equals the quantity $(\Delta B)^\lambda$ (20.48) arising from the variation of the Einstein-Hilbert action. As a consequence, there is no boundary term in the variation of the Einstein action for Dirichlet boundary conditions,

$$\delta S_E[g_{\alpha\beta}] = \int \sqrt{g} d^4x \ G_{\alpha\beta} \delta g^{\alpha\beta} \tag{20.60}$$

and the Einstein action leads to a well-defined variational principle (with a differentiable action).

Although this is progress, the boundary term that one adds to the Einstein-Hilbert action to obtain the Einstein action is not particularly attractive, in particular as it is non-covariant not only with respect to bulk coordinate transformations but also with respect to boundary coordinate transformations (i.e. the integrand is not a Σ-scalar in the terminology of section 15.1).

This is something one would have to live with if one could not do better. However, such a boundary term achieving this is not unique as it is evidently only defined up to terms whose variations vanish for Dirichlet boundary conditions, in particular up to terms that only depend on the intrinsic geometry of Σ. Among these candidates there is a preferred, geometrically transparent, boundary term, the *Gibbons-Hawking-York* boundary term,

$$S_{GHY}[g_{\alpha\beta}] = 2\epsilon \oint_{\Sigma} \sqrt{h} d^3y \, K \quad . \tag{20.61}$$

Here

$$\begin{aligned} K &= h^{ab}K_{ab} = h^{\alpha\beta}\nabla_\alpha N_\beta \\ &= (g^{\alpha\beta} + N^\alpha N^\beta)\nabla_\alpha N_\beta \\ &= g^{\alpha\beta}\nabla_\alpha N_\beta = \nabla_\alpha N^\alpha \end{aligned} \tag{20.62}$$

is the trace of the extrinsic curvature K_{ab} of Σ (discussed in section 18).

One way to prove that this is a good boundary term is to determine its variation and to show that it cancels against the variation arising from the Einstein-Hilbert action for Dirichlet boundary conditions. Alternatively one should show that the above boundary term differs from the boundary term for the Einstein action only by expressions that depend on the metric $g_{\alpha\beta}$ and its tangential derivatives.

Since usually one does the former, let us do the latter. The difference between the two boundary integrands is

$$2K + N_\lambda B^\lambda = 2\nabla_\lambda N^\lambda + N_\lambda B^\lambda \quad . \tag{20.63}$$

A calculation identical to the one leading to $N_\lambda(\Delta B)^\lambda$ in (20.52) shows that

$$N_\lambda B^\lambda = N^\rho h^{\mu\nu}\partial_\mu g_{\rho\nu} - N^\mu h^{\rho\nu}\partial_\mu g_{\rho\nu} \quad . \tag{20.64}$$

Here the first term only depends on the metric and its tangential derivatives, while the second term involves normal derivatives of the metric. On the other hand, for K we have

$$2K = h^{\nu\rho}L_N g_{\nu\rho} = h^{\nu\rho}N^\mu\partial_\mu g_{\nu\rho} + 2h^\mu_{\ \nu}\partial_\mu N^\nu \quad , \tag{20.65}$$

where we have used the non-covariant way (9.39) of writing the Lie derivative,

$$L_N g_{\mu\nu} = N^\lambda\partial_\lambda g_{\mu\nu} + \partial_\mu N^\lambda g_{\lambda\nu} + \partial_\nu N^\lambda g_{\mu\lambda} \quad . \tag{20.66}$$

Thus the first term cancels the normal derivative term in (20.64) and the remaining terms in (20.64) and (20.65) only involve fields $g_{\alpha\beta}$, i.e. $h_{\alpha\beta}$ and N_α, and their tangential derivatives that are fixed on the boundary.

This establishes that the Gibbons-Hawking-York boundary term is an acceptable choice of boundary term. This is also the standard choice, and gives rise to the standard gravitational action

$$\begin{aligned} S_g[g_{\alpha\beta}] &= S_{EH}[g_{\alpha\beta}] + S_{GHY}[g_{\alpha\beta}] \\ &= \int \sqrt{g} d^4x \, R + 2\epsilon \oint_\Sigma \sqrt{h} d^3y \, K \quad , \end{aligned} \tag{20.67}$$

expressed purely in terms of the intrinsic and extrinsic scalar curvatures R and K, with

$$\delta S_g[g_{\alpha\beta}] = \int \sqrt{g} d^4x \, G_{\alpha\beta} \delta g^{\alpha\beta} \tag{20.68}$$

for Dirichlet boundary conditions.

<u>Remarks:</u>

1. In addition, one can add terms to the action that do not depend on the dynamical fields (as this will certainly change neither the variation with respect to the dynamical fields nor the equations of motion). A common choice is a kind of "background subtraction", designed to associate the numerical value $S = 0$ to a particular background metric $g^0_{\alpha\beta}$.[38] Thus one could define the "physical" action to be

$$S[g_{\alpha\beta}] = S_g[g_{\alpha\beta}] - S_g[g^0_{\alpha\beta}] \quad . \tag{20.69}$$

In particular, if one is interested in asymptotically flat space-times, say, then the appropriate reference background metric is just the flat Minkowski metric, and then (20.69) takes the simple form

$$S[g_{\alpha\beta}] = \int \sqrt{g} d^4x \, R + 2\epsilon \oint_\Sigma \sqrt{h} d^3y \, (K - K^0) \quad , \tag{20.70}$$

where K^0 is the trace of the extrinsic curvature of the boundary (isometrically) embedded into Minkowski space. In section 21.12 we will be led to a similar subtraction prescription at the level of the Hamiltonian.

2. Another way to motivate (or arrive at) the Gibbons-Hawking-York boundary term is to start from the decomposition (18.73)

$$R = \bar{R} + \epsilon(K^2 - K^{\alpha\beta}K_{\alpha\beta}) + 2\epsilon\nabla_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) \quad . \tag{20.71}$$

of the Ricci scalar provided by the Gauss-Codazzi equations. This turns out to be a convenient starting point for the canonical (Hamiltonian) formulation of general relativity, and we will therefore discuss this in that context in section 21.2.

[38] See e.g. S. Hawking, G. Horowitz, *The Gravitational Hamiltonian, Action, Entropy, and Surface Terms*, `arXiv:gr-qc/9501014`, for a brief discussion.

20.6 General Covariance and Noether Identities

The variational (i.e. action or Lagrangian based) formulation of general relativity has a number of significant conceptual and technical advantages, and we will explore some of them in this section.

- I mentioned before, in section 19.7, that it is no accident that the Bianchi identities come to the rescue of the general covariance of the Einstein equations in the sense that they reduce the number of independent equations from ten to six. We will now see that indeed the Bianchi identities are a consequence of the general covariance of the Einstein-Hilbert action.

- Virtually the same calculation will show that the covariant (metric, Hilbert) energy-momentum tensor, as defined above, is automatically conserved (on shell) by virtue of the general covariance of the matter action.

- This argument can also be turned around to show that (generically, at least) the Einstein equations imply the matter equations of motion, a very characteristic feature of generally covariant gravitational field equations.

- Simple variants of these arguments will also provide us with the Noether currents associated with the general covariance of the Einstein-Hilbert action.

- Analogous considerations, but now applied to the minimally coupled generally covariant matter action, will provide us with some insight into the relation between the (Belinfante improved) canonical and covariant energy-momentum tensors introduced in section 7.

To set the stage, we need to discuss how to express general covariance of an action, either of the Einstein-Hilbert gravitational action $S_{EH}[g_{\alpha\beta}]$ or of some (minimally) gravitationally coupled general covariant matter action $S_M[\phi, g_{\alpha\beta}]$, in a form that allows us to explore its consequences in a Lagrangian formalism.

At first, the statement that an action is generally covariant means that it is invariant under transformations $x^\alpha \to x'^\alpha$ of the coordinates (for present purposes it will for once be more convenient to use the same indices on the old and new coordinates and to distinguish transformed objects by primes), and the accompanying tensorial transformation of the fields, given e.g. by

$$\begin{aligned} x \to x' \quad \Rightarrow \quad & \phi(x) \to \phi'(x') = \phi(x) \\ & g_{\alpha\beta}(x) \to g'_{\alpha\beta}(x') = \frac{\partial x^\gamma}{\partial x'^\alpha}\frac{\partial x^\delta}{\partial x'^\beta}\, g_{\gamma\delta}(x) \quad . \end{aligned} \tag{20.72}$$

So far, this is utterly familiar, but since in the action one is integrating over the coordinates x^α, say, we would like to express this invariance not as a statement between transformed fields at x' and the old fields at x but in terms of transformations of the fields at a point x (which we can then plug into the Lagrangian or action). This means that we do not want to consider the transformation $\phi(x) \to \phi'(x')$ (and its counterpart for the metric) but rather the transformations

$$\phi(x) \to \phi'(x) \quad , \quad g_{\alpha\beta}(x) \to g'_{\alpha\beta}(x) \ . \tag{20.73}$$

In particular, for infinitesimal coordinate transformations of the form

$$x'^{\alpha} = x^{\alpha} + \epsilon\xi^{\alpha}(x) \tag{20.74}$$

or (suppressing the ϵ)

$$\delta_{\xi} x^{\alpha} = \xi^{\alpha}(x) \ , \tag{20.75}$$

the infinitesimal variations of the fields are then precisely the Lie derivatives of the fields discussed in section 9,

$$\delta_{\xi}\phi(x) = L_{\xi}\phi(x) = \xi^{\alpha}\partial_{\alpha}\phi(x) \tag{20.76}$$

for a scalar field,

$$\delta_{\xi} g_{\alpha\beta}(x) = L_{\xi} g_{\alpha\beta}(x) = \nabla_{\alpha}\xi_{\beta}(x) + \nabla_{\beta}\xi_{\alpha}(x) \ , \tag{20.77}$$

for the metric, etc.

As a reminder, a quick way to derive this transformation of the metric is to start with the tensorial transformation behaviour in the form

$$(g_{\alpha\beta}(x') - g'_{\alpha\beta}(x'))dx'^{\alpha}dx'^{\beta} = g_{\alpha\beta}(x')dx'^{\alpha}dx'^{\beta} - g_{\alpha\beta}(x)dx^{\alpha}dx^{\beta} \ , \tag{20.78}$$

and to then apply this to the infinitesimal transformation (20.74). Expanding the differentials

$$dx'^{\alpha} = dx^{\alpha} + \epsilon(\partial_{\gamma}\xi^{\alpha})dx^{\gamma} \tag{20.79}$$

and the components $g_{\alpha\beta}(x')$ of the metric to first order in ϵ,

$$g_{\alpha\beta}(x') = g_{\alpha\beta}(x) + \epsilon\xi^{\gamma}\partial_{\gamma} g_{\alpha\beta}(x) \ , \tag{20.80}$$

one finds (20.77) (in its not manifestly covariant form (9.39)).

To see that this indeed leads to a symmetry for any generally covariant action, i.e. any action of the form

$$S = \int \sqrt{g} d^4x \ L(x) \tag{20.81}$$

where L is a scalar, note that for any density $\sqrt{g}F$, F a scalar, one has, by the by now familiar identity for the variation of $\sqrt{g}$,

$$\begin{aligned} \delta_{\xi}(\sqrt{g}F) &= (\delta_{\xi}\sqrt{g})F + \sqrt{g}\delta_{\xi}F = \tfrac{1}{2}\sqrt{g}g^{\alpha\beta}(L_{\xi}g_{\alpha\beta})F + \sqrt{g}L_{\xi}F \\ &= \sqrt{g}(\nabla_{\alpha}\xi^{\alpha})F + \sqrt{g}\xi^{\alpha}\partial_{\alpha}F = \partial_{\alpha}(\sqrt{g}\ \xi^{\alpha}F) \ , \end{aligned} \tag{20.82}$$

i.e.

$$\delta_\xi(\sqrt{g}F) = \partial_\alpha(\sqrt{g}\,\xi^\alpha F) = \sqrt{g}\,\nabla_\alpha(\xi^\alpha F) \quad , \tag{20.83}$$

a result previously obtained in (9.69). Thus the variation of a generally covariant action,

$$\delta_\xi S = \int \sqrt{g}d^4x\;\nabla_\alpha(\xi^\alpha L) \quad , \tag{20.84}$$

is a total derivative and $\delta_\xi = L_\xi$ certainly generates a symmetry in the usual Lagrangian / Noether sense. On the other hand, $\delta_\xi(\sqrt{g}L)$ can also be expressed in terms of the δ_ξ-variations of the fields, and combining this with the invariance of the action (up to boundary terms), one can now deduce the consequences of general covariance for a given theory, either for ξ that vanish at the boundary of the integration region or, more generally, for ξ that are non-trivial there. We will now discuss these possibilities in turn for the Einstein-Hilbert and matter actions.

1. Contracted Bianchi Identities

   As our first application we consider the Einstein-Hilbert action, and its associated invariance under all infinitesimal coordinate transformations generated by ξ that vanish on the boundary. In that case, from the considerations above we know that the Einstein-Hilbert action is strictly invariant,

   $$\delta_\xi S_{EH} = 0 \quad . \tag{20.85}$$

   On the other hand we also know that, modulo boundary terms,

   $$\delta S_{EH} = \int \sqrt{g}d^4x\; G_{\alpha\beta}\delta g^{\alpha\beta} = -\int \sqrt{g}d^4x\; G^{\alpha\beta}\delta g_{\alpha\beta} \tag{20.86}$$

   for any metric variation. Combining these two facts we arrive at the conclusion that

   $$\begin{aligned} 0 = \delta_\xi S_{EH} &= -\int \sqrt{g}d^4x\; G^{\alpha\beta}\delta_\xi g_{\alpha\beta} \\ &= -2\int \sqrt{g}d^4x\; G^{\alpha\beta}\nabla_\alpha\xi_\beta \\ &= +2\int \sqrt{g}d^4x\; (\nabla_\alpha G^{\alpha\beta})\xi_\beta \quad . \end{aligned} \tag{20.87}$$

   Since this has to hold for all ξ (vanishing on the boundary), we deduce

   $$\delta_\xi S_{EH} = 0 \quad \forall\,\xi \quad \Rightarrow \quad \nabla^\alpha G_{\alpha\beta} = 0 \quad . \tag{20.88}$$

   As promised and anticipated, the contracted Bianchi identities are a consequence of the general covariance of the Einstein-Hilbert action.

   To further strengthen the analogy with the gauge invariance of Maxwell theory emphasised in the discussion in section 19.7, note that also the relevant "contracted Bianchi" identity $\partial_a\partial_b F^{ab} = 0$ can be derived from gauge invariance of the Maxwell action (although this is of course not the most economical way of arriving at this identity in the case at hand).

Indeed, the general variation of the Maxwell action (for variations vanishing on the boundary) is

$$\delta S[A] = \int d^4x (\partial_a F^{ab}) \delta A_b \quad . \tag{20.89}$$

The action is invariant under gauge transformations

$$\delta_\Psi A_a = \partial_a \Psi \quad , \tag{20.90}$$

i.e. one has

$$0 = \delta_\Psi S[A] = \int d^4x (\partial_a F^{ab}) \partial_b \Psi = \int d^4x (\partial_a \partial_b F^{ab}) \Psi \tag{20.91}$$

(for any Ψ vanishing at the boundary), and thus $\partial_a \partial_b F^{ab} = 0$.

2. Identically Conserved Noether Currents

The fact that one obtains kinematical identities rather than non-trivially conserved currents (non-trivial in the sense that their conservation requires the validity of some dynamical equations of motion) is a characteristic feature of Noether's theorem applied to local (gauge) symmetries. We can also see this when we consider general vector fields ξ (not constrained by the requirement that they give rise to vanishing boundary terms for the given integration domain). In that case, $\delta_\xi S_{EH}$ will not be identically zero but will be a total derivative, and also the corresponding Noether currents will turn out to be identically conserved.

Thus we now consider again the Einstein-Hilbert action, but now with ξs that are allowed to be non-zero on the boundary,

$$\delta_\xi S_{EH} = \int \sqrt{g} d^4x \; \nabla_\alpha (\xi^\alpha R) \quad . \tag{20.92}$$

By explicitly performing this variation, as above, the bulk (Einstein tensor) term is identically zero by the contracted Bianchi identity, but we obtain one total derivative term from (20.19),

$$\begin{aligned} g^{\alpha\beta} \delta_\xi R_{\alpha\beta} &= \nabla_\mu [(g^{\mu\alpha} g^{\nu\beta} - g^{\mu\nu} g^{\alpha\beta}) \nabla_\nu L_\xi g_{\alpha\beta}] \\ &= \nabla_\mu [(g^{\mu\alpha} g^{\nu\beta} - g^{\mu\nu} g^{\alpha\beta}) \nabla_\nu (\nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha)] \quad , \end{aligned} \tag{20.93}$$

and a second total derivative term

$$\nabla_\mu (-2 G^\mu{}_\nu \xi^\nu) = \nabla_\mu (-2 R^\mu{}_\nu \xi^\nu + \xi^\mu R) \tag{20.94}$$

from the integration by parts performed in the course of the calculation in (20.87). The term involving the scalar curvature is identical to, and cancels against, the scalar curvature term arising from (20.92). One is thus left with the statement that for any vector field ξ^μ and any integration domain one has

$$\int \sqrt{g} d^4x \nabla_\mu[-2R^\mu_\nu\xi^\nu + (g^{\mu\alpha}g^{\nu\beta} - g^{\mu\nu}g^{\alpha\beta})\nabla_\nu(\nabla_\alpha\xi_\beta + \nabla_\beta\xi_\alpha)] = 0 \ . \tag{20.95}$$

Thus for any ξ^μ one has the conserved Noether current

$$J^\mu(\xi) = R^\mu_\nu\xi^\nu + \tfrac{1}{2}(g^{\mu\nu}g^{\alpha\beta} - g^{\mu\alpha}g^{\nu\beta})\nabla_\nu(\nabla_\alpha\xi_\beta + \nabla_\beta\xi_\alpha) \ . \tag{20.96}$$

This is precisely the identically conserved current (13.51), (13.53)

$$J^\mu(\xi) = \nabla_\nu(\nabla^{[\mu}\xi^{\nu]}) \quad \Rightarrow \quad \nabla_\mu J^\mu(\xi) = 0 \tag{20.97}$$

already mentioned in section 13.7. We thus learn that the generalised Komar currents of that section are indeed, as anticipated there, precisely the identically conserved Noether currents associated to the general covariance of the Einstein-Hilbert action.

Note that, as mentioned above, it is a general feature of Noether currents associated to local (gauge) symmetries that they are in fact identically conserved: the current $J^\mu(\xi)$ (or its conterpart for some local symmetry of another theory) cannot possibly be conserved for all possible $\xi^\mu(x)$ unless it is actually identically conserved.[39] In particular, this implies that the conserved charges associated with these currents can always be expressed as surface integrals.

3. On-Shell Covariant Conservation of the Energy-Momentum Tensor

   Now let us play the same game with the matter action S_M (20.24). Once again, the variation $\delta_\xi S_M$, expressed in terms of the Lie derivatives $L_\xi g_{\mu\nu}$ and $\delta_\xi\phi = L_\xi\phi$ of the matter fields should be identically zero, by general covariance of the matter action (for the time being we again at first only consider ξ which are such that any boundary terms vanish). Proceeding as before, and using the definition (7.105) of the energy-momentum tensor, we find

$$\begin{aligned} 0 &= \delta_\xi S_M \\ &= \int \sqrt{g} d^4x \, (-\tfrac{1}{2}T_{\mu\nu}\delta_\xi g^{\mu\nu} + \frac{\delta L_M}{\delta\phi}\delta_\xi\phi) \\ &= -\int \sqrt{g} d^4x \, (\nabla^\mu T_{\mu\nu})\xi^\nu + \int \sqrt{g} d^4x \, \frac{\delta L_M}{\delta\phi}\delta_\xi\phi \ . \end{aligned} \tag{20.98}$$

[39] See e.g. B. Julia, S. Silva, *Currents and Superpotentials in classical gauge invariant theories I: Local results with applications to Perfect Fluids and General Relativity*, `arXiv:gr-qc/9804029` for a rather explicit elementary argument, and R. Wald, *On identically closed forms locally constructed from a field*, J. Math. Phys. 31 (1990) 2378, R. Wald, *Black Hole Entropy is Noether Charge*, `arXiv:gr-qc/9307038`, V. Iyer and R. Wald, *Some properties of Noether charge and a proposal for dynamical black hole entropy*, `arXiv:gr-qc/9403028` and references therein for related considerations and further developments.

Now once again this has to hold for all ξ, and as the second term is identically zero 'on-shell', i.e. for ϕ satisfying the matter Euler-Lagrange equations of motion $\delta L_M/\delta\phi = 0$, we deduce that

$$\delta_\xi S_M = 0 \quad \forall\, \xi \quad \Rightarrow \quad \nabla^\mu T_{\mu\nu} = 0 \quad \text{on-shell} \ . \tag{20.99}$$

This should be contrasted with the contracted Bianchi identities which are valid 'off-shell'. The more general situation with the ξ not restricted to vanish on the boundary will be analysed in detail in section 22.2.

4. Einstein Equations Imply (generically) the Matter Equations of Motion

   Note that, to a certain extent, this argument can also be turned around to show that *the equations of motion of the gravitational field generated by some matter fields imply the equations of motion of these matter fields*!

   Indeed, we already know that the Einstein equations imply covariant conservation of the energy-momentum tensor,

$$G_{\alpha\beta} = 8\pi G_N\, T_{\alpha\beta} \quad \Rightarrow \quad \nabla_\alpha T^{\alpha\beta} = 0 \ . \tag{20.100}$$

   Hence, by the above reasoning, the Einstein equations imply

$$G_{\alpha\beta} = 8\pi G_N\, T_{\alpha\beta} \quad \Rightarrow \quad \int \sqrt{g} d^4x\, \frac{\delta L_M}{\delta\phi}\delta_\xi\phi = 0 \tag{20.101}$$

   for all compactly supported ξ. Excluding certain non-generic cases (like for example a constant scalar field for which $\delta_\xi\phi = 0$), one sees that

$$G_{\alpha\beta} = 8\pi G_N\, T_{\alpha\beta} \quad \stackrel{\text{generically}}{\Longrightarrow} \quad \frac{\delta L_M}{\delta\phi} = 0 \ . \tag{20.102}$$

   This should be contrasted with the Maxwell equations in the presence of (charged) matter fields, say, which only imply current conservation,

$$\partial_\mu F^{\mu\nu} \sim j^\nu \quad \Rightarrow \quad \partial_\nu j^\nu = 0 \ , \tag{20.103}$$

   but not the complete equations of motion of the matter fields.

   To explain what is special about the generally covariant gravitational field equations in this respect, I will conclude this section with a quote by Misner, Thorne and Wheeler, since I could not possibly state this more eloquently:

   > The Maxwell equations are so constructed that they automatically fulfill and demand the conservation of charge; but not everything has charge. The Einstein field equation is so constructed that it automatically fulfills and demands the conservation of momentum-energy; and everything does have energy. The Maxwell field equations are indifferent to the interposition of an "external" force, because that force in no way threatens

the principle of conservation of charge. The Einstein field equation cares about every force, because every force is a medium for the exchange of energy.

Electromagnetism has the motto, "I count all the electric charge that's here." All that bears no charge escapes its gaze.

"I weigh all that's here" is the motto of spacetime curvature. No physical entity escapes this surveillance.[40]

20.7 First Order Form of the Action, Torsion and the Palatini Principle

For certain purposes (e.g. as a precursor to a Hamiltonian formulation) it can be useful to put an action leading to 2nd order differential equations into "1st order form" by the introduction of some auxiliarly variables. The prototypical example is Maxwell theory, whose original action (cf. section 6.5)

$$S[A] = -\tfrac{1}{4}\int d^4\xi \; F_{ab}F^{ab} \tag{20.104}$$

depends quadratically on the 1st derivatives of the gauge field A_a, and leads to the 2nd order equations of motion

$$\partial_a F^{ab} = 0 \quad . \tag{20.105}$$

To put this into 1st order form, one treats A_a and $F_{ab} \to \mathcal{F}_{ab}$ as a priori *independent* fields and considers the action

$$\begin{aligned} S[A,\mathcal{F}] &= \int d^4\xi \; \left(-(\partial_a A_b)\mathcal{F}^{ab} + \tfrac{1}{4}\mathcal{F}_{ab}\mathcal{F}^{ab}\right) \\ &= \int d^4\xi \; \left(-\tfrac{1}{2}(\partial_a A_b - \partial_b A_a)\mathcal{F}^{ab} + \tfrac{1}{4}\mathcal{F}_{ab}\mathcal{F}^{ab}\right) \quad , \end{aligned} \tag{20.106}$$

which depends purely algebraically on $\mathcal{F}_{ab}$ and only linearly on the 1st derivatives of A_a. The equations of motion arising from the variation of A_a are the 1st order equations

$$\delta A \quad \Rightarrow \quad \partial_a \mathcal{F}^{ab} = 0 \quad . \tag{20.107}$$

However these are not (yet) the vacuum Maxwell equations because $\mathcal{F}_{ab}$ is not (yet) to be identified with the Maxwell field stength tensor of A_a. This identification now results from the (purely algebraic) equation of motion associated with variations of $\mathcal{F}_{ab}$,

$$\delta\mathcal{F} \quad \Rightarrow \quad \mathcal{F}_{ab} = \partial_a A_b - \partial_b A_a = F_{ab} \quad . \tag{20.108}$$

Plugging this result into the previous equations then gives rise to the standard Maxwell equations (and plugging it into the 1st order action $S[A,\mathcal{F}]$ reduces it to the standard Maxwell action $S[A]$).

[40] C. Misner, K. Thorne, J. Wheeler, *Gravitation*, section 20.6.

Something similar (but more interesting and somewhat more subtle) can be done in the case of general relativity.

Recall from sections 5.4 and 11.5 that a priori a metric $g_{\mu\nu}$ and a connection $\tilde{\Gamma}^{\lambda}_{\ \mu\nu}$ are independent concepts, and that the notion of curvature (curvature and Ricci tensors) can be defined for an arbitrary connection,

$$R^{\lambda}_{\ \sigma\mu\nu}(\tilde{\Gamma}) = \partial_\mu \tilde{\Gamma}^{\lambda}_{\ \sigma\nu} - \partial_\nu \tilde{\Gamma}^{\lambda}_{\ \sigma\mu} + \tilde{\Gamma}^{\lambda}_{\ \mu\rho}\tilde{\Gamma}^{\rho}_{\ \nu\sigma} - \tilde{\Gamma}^{\lambda}_{\ \nu\rho}\tilde{\Gamma}^{\rho}_{\ \mu\sigma} \quad , \quad R_{\mu\nu}(\tilde{\Gamma}) = R^{\lambda}_{\ \mu\lambda\nu}(\tilde{\Gamma}) \ . \tag{20.109}$$

General relativity employs and is formulated in terms of the canonical Levi-Civita connection described by the Christoffel symbols $\tilde{\Gamma}^{\lambda}_{\ \mu\nu} = \Gamma^{\lambda}_{\ \mu\nu}$, characterised by the fact that the connection is compatible with the metric and has no torsion. It is thus easy to come up with various generalisations of general relativity in which these requirements are relaxed. We will not get into these matters here.[41]

However, it is a curious, and occasionally calculationally or conceptually useful, fact that it is possible to relax somewhat the a priori identification of the connection with the Levi-Civita connection and nevertheless reproduce general relativity by treating the connection and metric as independent variables.

Specifically, we will consider an action of the generalised Einstein-Hilbert-like form

$$S[g_{\mu\nu}, \tilde{\Gamma}^{\lambda}_{\ \mu\nu}] = \int \sqrt{g} d^4x \ R(\tilde{\Gamma}) = \int \sqrt{g} d^4x \ g^{\mu\nu} R_{\mu\nu}(\tilde{\Gamma}) \tag{20.110}$$

for a (yet to be specified) class of connections $\tilde{\Gamma}^{\lambda}_{\ \mu\nu}$, with $g_{\mu\nu}$ and $\tilde{\Gamma}^{\lambda}_{\ \mu\nu}$ to be treated as independent variables. Since $R_{\mu\nu}(\tilde{\Gamma})$ does then not depend on the metric, the action depends purely algebraically on the metric, and on at most 1st derivatives of the connection (linearly!).

One key simplification of this kind of action is that the variation with respect to the metric is elementary (and identical to the variation of the $\sqrt{g} g^{\mu\nu}$ terms of the Einstein-Hilbert Lagrangian density $\sqrt{g} g^{\mu\nu} R_{\mu\nu}$), namely

$$\delta_g S[g_{\mu\nu}, \tilde{\Gamma}^{\lambda}_{\ \mu\nu}] = \int \sqrt{g} d^4x \ \left(R_{\mu\nu}(\tilde{\Gamma}) - \tfrac{1}{2} g_{\mu\nu} R(\tilde{\Gamma}) \right) \delta g^{\mu\nu} \tag{20.111}$$

(no integration by parts or identification of total derivative terms required). Thus, in the absence of the coupling of the metric (or of gravity) to other fields we find the equations of motion

$$G_{\mu\nu}(\tilde{\Gamma}) \equiv R_{\mu\nu}(\tilde{\Gamma}) - \tfrac{1}{2} g_{\mu\nu} R(\tilde{\Gamma}) = 0 \ . \tag{20.112}$$

These are, however, not yet the vacuum Einstein equations because the independent connection $\tilde{\Gamma}$ is not the Levi-Civita connection.

[41]Many of these generalisations (including theories with non-symmetric metrics) were originally explored by Einstein and his collaborators in their futile and (at least by the 1930s) ill-motivated attempts to find a unified field theory of gravity and Maxwell theory. For details, see e.g. the review H. Goenner, *On the History of Unified Field Theories*, Living Rev. Relativity 7 (2004) 2, `http://www.livingreviews.org/lrr-2004-2`.

It remains to look at the equations of motion imposed by stationarity of the action with respect to variations of $\tilde{\Gamma}$. It turns out (the *Palatini principle*) that

1. if one chooses the connections to be torsion-free and imposes the $\tilde{\Gamma}$-equations of motion, then the connections are forced to also be compatible with the metric and thus $\tilde{\Gamma}$ is uniquely determined to be the Levi-Civita connection

2. if one chooses the connections to be compatible with the metric and imposes the $\tilde{\Gamma}$-equations of motion, then the connections are forced to also be torsion-free and thus $\tilde{\Gamma}$ is uniquely determined to be the Levi-Civita connection

In terms of the notation introduced in section 11.5, this amounts to the assertions

$$\begin{aligned} T_{\lambda\mu\nu}(\tilde{\Gamma}) = 0 \quad \text{and} \quad \delta_{\tilde{\Gamma}} S[g,\tilde{\Gamma}] = 0 \quad &\Rightarrow \quad Q_{\lambda\mu\nu}(\tilde{\Gamma}) = 0 \quad \Rightarrow \quad \tilde{\Gamma}^{\lambda}_{\ \mu\nu} = \Gamma^{\lambda}_{\ \mu\nu} \\ Q_{\lambda\mu\nu}(\tilde{\Gamma}) = 0 \quad \text{and} \quad \delta_{\tilde{\Gamma}} S[g,\tilde{\Gamma}] = 0 \quad &\Rightarrow \quad T_{\lambda\mu\nu}(\tilde{\Gamma}) = 0 \quad \Rightarrow \quad \tilde{\Gamma}^{\lambda}_{\ \mu\nu} = \Gamma^{\lambda}_{\ \mu\nu} \end{aligned} \tag{20.113}$$

In either case, the metric equations of motion (20.112) then reduce to the vacuum Einstein equations.

In order to establish the assertions (20.113), we need two preparatory results. The first is that the generalisation of the formula (20.15),

$$\delta R_{\mu\nu}(\Gamma) = \nabla_\lambda \delta\Gamma^{\lambda}_{\ \mu\nu} - \nabla_\nu \delta\Gamma^{\lambda}_{\ \mu\lambda} \tag{20.114}$$

for the variation of the Ricci tensor in terms of the variation of the connection is

$$\delta R_{\mu\nu}(\tilde{\Gamma}) = \tilde{\nabla}_\lambda \delta\tilde{\Gamma}^{\lambda}_{\ \mu\nu} - \tilde{\nabla}_\nu \delta\tilde{\Gamma}^{\lambda}_{\ \mu\lambda} + \left(\tilde{\Gamma}^{\rho}_{\ \nu\lambda} - \tilde{\Gamma}^{\rho}_{\ \lambda\nu}\right) \delta\tilde{\Gamma}^{\lambda}_{\ \mu\rho} \quad . \tag{20.115}$$

The second is that when the connection is not the Levi-Civita connection, an expression like $\tilde{\nabla}_\lambda J^\lambda$ is not a total derivative in the integral, this being only true for the Levi-Civita connection thanks to the identity

$$\int \sqrt{g} d^4x \, \nabla_\lambda J^\lambda = \int d^4x \, \partial_\lambda(\sqrt{g} J^\lambda) \quad . \tag{20.116}$$

Writing

$$\tilde{\Gamma}^{\lambda}_{\ \mu\nu} = \Gamma^{\lambda}_{\ \mu\nu} + C^{\lambda}_{\ \mu\nu} \quad , \tag{20.117}$$

we have

$$\tilde{\nabla}_\lambda V^\mu = \nabla_\lambda V^\mu + C^{\mu}_{\ \nu\lambda} V^\nu \tag{20.118}$$

and, in particular,

$$\tilde{\nabla}_\lambda J^\lambda = \nabla_\lambda J^\lambda + C^{\lambda}_{\ \nu\lambda} J^\nu \quad , \tag{20.119}$$

only the first term on the right-hand side giving rise to a total derivative.

What we are interested in is $g^{\mu\nu}\delta R_{\mu\nu}(\tilde{\Gamma})$, and with the above results and notation we can write this as

$$\begin{aligned} g^{\mu\nu}\delta R_{\mu\nu}(\tilde{\Gamma}) = & g^{\mu\nu}\left(C^{\lambda}_{\ \rho\lambda}\delta C^{\rho}_{\ \mu\nu} - C^{\rho}_{\ \mu\lambda}\delta C^{\lambda}_{\ \rho\nu} - C^{\rho}_{\lambda\nu}\delta C^{\lambda}_{\ \mu\rho} + C^{\rho}_{\mu\nu}\delta C^{\lambda}_{\ \rho\lambda}\right) \\ & + \text{total derivative terms} \end{aligned} \tag{20.120}$$

Thus the variation of the action with respect to the $\tilde{\Gamma}$ is

$$
\begin{aligned}
\delta_{\tilde{\Gamma}} S[g,\tilde{\Gamma}] &= \int \sqrt{g} d^4x \; g^{\mu\nu} \left(C^{\lambda}_{\;\rho\lambda} \delta C^{\rho}_{\;\mu\nu} - C^{\rho}_{\;\mu\lambda} \delta C^{\lambda}_{\;\rho\nu} - C^{\rho}_{\;\lambda\nu} \delta C^{\lambda}_{\;\mu\rho} + C^{\rho}_{\;\mu\nu} \delta C^{\lambda}_{\;\rho\lambda} \right) \\
&= \int \sqrt{g} d^4x \; \left(g^{\alpha\beta} C^{\lambda\gamma}{}_{\lambda} + g^{\gamma\beta} C^{\alpha\lambda}{}_{\lambda} - C^{\beta\gamma\alpha} - C^{\alpha\beta\gamma} \right) \delta C_{\gamma\alpha\beta}
\end{aligned}
\tag{20.121}
$$

If we were to consider arbitrary $\tilde{\Gamma}$ and hence unconstrained variations $\delta C_{\gamma\alpha\beta}$, the condition for the action to be stationary with respect to variations of $\tilde{\Gamma}$ would be

$$
\delta_{\tilde{\Gamma}} S[g,\tilde{\Gamma}] = 0 \quad \Leftrightarrow \quad g^{\alpha\beta} C^{\lambda\gamma}{}_{\lambda} + g^{\gamma\beta} C^{\alpha\lambda}{}_{\lambda} - C^{\beta\gamma\alpha} - C^{\alpha\beta\gamma} = 0 \;\; . \tag{20.122}
$$

However, these equations do not determine the $C^{\alpha}_{\;\beta\gamma}$ uniquely (we will explicitly parametrise this non-uniqueness below), and hence in this case the Einstein-Hilbert-like action (20.110) alone does not give rise to acceptable equations of motion for the fields.

The situation changes if one imposes some a priori constraints on the allowed $\tilde{\Gamma}$, and hence on their variations $\delta C_{\gamma\alpha\beta}$. We now consider separately the two cases mentioned above:

1. $\tilde{\Gamma}$ are restricted to be symmetric (torsion-free)

   In terms of the coefficients $C^{\gamma}_{\;\alpha\beta}$, this amounts to the condition

$$
C^{\gamma}_{\;\alpha\beta} = C^{\gamma}_{\;\beta\alpha} \tag{20.123}
$$

   and the same condition should be imposed on their variations,

$$
\delta C^{\gamma}_{\;\alpha\beta} = \delta C^{\gamma}_{\;\beta\alpha} \tag{20.124}
$$

   Thus, symmetrising appropriately, from (20.121) one obtains the constraints

$$
2g^{\alpha\beta} C^{\lambda\gamma}{}_{\lambda} + g^{\gamma\beta} C^{\alpha\lambda}{}_{\lambda} + g^{\gamma\alpha} C^{\beta\lambda}{}_{\lambda} - 2C^{\beta\alpha\gamma} - 2C^{\alpha\beta\gamma} = 0 \;\; . \tag{20.125}
$$

   Taking traces, once by contraction with $g_{\alpha\beta}$ and once by contraction with $g_{\alpha\gamma}$ (or, equivalently, with $g_{\beta\gamma}$), one obtains two linearly independent conditions on the traces $C^{\lambda\gamma}{}_{\lambda}$ and $C^{\alpha\lambda}{}_{\lambda}$ requiring both to vanish,

$$
\text{traces} \quad \Rightarrow \quad C^{\lambda\gamma}{}_{\lambda} = 0 \quad , \quad C^{\alpha\lambda}{}_{\lambda} = 0 \;\; . \tag{20.126}
$$

   Then, upon symmetrisation, (20.125) reduces to the condition

$$
C^{\beta\alpha\gamma} + C^{\alpha\beta\gamma} = 0 \quad \Leftrightarrow \quad C_{\beta\alpha\gamma} + C_{\alpha\beta\gamma} = 0 \tag{20.127}
$$

   As we have seen in (11.103) of section 11.5, this is precisely the condition that the non-metricity tensor is zero,

$$
Q_{\alpha\beta\gamma} = C_{\beta\alpha\gamma} + C_{\alpha\beta\gamma} = 0 \;\; , \tag{20.128}
$$

i.e. that the connection is compatible with the metric.

Since we started off with a torsion-free (symmetric) connection, this means that the equations of motion fix the connection $\tilde{\Gamma}$ to be the Levi-Civita connection. Alternatively, (20.123) and (20.127) imply that $C_{\gamma\alpha\beta} = 0$ This concludes the proof of the first assertion in (20.113).

2. $\tilde{\Gamma}$ are restricted to be compatible with the metric

This is largely analogous. In terms of the coefficients $C^{\gamma}_{\alpha\beta}$, metric compatibility amounts (as just recalled) to the condition

$$C_{\gamma\alpha\beta} = -C_{\alpha\gamma\beta} \tag{20.129}$$

and the same condition should be imposed on their variations,

$$\delta C_{\gamma\alpha\beta} = -\delta C_{\alpha\gamma\beta} \tag{20.130}$$

Thus

$$C_{\gamma\alpha\beta} = C_{\gamma(\alpha\beta)} + C_{\gamma[\alpha\beta]} = T_{(\alpha\beta)\gamma} + \tfrac{1}{2}T_{\gamma\alpha\beta} \tag{20.131}$$

is the contorsion tensor. In this case, anti-symmetrisation of (20.121) leads to

$$g^{\alpha\beta}C^{\lambda\gamma}{}_{\lambda} - g^{\gamma\beta}C^{\lambda\alpha}{}_{\lambda} + C^{\beta\alpha\gamma} - C^{\beta\gamma\alpha} = 0 \ \ . \tag{20.132}$$

Taking traces, one finds $C_{\lambda\gamma}{}^{\lambda} = 0$ (the other trace $C^{\lambda}_{\lambda\gamma}$ is identically zero because of anti-symmetry), and thus (20.132) reduces to

$$C^{\beta\alpha\gamma} - C^{\beta\gamma\alpha} = 0 \quad \Leftrightarrow \quad C_{\beta\alpha\gamma} = C_{\beta\gamma\alpha} \ \ , \tag{20.133}$$

which is precisely the symmetry (no torsion) condition

$$T_{\beta\alpha\gamma} = C_{\beta\alpha\gamma} - C_{\beta\gamma\alpha} = 0 \ \ . \tag{20.134}$$

Since we started off with metric-compatible connection, this means that the equations of motion fix the connection $\tilde{\Gamma}$ to be the Levi-Civita connection. Alternatively, (20.129) and (20.133) imply that $C_{\gamma\alpha\beta} = 0$. This concludes the proof of the second assertion in (20.113).

Alternatively, and perhaps somewhat more insightfully, one can first determine the general solution to the (under-determined) equation (20.122),

$$g^{\alpha\beta}C^{\lambda\gamma}{}_{\lambda} + g^{\gamma\beta}C^{\alpha\lambda}{}_{\lambda} - C^{\beta\gamma\alpha} - C^{\alpha\beta\gamma} = 0 \ \ , \tag{20.135}$$

and then analyse the properties of the solution and the consequences of imposing some conditions on $C_{\alpha\beta\gamma}$.[42] To disentangle this equation, we proceed as in the proof of the

[42]See e.g. A. Bernal et al., *On the (non-)uniqueness of the Levi-Civita solution in the Einstein-Hilbert-Palatini formalism*, `arXiv:1606.08756 [gr-qc]` and references therein.

uniqueness of the Levi-Civita connection in section 5.4 and take sums and differences of cyclic permutations of the above equation. Then one ends up with the equation

$$2C^{\alpha\beta\gamma} = g^{\alpha\beta}(A^\gamma + B^\gamma) + g^{\beta\gamma}(A^\alpha - B^\alpha) + g^{\gamma\alpha}(B^\beta - A^\beta) \quad , \tag{20.136}$$

where

$$A^\alpha = C^{\alpha\beta}{}_\beta \quad , \quad B^\alpha = C^{\beta\alpha}{}_\beta \tag{20.137}$$

are two of the (a priori independent) traces of $C^{\alpha\beta\gamma}$. Performing either of these contractions in (20.136), one finds the condition

$$A^\alpha = B^\alpha \quad , \tag{20.138}$$

and therefore

$$C^{\alpha\beta\gamma} = g^{\alpha\beta}A^\gamma \quad \Leftrightarrow \quad C^\alpha_{\ \beta\gamma} = \delta^\alpha_\beta A_\gamma \quad . \tag{20.139}$$

Thus the general solution to the equation of motion (20.122) is

$$\tilde{\Gamma}^\alpha_{\ \beta\gamma} = \Gamma^\alpha_{\ \beta\gamma} + \delta^\alpha_\beta A_\gamma \quad , \tag{20.140}$$

with A_γ an arbitrary covector.

This family of connections has the properties

$$\tilde{\nabla}_\gamma g_{\alpha\beta} = -2A_\gamma g_{\alpha\beta} \tag{20.141}$$

and

$$\tilde{\Gamma}^\alpha_{\ \beta\gamma} - \tilde{\Gamma}^\alpha_{\ \gamma\beta} = \delta^\alpha_\beta A_\gamma - \delta^\alpha_\gamma A_\beta \quad . \tag{20.142}$$

It is now obvious that requiring either metric compatibility or the symmetry of the connection enforces $A_\gamma = 0$ and thus $\tilde{\Gamma} = \Gamma$.

REMARKS:

1. When one couples either of these theories to matter, one will find the standard Einstein equations with source the usual matter energy-momentum tensor, provided that the minimally coupled matter action depends only on the metric and not on the connection. As we have seen, this is satisfied in the case of scalar or Maxwell gauge fields (for which the minimally coupled action in the usual setting could be written in such a way that it depends only on the metric but not on the derivatives of the metric). However, typically the connection appears explicitly in the action for spinors, and in this case variation of the matter action will produce a non-zero contribution to the torsion, say. Thus in that case the Einstein-Hilbert approach (no torsion as a kinematical constraint) and the Palatini approach (torsion determined dynamically) are inequivalent.

2. Within the present framework, it is not possible to relax both the no-torsion and the metricity constraints simultaneously and to simultaneously regain them dynamically, but one can attempt to achieve this either with the aid of additional auxiliary (Lagrange multiplier) fields, or "spontaneously" by adding potentials that force the connection in the ground state to either zero torsion or zero non-metricity.[43]

3. Once one contemplates and permits the presence of non-metricity and/or torsion, there are many more terms that one can in principle use to build an action (by using scalars built from the torsion and non-metricity tensors $T_{\mu\nu\lambda}$ and $Q_{\mu\nu\lambda}$ and their covariant derivatives). Thus, unless one imposes additional symmetry requirements, say, there is no good reason to focus attention exclusively on the Einstein-Hilbert-like action (20.110), and many generalisations of general relativity suggested and discussed in the literature can and should be either rejected or amended simply on these grounds.[44]

---

[43]See e.g. R. Percacci, *Geometry of Nonlinear Field Theories* for an exploration of some of these ideas.

[44]In his review of gauge theories of gravity, F. Hehl, *Gauge Theories of Gravity and Spacetime*, `arXiv:1204.3672 [gr-qc]`, also emphasises this: "Numerous pages of printed pages could be saved if our colleagues would [...] just motivate their choice of the unknown constants." In that review it is also pointed out that what is generally referred to (and I also called and will continue to call) the Palatini formalism should properly also be attributed to Einstein (1925).

21 Hamiltonian Formulation of General Relativity

In mechanics or field theory, a common alternative to the standard Lagrangian or action-based formulation is the canonical or Hamiltonian formulation. Over the years, a lot of effort has gone into developing a framework for the canonical Hamiltonian (phase space) formulation of General Relativity. The most well known and most influential of these is the so-called ADM (Arnowitt, Deser, Misner 1959 - 1962) formalism.[45]

No other body of work on classical general relativity has single-handedly had such an impact and influence on research in this field: it awoke general relativity from its (to a large extent rather uninspiring and uninspired) "finding exact solutions" phase; it brought it to the renewed attention of a wider theoretical physics community since it provided a field theorists' analysis, perspective and understanding of the basic structures of general relativity; it presented a clean and clear 1st-order (canonical) formulation of the theory (the ADM formalism), which is crucial in understanding the Cauchy (initial value) problem in general relativity and which also provided the basis for (in)numerous subsequent attempts at a canonical quantisation of gravity; it provided groundbreaking work and insights on questions related to the notions of energy and radiation in general relativity, etc. etc.

In this section, I will sketch some aspects of the Hamiltonian or canonical formulation of general relativity, without however attempting to develop this in a completely systematic way and without being able to do justice to the depth and importance of this subject and body of work.[46]

For a concrete illustration of some of the facts and statements encountered in this section, see the discussion of the simple cosmological toy model (or "minisuperspace model" in more fancy terminology) in sections 35.8 and 35.9.

The canonical formalism has been developed in particular with an eye towards canonical quantisation of gravity and in recent years a variant of the ADM canonical variables (the Ashtekar variables) has become very popular and forms the basis of the so-called loop quantum gravity approach to quantum gravity (but I will have nothing more to say about this in these notes).[47]

In this section, we will (have to) freely make use of the results on the geometry of hypersurfaces obtained in sections 15 - 18, in particular the extrinsic geometry and

[45] This body of research is summarised in the 1962 article R. Arnowitt, S. Deser, C. Misner, *The Dynamics of General Relativity*, kindly made available on the `arXiv` as `arXiv:gr-qc/0405109`.

[46] See e.g. section 10 and Appendix E of R. Wald, *General Relativity*, or sections 3.6 and 4 of E. Poisson, *A Relativist's Toolkit* (which I have found enormously helpful in preparing this section), or sections 3 and 4 of C. Kiefer, *Quantum Gravity* (2nd edition) for modern textbook treatments of this subject.

[47] See e.g. A. Ashtekar, *Lectures on non-perturbative canonical gravity*, or C. Rovelli, *Quantum Gravity* for textbook accounts, as well as numerous review articles by these and other authors on the `arXiv`.

Gauss-Codazzi equations described in section 18.

21.1 General Covariance and Constraints

We had previously seen in sections 19.7 and 20.6 that general covariance of the Einstein equations is related to the Bianchi identities, i.e. to the existence of 4 differential relations among the 10 components of the Einstein equations. We had also seen in section 19.7 that this is reflected in the fact that the Bianchi identities imply that only six of the ten equations are truly dynamical 2nd-order differential equations while four of them constrain the initial values of the fields on some spacelike hypersurface.

We can also see this directly (i.e. without going via the Bianchi identities) from the Gauss-Codazzi equations we derived in section 18.4. We choose a foliation of the space-time by spacelike hypersurfaces Σ and choose one of them as the surface on which we will specify initial data, with the time direction pointing off (but not necessarily normal to) the hypersurface Σ. These initial data will be the spatial metric h_{ab} on Σ and something like the time-derivative of h_{ab}, i.e. something like the extrinsic curvature tensor K_{ab}. The Einstein equations should then evolve these initial data forward from Σ, i.e. they should determine the space-time metric $g_{\alpha\beta}$ in such a way that h_{ab} is the induced metric on Σ and K_{ab} its extrinsic curvature.

It turns out, however, that these initial data cannot be specified freely but are subject to some constraints. This can be immediately seen from the expressions (18.66) and (18.69) for the "time-time" and "time-space" components of the Einstein tensor we had obtained in section 18.4, namely

$$\begin{aligned} G_{NN} &\equiv G_{\alpha\beta}N^\alpha N^\beta = \tfrac{1}{2}\bar{R} + \tfrac{1}{2}(K^2 - K_{ab}K^{ab}) \\ G_{aN} &\equiv G_{\alpha\beta}E^\alpha_a N^\beta = \bar{\nabla}^b K_{ba} - \bar{\nabla}_a K^b_b \end{aligned} \tag{21.1}$$

These just depend on the values of h_{ab} and K_{ab} on Σ, and therefore these components of the Einstein equations are not evolution equations at all but rather provide 4 constraints among the initial data. These constraints

$$\begin{aligned} \bar{R} + K^2 - K_{ab}K^{ab} &= 16\pi G_N\, T_{\alpha\beta}N^\alpha N^\beta \\ \bar{\nabla}^b K_{ba} - \bar{\nabla}_a K^b_b &= 8\pi G_N\, T_{\alpha\beta}E^\alpha_a N^\beta \end{aligned} \tag{21.2}$$

(on Σ) reflect the underlying general covariance of the Einstein equations.

The remaining (space-space) components of the Einstein tensor depend on the 2nd time derivatives of the metric, i.e. on the time-derivatives of K_{ab}, and therefore the remaining 6 space-space components

$$G_{ab} = 8\pi G_N\, T_{ab} \tag{21.3}$$

of the Einstein equations are true evolution equations for h_{ab}. Due to their non-linearity, and due to the presence of the constraints, these equations are highly non-trivial and mathematically extremely challenging.

In a canonical (Hamiltonian, first-order) formulation of the problem, the first step is a 3+1 dimensional decomposition of the space-time into "space" (a hypersurface or family of spacelike hypersurfaces Σ) and "time", and a corresponding decomposition of the dynamical variables. Among the dynamical variables one would then have the "spatial" metric h_{ab} on Σ, and phase space variables and initial data would then include the configuration variable h_{ab} and its canonically conjugate momentum π^{ab}.

In the ADM formalism, a more detailed analysis, starting from the Lagrangian formulation of the theory and then using the Gauss-Codazzi expression (18.73) for the Ricci scalar (Einstein-Hilbert Lagrangian) R shows that, more specifically, the canonically conjugate variables are h_{ab} and

$$\pi^{ab} = \sqrt{h}(K^{ab} - Kh^{ab}) \tag{21.4}$$

(see (21.70) in section 21.6). Since π^{ab} can be expressed in terms of h_{ab} and K_{ab} (and conversely K_{ab} can be expressed in terms of h_{ab} and π^{ab}), initial data can also be specified by specifying h_{ab} and π^{ab} on Σ (so these variables span the phase space of the theory).

Of course, these variables need to satisfy the constraints. In a Hamiltonian formulation these constraints are known as the *Hamiltonian constraint* and the *Momentum constraints* respectively. Presence of such constraints in the Hamiltonian formulation is a characteristic feature of gauge theories and/or generally covariant theories, and we will see below how precisely they arise from a canonical formulation of the theory, and what their significance is from this perspective. Roughly speaking, they turn out to generate the time evolution and the action of spatial coordinate transformations on the fields via Poisson brackets. This is the way 4-dimensional general covariance is implemented in a foliation-dependent way in the (foliation-dependent) 3+1 dimensional Hamiltonian formulation of the theory.

21.2 Gauss-Codazzi Action and the Gibbons-Hawking-York Boundary Term

As we saw above, the Gauss-Codazzi decomposition of the curvature tensor already provides a reasonably clean separation of the Einstein equations into constraints and true dynamical evolution equations. It is therefore also natural to take the corresponding decomposition (18.73) of the Ricci scalar, i.e. the of the Einstein-Hilbert Lagrangian,

$$R = \bar{R} + \epsilon(K^2 - K^{\alpha\beta}K_{\alpha\beta}) + 2\epsilon\nabla_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) \ , \tag{21.5}$$

as the starting point of a canonical Hamiltonian analysis of the theory.

In order to be able to use this, let us assume that we do not just have a single hypersurface Σ but a foliation of the space-time into such hypersurfaces. We thus assume that the space-time is of the form $\Sigma \times \mathbb{R}$, with $\mathbb{R}$ representing the time-direction, and Σ are the constant time slices equipped with some fixed spatial coordinates y^a.

On each of these slices of constant time the scalar curvature takes the above form (21.5). The first term $\bar{R}$ only depends on the intrinsic geometry of Σ (and thus contains no normal derivatives), while the extrinsic curvature term contains squares of terms with first normal derivatives but no second normal derivatives. These second normal derivatives can then only appear in the third term, which is a total derivative. Thus this decomposition is reminiscent of, and serves the same purpose as, say the addition of the Gibbons-Hawking-York boundary term (20.61) to the Einstein-Hilbert action discussed in section 20.5.

Indeed, if the boundary consists of one (or two, initial and final, say) of these spacelike hypersurfaces Σ, this already leads to an appropriate decomposition of the Einstein-Hilbert action, namely the Gauss-Codazzi form of the action

$$\begin{aligned} S_{GC}[g_{\alpha\beta}] &= \int \sqrt{g}d^4x\ (\bar{R} + \epsilon(K^2 - K^{\alpha\beta}K_{\alpha\beta})) \\ &= S_{EH}[g_{\alpha\beta}] - 2\epsilon \oint_\Sigma d\sigma_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) \\ &= S_{EH}[g_{\alpha\beta}] - 2\oint_\Sigma \sqrt{h}d^3y\ N_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta)\ \ . \end{aligned} \tag{21.6}$$

As we will now see, for spacelike boundaries addition of this total derivative term is equivalent to the addition of the Gibbons-Hawking-York term. Indeed, looking at the boundary term more closely, we see that, as a consequence of $N^\alpha\nabla_\beta N_\alpha = \nabla_\beta(N^\alpha N_\alpha)/2 = 0$, it reduces to

$$-2\oint \sqrt{h}d^3y\ N_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) = 2\epsilon\oint \sqrt{h}d^3y\ \nabla_\beta N^\beta\ \ , \tag{21.7}$$

which is precisely the Gibbons-Hawking-York term (20.61), (20.62).

With respect to such a foliation and boundary, the Gauss-Codazzi form of the action is therefore identical to the standard gravitational action (20.67),

$$S_{GC}[g_{\alpha\beta}] = S_g[g_{\alpha\beta}] = S_{EH}[g_{\alpha\beta}] + S_{GHY}[g_{\alpha\beta}]\ \ . \tag{21.8}$$

<u>Remarks:</u>

1. Thus another way to motivate (or arrive at) the Gibbons-Hawking-York boundary term is to start from the decomposition (18.73) of the Ricci scalar provided by the Gauss-Codazzi equations.

2. Once expressed in terms of the so-called ADM variables - see section 21.4 below - I will refer to the form (21.6) of the action as the ADM action.

3. If in addition there are timelike (asymptotic) boundaries $\mathcal{B}$, then additional boundary terms are required, because the contribution of such a boundary to the boundary term in the action, schematically something like

$$-2\oint_{\mathcal{B}} \sqrt{h_{\mathcal{B}}}d^3y\ r_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta)$$

(with r^α the normal to $\mathcal{B}$ and $h_\mathcal{B}$ the absolute value of the determinant of the (Lorentzian signature) metric induced on $\mathcal{B}$), will not equal the standard Gibbons-Hawking-York boundary term for this boundary (the integral over $\mathcal{B}$ of the trace of the extrinsic curvature of $\mathcal{B}$). In the following we will, until further notice, assume that there is no such boundary component $\mathcal{B}$, and will then return to this issue in sections 21.10 and 21.11.

21.3 ADM Decomposition of the Metric (ADM Variables)

The next step is to find a parametrisation of the space-time metric adapted to a given choice of foliation of the space-time by (constant time) hypersurfaces. In order to achieve this, we first assume that the spatial hypersurfaces of this foliation of the space-time are hypersurfaces of "constant time", i.e. they are the level sets of some *time function* $t(x^\alpha)$,

$$\Sigma_{t_0} = \{x^\alpha :\ t(x^\alpha) = t_0\} \quad , \tag{21.9}$$

with timelike (future-oriented) normal vector N^α, $N_\alpha \sim \partial_\alpha t$.

We can now introduce coordinates (t, y^a) on the space-time via a coordinate transformation

$$x^\alpha = x^\alpha(t, y^a) \tag{21.10}$$

in the following way:

- We stipulate that for any fixed value $t = t_0$,

  $$x^\alpha_{t_0}(y^a) := x^\alpha(t_0, y^a) \tag{21.11}$$

  gives us the embedding (cf. (15.4) and the discussion in sections 15.1 and 15.2) of a hypersurface Σ (with coordinates y^a) as the hypersurface Σ_{t_0} in space-time,

  $$x_{t_0} :\ \Sigma \to \Sigma_{t_0} \subset M \quad . \tag{21.12}$$

- The curves

  $$x^\alpha_{y_0}(t) := x^\alpha(t, y^a_0) \tag{21.13}$$

  then connect points on different hypersurfaces with the same values of the spatial coordinates $y^a = y^a_0$, and thus provide us with a notion (or encode a choice) of "time evolution" from one hypersurface to the next.

Given $x^\alpha = x^\alpha(t, y^a)$ for some choice of foliation and time-evolution,

$$E^\alpha_a = \left(\frac{\partial x^\alpha}{\partial y^a}\right)_t \tag{21.14}$$

gives us the tangent vectors (15.16) to the surfaces Σ_t, while

$$(\partial_t)^\alpha = \left(\frac{\partial x^\alpha}{\partial t}\right)_y \tag{21.15}$$

gives us the components of the time-evolution vector field ∂_t. The curves (21.13) are not required to be normal to the hypersurface. In general, therefore, ∂_t can be decomposed into a normal and tangential part as

$$(\partial_t)^\alpha = N N^\alpha + E_a^\alpha \mathcal{N}^a \quad . \tag{21.16}$$

The function N and spatial vector field $\mathcal{N}^a$ appearing in this expression are known as the *lapse function* and *shift vector field* respectively. They parametrise the freedom in the choice of the time-evolution vector.

We thus have

$$\begin{aligned} dx^\alpha &= (N N^\alpha + E_a^\alpha \mathcal{N}^a)dt + E_a^\alpha dy^a \\ &= N N^\alpha dt + E_a^\alpha (dy^a + \mathcal{N}^a dt) \quad . \end{aligned} \tag{21.17}$$

Plugging this into the line element for the space-time metric and using $g_{\alpha\beta}N^\alpha N^\beta = -1$, one finds

$$ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta = -N^2 dt^2 + h_{ab}(dy^a + \mathcal{N}^a dt)(dy^b + \mathcal{N}^b dt) \quad , \tag{21.18}$$

where

$$h_{ab} = g_{\alpha\beta} E_a^\alpha E_b^\beta \tag{21.19}$$

is the induced metric. This is the so-called *ADM decomposition* of the metric, and is the usual point of departure for developing the Hamiltonian formulation of general relativity (and of field theories in a gravitational background).

The following facts are easy to establish:

1. The components of the metric and its inverse are explicitly

   $$g_{tt} = -N^2 + h_{ab}\mathcal{N}^a\mathcal{N}^b \quad , \quad g_{at} = h_{ab}\mathcal{N}^b \equiv \mathcal{N}_a \quad , \quad g_{ab} = h_{ab} \tag{21.20}$$

   and

   $$g^{tt} = -N^{-2} \quad , \quad g^{at} = N^{-2}\mathcal{N}^a \quad , \quad g^{ab} = h^{ab} - N^{-2}\mathcal{N}^a\mathcal{N}^b \quad . \tag{21.21}$$

2. The normalised timelike normal vector field to the surfaces of constant t, thus $N_\alpha \sim \partial_\alpha t$ is given by

   $$N_\alpha = -N\partial_\alpha t \quad . \tag{21.22}$$

3. Thus in the ADM coordinates (t, y^a) one has

   $$N_t = -N \quad , \quad N_a = 0 \tag{21.23}$$

   as well as

   $$E_a^t = 0 \quad , \quad E_a^b = \delta_a^b \quad . \tag{21.24}$$

4. In terms of these variables, the 4-dimensional volume element $\sqrt{g}$ takes the simple factorised form
$$\sqrt{g} = N\sqrt{h} \quad . \tag{21.25}$$

5. Moreover, in terms of these variables the extrinsic curvature tensor of the surfaces of constant t can be written as
$$K_{ab} = \frac{1}{2N}(\dot{h}_{ab} - L_{\mathcal{N}}h_{ab}) \tag{21.26}$$
where
$$\dot{h}_{ab} = L_{\partial_t}h_{ab} = \partial_t h_{ab} \tag{21.27}$$
is the time (Lie) derivative of h_{ab}, and in terms of the intrinsic = induced covariant derivative $\bar{\nabla}$ the 2nd term can be written as
$$L_{\mathcal{N}}h_{ab} = \bar{\nabla}_a\mathcal{N}_b + \bar{\nabla}_b\mathcal{N}_a \quad . \tag{21.28}$$

It is perhaps only the derivation of the last result (21.26) that requires some comment. Here is a sketch of 2 derivations:

5a Start with the definition (18.27),
$$K_{ab} = E_a^\alpha E_b^\beta \nabla_\alpha N_\beta \tag{21.29}$$
and use (21.22) and $E^t_\alpha = 0$ (21.24) to write this as
$$K_{ab} = -NE_a^\alpha E_b^\beta \nabla_\alpha\partial_\beta t = NE_a^\alpha E_b^\beta \Gamma^t{}_{\alpha\beta} \quad . \tag{21.30}$$
Now use the explicit expressions for the components of the metric and its inverse to write this as
$$K_{ab} = -N^{-1}\Gamma_{tab} + N^{-1}\mathcal{N}^c\Gamma_{cab} \quad . \tag{21.31}$$
Noting that
$$2\Gamma_{tab} = \partial_a\mathcal{N}_b + \partial_b\mathcal{N}_a - \partial_t h_{ab} \quad , \quad \Gamma_{cab} = \bar{\Gamma}_{cab} \tag{21.32}$$
this leads directly to (21.26).

5b Alternatively, start with
$$\dot{h}_{ab} = L_{\partial_t}h_{ab} = L_{\partial_t}(g_{\alpha\beta}E_a^\alpha E_b^\beta) \tag{21.33}$$
and use
$$L_{\partial_t}E^a_\alpha = [\partial_t, \partial_{y^a}]^\alpha = 0 \tag{21.34}$$
to write this as
$$\dot{h}_{ab} = (L_{\partial_t}g_{\alpha\beta})E_a^\alpha E_b^\beta = (\nabla_\alpha(\partial_t)_\beta + \nabla_\beta(\partial_t)_\alpha)E_a^\alpha E_b^\beta \quad . \tag{21.35}$$

Now use (21.16) in the form

$$(\partial_t)_\beta = N N_\beta + \mathcal{N}_\beta \tag{21.36}$$

with

$$\mathcal{N}_\beta = g_{\beta\alpha}\mathcal{N}^\alpha = g_{\beta\alpha}E^\alpha_a\mathcal{N}^a \tag{21.37}$$

and $N_\alpha E^\alpha_a = 0$ and (21.29) to deduce

$$\dot{h}_{ab} = 2NK_{ab} + \bar{\nabla}_a\mathcal{N}_b + \bar{\nabla}_b\mathcal{N}_a \;\; , \tag{21.38}$$

which is equivalent to (21.26).

21.4 ADM Action and the DeWitt Metric

With these preliminaries out of the way, let us now turn our attention to the gravitational action. The starting point is the Gauss-Codazzi action (21.6), but now of course viewed as a functional of the ADM variables $(h_{ab}, N, \mathcal{N}^a)$. Since the extrinsic curvature tensor $K_{\alpha\beta}$ is a spatial tensor, one has

$$K = g^{\alpha\beta}K_{\alpha\beta} = h^{ab}K_{ab} \tag{21.39}$$

and

$$K^{\alpha\beta}K_{\alpha\beta} = K^{ab}K_{ab} \tag{21.40}$$

where

$$K_{ab} = E^\alpha_a E^\beta_b K_{\alpha\beta} \;\; . \tag{21.41}$$

Thus we can write the action (21.6) as

$$S_{ADM}[h_{ab}, N, \mathcal{N}_k] = \int dt\; d^3x\; \sqrt{h}N(\bar{R} + K^{ab}K_{ab} - K^2) \tag{21.42}$$

This is what I will refer to as the (2nd order form of the) ADM action. We can write this as an integral of a Lagrangian L_{ADM} or a Lagrangian density $\mathcal{L}_{ADM}$ as

$$S_{ADM}[h_{ab}, N, \mathcal{N}_k] = \int dt\; L_{ADM} = \int dt\; d^3x\; \mathcal{L}_{ADM} \;\; . \tag{21.43}$$

Note that in terms of the so-called *DeWitt metric*

$$G^{abcd} = \tfrac{1}{2}(h^{ac}h^{bd} + h^{ad}h^{bc} - 2h^{ab}h^{cd}) \tag{21.44}$$

the "kinetic" (extrinsic curvature) term

$$K_{\alpha\beta}K^{\alpha\beta} - K^2 = K_{ab}K^{ab} - K^2 \tag{21.45}$$

can be written as

$$K_{ab}K^{ab} - K^2 = G^{abcd}K_{ab}K_{cd} \;\; . \tag{21.46}$$

Thus the Lagrangian density

$$\mathcal{L}_{ADM} = \sqrt{h}N(G^{abcd}K_{ab}K_{cd} + \bar{R}) \quad . \tag{21.47}$$

now has the standard "kinetic minus potential energy" form.

Remarks:

1. As this DeWitt metric determines the form of the kinetic term, it also plays the role of a natural metric on the space of spatial metrics or, better, metric deformations δh_{ab}, in the sense that one can define

$$\langle \delta_1 h, \delta_2 h \rangle = \int_\Sigma \sqrt{h} d^3x \; G^{abcd}(\delta_1 h_{ab})(\delta_2 h_{cd}) \quad . \tag{21.48}$$

   This metric is not positive definite. The "negative" direction in the space of deformations of a spatial metric h_{ab} turns out to be associated with overall volume deformations.

2. This can be seen very explicitly in the case of simple cosmological models, where this overall scaling of the spatial metric is the only degree of freedom (the *cosmic scale factor*) and thus the gravitational kinetic contribution to the action is strictly non-positive (see the discussion in section 35.8 for an explicit illustration of this fact).

21.5 Synopsis of the Canonical Formulation of Maxwell Theory

At this point, for comparison purposes it will be useful to have some at least very superficial familiarity with the canonical formulation of Maxwell theory. I will therefore briefly summarise this here (more sophisticated treatments of this standard subject can be found in many places).

We start with the Lorentz-invariant Lagrangian (density) of Maxwell theory,

$$\mathcal{L} = -\tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta} \tag{21.49}$$

but break manifest Lorentz invariance by choosing a particular inertial frame with coordinates $(x^0 = t, x^a)$, and a corresponding slicing of space-time by constant time hypersurfaces. Then the Lagrangian takes the form

$$\mathcal{L} = \frac{1}{2}(\vec{E}^2 - \vec{B}^2) \tag{21.50}$$

where

$$F_{0a} = \partial_0 A_a - \partial_a A_0 = -E_a \quad , \quad F_{ab} = \epsilon_{abc}B_c \quad . \tag{21.51}$$

Now we proceed as follows:

- The canonical momenta conjugate to the fields (A_0, A_a) are

$$\Pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = 0 \quad , \quad \Pi^a = \frac{\partial \mathcal{L}}{\partial \dot{A}_a} = -E^a \quad . \tag{21.52}$$

  The latter allows us to express the "velocities" $\dot{A}_a$ in terms of the momenta,

$$\dot{A}_a = \Pi_a + \partial_a A_0 \quad , \tag{21.53}$$

  The former, on the other hand, is a constraint, known as a *primary constraint*, that arises because the action does not depend on $\dot{A}_0$.

- The Legendre transform of the Lagrangian density is thus the Hamiltonian density

$$\begin{aligned} \mathcal{H} &= \Pi^a \dot{A}_a - \mathcal{L} = \Pi^a \Pi_a + \Pi^a \partial_a A_0 - \mathcal{L} \\ &= \tfrac{1}{2}(\vec{\Pi}^2 + \vec{B}^2) - A_0(\partial_a \Pi^a) + \partial_a(A_0 \Pi^a) \quad . \end{aligned} \tag{21.54}$$

  In constructing the Hamiltonian

$$H = \int d^3x\, \mathcal{H} \quad , \tag{21.55}$$

  with suitable boundary conditions we can ignore the total derivative term. Thus we can work instead with the Hamiltonian density

$$\begin{aligned} \mathcal{H} &= \tfrac{1}{2}(\vec{\Pi}^2 + \vec{B}^2) - A_0(\partial_a \Pi^a) \\ &= \tfrac{1}{2}(\vec{E}^2 + \vec{B}^2) + A_0(\partial_a E^a) \quad . \end{aligned} \tag{21.56}$$

- (A_a, E^a) are standard canonically conjugate variables satisfying the canonical Poisson bracket relations

$$\{A_a(\vec{x}), E^b(\vec{y})\} = -\delta_a^b\, \delta^{(3)}(\vec{x}, \vec{y}) \quad . \tag{21.57}$$

- Hamilton's equations of motion

$$\begin{aligned} \dot{A}_a &= \{A_a, H\} = \Pi_a + \partial_a A_0 \\ \dot{\Pi}^a &= \{\Pi^a, H\} \end{aligned} \tag{21.58}$$

  reproduce the relation (21.53) between the velocities and momenta, and the spatial components of the Maxwell equations $\partial_\alpha F^{\alpha\beta} = 0$,

$$\dot{\Pi}^a = \{\Pi^a, H\} \quad \Leftrightarrow \quad \partial_\alpha F^{\alpha a} = 0 \quad . \tag{21.59}$$

- A_0 acts as a Lagrange multiplier, imposing the time-component of the Maxwell equations,

$$\partial_a E^a = 0 \quad \Leftrightarrow \quad \partial_\alpha F^{\alpha 0} = 0 \quad . \tag{21.60}$$

Alternatively, this equation arises from requiring that the primary constraint $\Pi^0 = 0$ be preserved under time-evolution,

$$\dot{\Pi}^0 = \{\Pi^0, H\} \stackrel{!}{=} 0 \quad \Rightarrow \quad \mathcal{G} = \partial_a E^a = 0 \ . \tag{21.61}$$

This condition does not contain time-derivatives of the canonical variables, and therefore it is not an evolution equation for the phase space variables but rather a *secondary constraint* on the initial data on a fixed time hypersurface, the *Gauss Law Constraint.*

The name derives from the fact that in the presence of matter one would instead have

$$\partial_a E^a = \rho \ , \tag{21.62}$$

with ρ the charge density, and this relation allows one to express the total charge contained in a spatial volume as a surface integral (a statement usually known as the Gauss Law).

- This Gauss law constraint reflects the underlying gauge invariance of Maxwell theory. In particular, via Poisson brackets it generates the action of the gauge transformations on A_a (and E^a),

$$\begin{aligned} \delta_\Psi A_a(\vec{x}) &= \{A_a(\vec{x}), \int d^3y\ \Psi(\vec{y})\partial_b E^b(\vec{y})\} = \partial_a \Psi(\vec{x}) \\ \delta_\Psi E^a(\vec{x}) &= \{E^a(\vec{x}), \int d^3y\ \Psi(\vec{y})\partial_b E^b(\vec{y})\} = 0 \ . \end{aligned} \tag{21.63}$$

  Thus the physical (reduced) phase space of the system consists of the pairs (A_a, E^a) satisfying the Gauss law constraint, modulo gauge transformations. Gauge invariant observables are those functions on phase space that have vanishing Poisson brackets with the Gauss law constraint.

- Since these smeared Gauss law constraints

$$\mathcal{G}[\Psi] = \int d^3y\ \Psi(\vec{y})\partial_b E^b(\vec{y}) \tag{21.64}$$

  depend only on the electric field, not on the vector potential, they satisfy the constraint algebra

$$\{\mathcal{G}[\Psi_1], \mathcal{G}[\Psi_2]\} = 0 \tag{21.65}$$

  which reflects the Abelian $U(1)$ gauge invariance of Maxwell theory (whereas the corresponding Gauss law generators in a non-Abelian gauge theory would have formed a Poisson bracket realisation of the gauge algebra).

- Finally we note that the on-shell value of the Hamiltonian gives the energy (density) of a solution,

$$\begin{aligned} \mathcal{G} = 0 \quad \Rightarrow \quad \mathcal{H} &= \tfrac{1}{2}(\vec{E}^2 + \vec{B}^2) = T_{00} \\ H &= \int d^3x\ T_{00} = E \ . \end{aligned} \tag{21.66}$$

In the following, you should see that there is a close analogy betweeen the Gauss Law constraint $\mathcal{G}$ (and its associated Lagrange multiplier A_0) of Maxwell theory, and the so-called *Momentum Constraint* $\mathcal{H}_a$ (and its associated Lagrange multiplier, the shift vector $\mathcal{N}^a$) on the gravity side (but I will refrain from constantly pointing out these analogies in the following, as that can become rather obnoxious).

On the other hand, there is no good Maxwell analogue of the so-called *Hamiltonian Constraint*, whose presence is instead a characteristic feature of general relativity (and other parametrisation invariant theories).

21.6 Back to Gravity: Conjugate Momenta and Primary Constraints

Let us now return to gravity, in particular to the ADM Lagrangian (21.47)

$$\mathcal{L}_{ADM} = \sqrt{h}N(G^{abcd}K_{ab}K_{cd} + \bar{R}) \quad . \tag{21.67}$$

The (genuine, unconstrained) conjugate momenta to h_{ab} are (the definition adopted here is that the canonical momenta are tensor densities)

$$\pi^{ab} = \frac{\partial \mathcal{L}_{ADM}}{\partial \dot{h}_{ab}} \quad . \tag{21.68}$$

From (21.67) and (21.26), one sees that

$$\mathcal{L}_{ADM} = \sqrt{h}NG^{abcd}(\dot{h}_{ab}/2N + \ldots)K_{cd} + \ldots \quad . \tag{21.69}$$

so that explicitly the conjugate momenta are

$$\pi^{ab} = \sqrt{h}G^{abcd}K_{cd} = \sqrt{h}(K^{ab} - h^{ab}K) \quad , \tag{21.70}$$

as anticipated in (21.4).

By taking the trace, one finds that

$$\pi \equiv h_{ab}\pi^{ab} = -2\sqrt{h}K \tag{21.71}$$

so that one can invert (21.70) and express K_{ab} in terms of π^{ab} as

$$\sqrt{h}K_{ab} = \pi_{ab} - \tfrac{1}{2}h_{ab}\pi \quad . \tag{21.72}$$

This can also be written as

$$\sqrt{h}K_{ab} = G_{abcd}\pi^{cd} \quad , \tag{21.73}$$

where

$$G_{abcd} = \tfrac{1}{2}(h_{ac}h_{bd} + h_{ad}h_{bc} - h_{ab}h_{cd}) \tag{21.74}$$

is indeed the inverse of the DeWitt metric (21.44) in the sense of a metric on symmetric 2-tensors,

$$G^{abcd}G_{cdef} = \tfrac{1}{2}(\delta^a_e\delta^b_f + \delta^a_f\delta^b_e) \quad . \tag{21.75}$$

Using (21.26) in the form

$$\dot{h}_{ab} = 2NK_{ab} + L_{\mathcal{N}}h_{ab} \quad , \tag{21.76}$$

one sees that the "velocities" $\dot{h}_{ab}$ can be written in terms of the coordinates and momenta as

$$\dot{h}_{ab} = \frac{2N}{\sqrt{h}}G_{abcd}\pi^{cd} + L_{\mathcal{N}}h_{ab} \quad . \tag{21.77}$$

Turning now to the other variables N and $\mathcal{N}^a$, note that the action does not depend on their time-derivatives at all since the intrinsic scalar curvature $\bar{R}$ is completely independent of these variables while the extrinsic curvature involves only N and the spatial covariant derivatives of $\mathcal{N}^a$. Thus the conjugate momenta to these variables are zero,

$$p_N = \frac{\partial \mathcal{L}_{ADM}}{\partial \dot{N}} = 0 \quad , \quad p_{\mathcal{N}^a} = \frac{\partial \mathcal{L}_{ADM}}{\partial \dot{\mathcal{N}}^a} = 0 \quad . \tag{21.78}$$

Since the action does not depend on the time-derivatives of these variables, they act as Lagrange multipliers and variation of the action with respect to the lapse function and shift vector gives rise to the *Hamiltonian constraint* $\mathcal{H} = 0$ and the *Momentum Constraints* $\mathcal{H}_a = 0$ already mentioned in section 21.1.

- Variation of the lapse function N

  Variation of the lapse function N leads to

$$\mathcal{H} = \sqrt{h}(K_{ab}K^{ab} - K^2 - \bar{R}) = \sqrt{h}(G^{abcd}K_{ab}K_{cd} - \bar{R}) = 0 \tag{21.79}$$

  (it is convenient to define the constraints as tensor densities; this accounts for the factor of $\sqrt{h}$). Comparison with (21.1) shows that

$$\mathcal{H} = -2\sqrt{h}G_{NN} \tag{21.80}$$

  so vanishing of this constraint is precisely this component of the Einstein equations.

  Note the relative sign flip between the "kinetic" (extrinsic curvature) and "potential" (intrinsic scalar curvature) terms between $\mathcal{L}_{ADM}$ (21.47) and $\mathcal{H}$. This arises because due to the N^{-1} in the expression (21.26) for the extrinsic curvature

$$\sqrt{h}NG^{abcd}K_{ab}K_{cd} \sim N^{-1} \quad , \quad \sqrt{h}N\bar{R} \sim N \quad . \tag{21.81}$$

  Thus variation of the action with respect to N (i.e. differentiation of the Lagrangian density with respect to N in the case at hand) simply changes the relative sign of the 2 terms, giving rise to the Hamiltonian constraint $\mathcal{H} = 0$. This Hamiltonian constraint will indeed turn out to be part of the Hamiltonian of the theory. In this sense, variation with respect to N implements the Legendre transformation.

- Variation of the shift vector $\mathcal{N}_a$

  The Momentum constraint arises from the variation of the shift vector $\mathcal{N}_a$ in

$$\mathcal{L}_{ADM} = \sqrt{h}NG^{abcd}(-N^{-1}\bar{\nabla}_a\mathcal{N}_b \pm \ldots)K_{cd} + \ldots \tag{21.82}$$

  (and a (spatial) integration by parts). Thus the Momentum constraint is

$$\mathcal{H}^b = -2\sqrt{h}\bar{\nabla}_a(K^{ab} - h^{ab}K) = -2\sqrt{h}\bar{\nabla}_a(G^{abcd}K_{cd}) = 0 \quad . \tag{21.83}$$

  Again comparison with (21.1) shows that

$$\mathcal{H}_a = -2\sqrt{h}G_{Na} \quad , \tag{21.84}$$

  so that the Momentum constraint impose these components of the Einstein equations. Written in terms of the canonical momenta (21.70), the Momentum constraint is simply

$$\mathcal{H}^b = -2\bar{\nabla}_a\pi^{ab} \quad . \tag{21.85}$$

  Note that this is a tensor density because of the $\sqrt{h}$ in the definition (21.70) of π^{ab}.

21.7 Legendre Transform and ADM Hamiltonian

One can now pass to a Hamiltonian formulation in the standard manner, by

- performing the Legendre transformation,
- expressing the velocities in terms of the momenta,
- and thinking about how the constraints are to be implemented.

We start with the Legendre transform

$$\mathcal{H}_{ADM} = \pi^{ab}\dot{h}_{ab} - \mathcal{L}_{ADM} \tag{21.86}$$

(because of (21.78), whether or not we also formally include $p_N\dot{N}$ etc. in this expression makes no difference). Now from (21.26) $\pi^{ab}\dot{h}_{ab}$ consists of 2 kinds of terms, namely

$$\pi^{ab}\dot{h}_{ab} = \pi^{ab}(2NK_{ab} + 2\bar{\nabla}_a\mathcal{N}_b) \tag{21.87}$$

The first of these, combined with the kinetic term, gives

$$2N\pi^{ab}K_{ab} - N\sqrt{h}G^{abcd}K_{ab}K_{cd} = N\sqrt{h}G^{abcd}K_{ab}K_{cd} \quad . \tag{21.88}$$

To write this in terms of the momenta π^{ab}, note that

$$\pi \equiv h_{ab}\pi^{ab} = -2\sqrt{h}K \quad , \quad \pi^{ab}\pi_{ab} = h(K^{ab}K_{ab} + K^2) \quad , \tag{21.89}$$

so that

$$\sqrt{h}(K^{ab}K_{ab} - K^2) = (\pi^{ab}\pi_{ab} - \tfrac{1}{2}\pi^2)/\sqrt{h} \ . \tag{21.90}$$

We can also write this in terms of the inverse DeWitt metric (21.74) as

$$\sqrt{h}G^{abcd}K_{ab}K_{cd} = G_{abcd}\pi^{ab}\pi^{cd}/\sqrt{h} \ . \tag{21.91}$$

Taking into account the scalar curvature term in the Lagrangian, we thus arrive at

$$2N\pi^{ab}K_{ab} - \mathcal{L}_{ADM} = N\sqrt{h}(G_{abcd}\pi^{ab}\pi^{cd}/h - \bar{R}) \tag{21.92}$$

This is precisely (N times the) the Hamiltonian constraint (21.79), now expressed in terms of the canonical variables (h_{ab}, π^{ab}),

$$2N\pi^{ab}K_{ab} - \mathcal{L}_{ADM} = N\mathcal{H} \ , \tag{21.93}$$

with

$$\begin{aligned}\mathcal{H} &= \sqrt{h}(G_{abcd}\pi^{ab}\pi^{cd}/h - \bar{R}) \\ &= (\pi^{ab}\pi_{ab} - \tfrac{1}{2}\pi^2)/\sqrt{h} - \sqrt{h}\bar{R} \ .\end{aligned} \tag{21.94}$$

The other contribution from (21.87) is up to an integration by part simply equal to

$$2\pi^{ab}\bar{\nabla}_a\mathcal{N}_b \quad \Rightarrow \quad -2(\bar{\nabla}_a\pi^{ab})\mathcal{N}_b = \mathcal{N}_b\mathcal{H}^b \tag{21.95}$$

Thus the Hamiltonian density has the striking form

$$\mathcal{H}_{ADM} = N\mathcal{H} + \mathcal{N}^a\mathcal{H}_a \tag{21.96}$$

and the Hamiltonian is

$$H_{ADM} = \int d^3x \ \mathcal{H}_{ADM} = \int d^3x \ (N\mathcal{H} + \mathcal{N}^a\mathcal{H}_a) \ . \tag{21.97}$$

We can also use these results to write the ADM action in 1st-order form as

$$S_{ADM} = \int dt \ d^3x \ (\pi^{ab}\dot{h}_{ab} - N\mathcal{H} - \mathcal{N}^a\mathcal{H}_a) \ . \tag{21.98}$$

Remarks:

1. As anticipated in the previous section, the Hamiltonian constraint (21.94) looks exactly like a standard Legendre transform of the Lagrangian (21.47)

$$\begin{aligned}\mathcal{L}_{ADM} &= N\sqrt{h}(G^{abcd}K_{ab}K_{cd} + \bar{R}) \\ \longrightarrow \quad N\mathcal{H} &= N\sqrt{h}(G_{abcd}\pi^{ab}\pi^{cd}/h - \bar{R})\end{aligned} \tag{21.99}$$

2. If one starts with the correctly normalised action

$$S_{EH} = \frac{1}{16\pi G_N}\int \sqrt{g}d^4x \ R \ , \tag{21.100}$$

then what is only an overall normalisation factor of the action manifests itself as a relative factor between the kinetic and potential terms of the Hamiltonian constraint. Indeed, with this normalisation the momenta π^{ab} (21.70) carry an additional factor of $1/16\pi G_N$,

$$\pi^{ab} = \frac{1}{16\pi G_N}\sqrt{h}G^{abcd}K_{cd} = \frac{\sqrt{h}}{16\pi G_N}(K^{ab} - h^{ab}K) \ . \tag{21.101}$$

It is then common to also rescale the constraints by a factor of $(16\pi G_N)^{-1}$, so that the (rescaled) Hamiltonian constraint is

$$\mathcal{H} = \frac{(16\pi G_N)}{\sqrt{h}}G_{abcd}\pi^{ab}\pi^{cd} - \frac{\sqrt{h}}{(16\pi G_N)}\bar{R} \ , \tag{21.102}$$

while the rescaled Momentum constraint continues to take the form $\mathcal{H}_a = -\bar{\nabla}_a\pi^{ab}$ with respect to the rescaled momenta.

3. With this rescaling, there is then no explicit factor of $(16\pi G_N)$ in the ADM Hamiltonian, i.e. one continues to have (21.97)

$$H_{ADM} = \int d^3x \ \mathcal{H}_{ADM} = \int d^3x \ (N\mathcal{H} + \mathcal{N}^a\mathcal{H}_a) \ . \tag{21.103}$$

4. If one now includes matter, then a comparison with the (rescaled) (21.80) and (21.84),

$$(16\pi G_N)\mathcal{H} = -2\sqrt{h}G_{NN} \quad , \quad (16\pi G_N)\mathcal{H}_a = -2\sqrt{h}G_{Na} \ , \tag{21.104}$$

shows that correspondingly there is no explicit factor of $(16\pi G_N)$ in the constraints either, which will be modified to

$$\mathcal{H} + \sqrt{h}T_{NN} = 0 \quad , \quad \mathcal{H}_a + \sqrt{h}T_{Na} = 0 \ . \tag{21.105}$$

21.8 Secondary Constraints: the Hamiltonian and Momentum Constraints

From the 1st-order form (21.98)

$$S_{ADM} = \int dt \ d^3x \ (\pi^{ab}\dot{h}_{ab} - N\mathcal{H} - \mathcal{N}^a\mathcal{H}_a) \tag{21.106}$$

of the ADM action, it is now manifest that variations of the action with respect to the lapse and shift give rise to the Hamiltonian and Momentum constraints respectively,

$$\begin{aligned} \frac{\delta S_{ADM}}{\delta N} = 0 \quad &\Rightarrow \quad \mathcal{H} = 0 \\ \frac{\delta S_{ADM}}{\delta \mathcal{N}^a} = 0 \quad &\Rightarrow \quad \mathcal{H}_a = 0 \end{aligned} \tag{21.107}$$

In the Hamiltonian picture, these constraints arise and are implemented by demanding that the so-called *primary constraints* (21.78)

$$p_N = 0 \quad , \quad p_{\mathcal{N}^a} = 0 \quad , \tag{21.108}$$

are preserved under the Hamiltonian time-evolution. This indeed gives rise precisely to the Momentum and Hamiltonian constraints as *secondary constraints*,

$$\begin{aligned} \dot{p}_N = \{p_N, H_{ADM}\} = 0 \quad &\Leftrightarrow \quad \mathcal{H} = 0 \quad \Leftrightarrow \quad G_{NN} = 0 \\ \dot{p}_{\mathcal{N}^a} = \{\dot{p}_{\mathcal{N}^a}, H_{ADM}\} = 0 \quad &\Leftrightarrow \quad \mathcal{H}_a = 0 \quad \Leftrightarrow \quad G_{Na} = 0 \quad . \end{aligned} \tag{21.109}$$

The remaining (true evolution) equations $G_{ab} = 0$ are then the Hamilton equations for the spatial metric (configuration variable) h_{ab} and its conjugate momentum π^{ab}. These can either be written in the form

$$\dot{h}_{ab} = \frac{\delta H_{ADM}}{\delta \pi^{ab}} \quad , \quad \dot{\pi}^{ab} = -\frac{\delta H_{ADM}}{\delta h_{ab}} \tag{21.110}$$

or in terms of Poisson brackets as

$$\dot{h}_{ab} = \{h_{ab}, H_{ADM}\} \quad , \quad \dot{\pi}^{ab} = \{\pi^{ab}, H_{ADM}\} \quad , \tag{21.111}$$

where the non-vanishing Poisson brackets between the canonical variables h_{ab} and π^{ab} are

$$\{h_{ab}(x), \pi^{cd}(y)\} = \tfrac{1}{2}(\delta_a^b \delta_c^d + \delta_a^d \delta_b^c)\delta(x,y) \quad . \tag{21.112}$$

Inserting the explicit expression for the Hamiltonian, one finds (unsurprisingly) that the equation for $\dot{h}_{ab}$ simply reproduces the definition of π^{ab}, i.e. the relation (21.77). Indeed, from the kinetic term of the Hamiltonian constraint one finds

$$\{h_{ab}, \int d^3x \; N G_{abcd} \pi^{ab} \pi^{cd}/\sqrt{h}\} = \frac{2N}{\sqrt{h}} G_{abcd} \pi^{cd} \quad , \tag{21.113}$$

and the Poisson bracket with the potential term is zero

$$\{h_{ab}, \int d^3x \; N\sqrt{h}\bar{R}\} = 0 \tag{21.114}$$

(because $\bar{R} = \bar{R}(h)$ is only a function of h_{ab} and its spatial derivatives). Finally, the Poisson bracket with the momentum constraint part of the ADM Hamiltonian gives

$$\begin{aligned} \{h_{ab}, \int d^3x \; (-2\mathcal{N}_c \bar{\nabla}_d \pi^{cd})\} &= \{h_{ab}, \int d^3x \; (+2(\bar{\nabla}_d \mathcal{N}_c)\pi^{cd})\} \\ &= \bar{\nabla}_a \mathcal{N}_b + \bar{\nabla}_b \mathcal{N}_a \quad . \end{aligned} \tag{21.115}$$

Putting everything together, one sees that this indeed reproduces (21.77) in the form

$$\dot{h}_{ab} = \frac{2N}{\sqrt{h}} G_{abcd} \pi^{cd} + \bar{\nabla}_a \mathcal{N}_b + \bar{\nabla}_b \mathcal{N}_a \quad . \tag{21.116}$$

The equation for $\dot{\pi}^{ab}$,

$$\dot{\pi}^{ab} = -\frac{\delta H_{ADM}}{\delta h_{ab}} = \{\pi^{ab}, H_{ADM}\} \quad , \tag{21.117}$$

is now equivalent to $G_{ab} = 0$. The explicit expression can of course be worked out in analogy with the above and from the results obtained so far, but it is rather complicated (these are, after all, the non-linear coupled Einstein equations) and not particularly enlightning, at least not upon first sight, and will not be given here. A partial result, however, namely the Poisson bracket of π^{ab} with the momentum constraint part of the Hamiltonian, $\{\pi^{ab}, \int \mathcal{N}^a \mathcal{H}_a\}$, will be given below as it illustrates the significance of the momentum constraint.

21.9 Properties and Significance of the Constraints

Above we saw that consistency of the primary constraints (21.108)

$$p_N = 0 \quad , \quad p_{\mathcal{N}^a} = 0 \quad , \tag{21.118}$$

i.e. the condition that they are preserved under the time-evolution generated by the ADM Hamiltonian, leads to the secondary Hamiltonian and Momentum constraints (21.109),

$$\dot{p}_N = 0 \quad , \quad \dot{p}_{\mathcal{N}^a} = 0 \quad \Rightarrow \quad \mathcal{H} = 0 \quad , \quad \mathcal{H}_a = 0 \ . \tag{21.119}$$

One now needs to inquire whether further (tertiary, ...) constraints are generated by the requirement that these secondary constraints are preserved under time-evolution. It turns out that the story ends here and that no further constraints are required. Thus the Hamiltonian and Momentum constraints will be satisfied at all times provided that they are satisfied on the initial value surface. This is the Hamiltonian counterpart of the statement about propagation of the constraints discussed in section 19.7 in connection with the Bianchi identities and their implications.

From the form of the ADM Hamiltonian,

$$H_{ADM} = \int d^3y \, (N(y)\mathcal{H}(y) + \mathcal{N}^b(y)\mathcal{H}_b(y)) \tag{21.120}$$

and

$$\begin{aligned} \frac{d}{dt}\mathcal{H}(x) &= \int d^3y \, N(y)\{\mathcal{H}(x), \mathcal{H}(y)\} + \int d^3y \, \mathcal{N}^b(y)\{\mathcal{H}(x), \mathcal{H}_b(y)\} \\ \frac{d}{dt}\mathcal{H}_a(x) &= \int d^3y \, N(y)\{\mathcal{H}_a(x), \mathcal{H}(y)\} + \int d^3y \, \mathcal{N}^b(y)\{\mathcal{H}_a(x), \mathcal{H}_b(y)\} \quad , \end{aligned} \tag{21.121}$$

it is clear that checking this amounts to calculating the Poisson brackets among the constraints and verifying that these are zero when the constraints themselves are satisfied, i.e. that the Poisson bracket algebra of the secondary constraints actually "closes" on the secondary constraints.

Even though this fact, and the resulting "surface deformation algebra", are in some sense one of the most interesting aspects of this entire story, we will skip the direct calculation of the Poisson bracket algebra of the constraints here as it is somewhat painful. However, it is useful to at least display the algebra of constraints (and we will then at least partially verify it afterwards).

The Poisson brackets among the "naked" constraints $\mathcal{H}(x), \mathcal{H}(y), \mathcal{H}_a(x), \mathcal{H}_b(y)$, as they appear in (21.121), will involve delta-functions $\delta(x,y)$ and their derivatives and are a bit unattractive (see e.g. (21.130) and (21.138) below). In order to exhibit the Poisson bracket algebra and clarify its structure, it is more instructive and convenient to explicitly introduce the "smeared" constraints

$$H[N] = \int d^3x\, N\mathcal{H} \quad , \quad P[\mathcal{N}] = \int d^3x\, \mathcal{N}^a\mathcal{H}_a \ , \tag{21.122}$$

in terms of which the ADM Hamiltonian takes the form

$$H_{ADM}[N,\mathcal{N}] = H[N] + P[\mathcal{N}] \ . \tag{21.123}$$

Then the Poisson bracket algebra of the constraints is found to be

$$\begin{aligned} \{H[N_1], H[N_2]\} &= P[N_1\bar{\nabla}N_2 - N_2\bar{\nabla}N_1] \\ \{P[\mathcal{N}], H[N]\} &= H[L_{\mathcal{N}}N] \\ \{P[\mathcal{N}_1], P[\mathcal{N}_2]\} &= P[[\mathcal{N}_1,\mathcal{N}_2]] \ . \end{aligned} \tag{21.124}$$

Here the new lapse function and shift vectors appearing on the right-hand side are

1. the lapse function
$$L_{\mathcal{N}}N = \mathcal{N}^a\partial_a N \ , \tag{21.125}$$
i.e. the (Lie) derivative of the lapse N along the shift vector field $\mathcal{N}^a$;

2. the shift vector field
$$[\mathcal{N}_1,\mathcal{N}_2] = L_{\mathcal{N}_1}\mathcal{N}_2 = -L_{\mathcal{N}_2}\mathcal{N}_1 \ , \tag{21.126}$$
i.e. the Lie bracket (9.22) among the shift vector fields $\mathcal{N}_1$ and $\mathcal{N}_2$, with components
$$[\mathcal{N}_1,\mathcal{N}_2]^a = \mathcal{N}_1^b\partial_b\mathcal{N}_2^a - \mathcal{N}_2^b\partial_b\mathcal{N}_1^a \tag{21.127}$$

3. and finally a shift vector constructed from the two lapse functions N_1 and N_2 (*and the metric*!), denoted by $N_1\bar{\nabla}N_2 - N_2\bar{\nabla}N_1$, which has the components
$$\begin{aligned} (N_1\bar{\nabla}N_2 - N_2\bar{\nabla}N_1)^a &= N_1\bar{\nabla}^a N_2 - N_2\bar{\nabla}^a N_1 \\ &\equiv h^{ab}(N_1\partial_b N_2 - N_2\partial_b N_1) \end{aligned} \tag{21.128}$$

There are many things that can and should be said about this algebra, its properties, its interpretation, its deeeper meaning, and its consequences, but in the following I will just make some rather elementary and simplistic comments.

First of all, recall that we expect the Hamiltonian and Momentum constraints to reflect the general covariance of general relativity. This general covariance is manifest in the covariant 4-dimensional Lagrangian formulation, but the Hamiltonian formulation requires a split of the 4-dimensional space-time into space and time via the choice of a foliation of space-time by spacelike hypersurfaces Σ_t, encoded in the choice of a time evolution vector field ∂_t with (21.16)

$$(\partial_t)^\alpha = NN^\alpha + E^\alpha_a \mathcal{N}^a \quad . \tag{21.129}$$

In this Hamiltonian formulation, spatial general covariance is still manifest, and this is reflected in the fact that the part of the constraint algebra that is easiest to understand is the algebra among the momentum constraints, or its "naked" counterpart

$$\{\mathcal{H}_a(x), \mathcal{H}_b(y)\} = \mathcal{H}_b(x)\partial_{x^a}\delta(x,y) + \mathcal{H}_a(x)\partial_{x^b}\delta(x,y) \quad . \tag{21.130}$$

Indeed, the Momentum constraint $P[\xi]$ associated to some vector field ξ on Σ implements the action of (infinitesimal) spatial diffeomorphisms

$$\delta_\xi x^a = \xi^a \tag{21.131}$$

on the phase space variables, and thus on functions on phase space, namely the Lie derivative L_ξ, via Poisson brackets,

$$\delta_\xi F(h_{ab}, \pi^{ab}) \equiv \{F(h_{ab}, \pi^{ab}), P[\xi]\} = L_\xi F(h_{ab}, \pi^{ab}) \tag{21.132}$$

(a proof of this is postponed to the very end of this section). As discussed in various ways in section 9, the Lie derivative provides a representation of the Lie algebra of vector fields (with respect to the Lie bracket) on tensors,

$$[L_{\xi_1}, L_{\xi_2}] = L_{[\xi_1,\xi_2]} \quad , \tag{21.133}$$

and the momentum constraint algebra shows that on phase space variables this representation is lifted to a representation at the level of Poisson brackets,

$$\{P[\xi_1], P[\xi_2]\} = P[[\xi_1,\xi_2]] \quad . \tag{21.134}$$

Correspondingly, the Poisson bracket

$$\{H[N], P[\xi]\} = -H[L_\xi N] \tag{21.135}$$

among the Hamiltonian and Momentum constraints, in its "half-naked" form

$$\{\mathcal{H}(x), P[\xi]\} = \partial_a(\xi^a \mathcal{H})(x) \quad , \tag{21.136}$$

simply expresses the fact that $\mathcal{H}$ is (and hence transforms as) a scalar density (9.70),

$$\delta_\xi \mathcal{H} = L_\xi \mathcal{H} = \partial_a(\xi^a \mathcal{H}) \ . \tag{21.137}$$

The remaining Poisson bracket relation, among the Hamiltonian constraints,

$$\begin{aligned} &\{H[N_1], H[N_2]\} = P[N_1 \bar{\nabla} N_2 - N_2 \bar{\nabla} N_1] \\ \Leftrightarrow \quad &\{\mathcal{H}(x), \mathcal{H}(y)\} = (h^{ab}(x)\mathcal{H}_b(x) + h^{ab}(y)\mathcal{H}_b(y))\partial_{x^a}\delta(x,y) \end{aligned} \tag{21.138}$$

is somewhat more enigmatic. In particular, since the right-hand side is field-dependent (as it depends on h^{ab}), the Hamiltonian constraint does not complete the 3-dimensional diffeomorphism Lie algebra of spatial vector fields (represented by the momentum constraints) to the 4-dimensional diffeomorphism Lie algebra of space-time vector fields.

Rather, the Poisson bracket algebra of the constraints represents what is known as the *surface deformation algebra*, subtly different from the algebra of space-time diffeomorphisms, as it acts not on the space-time but on the space of embeddings of spatial hypersurfaces.

Thinking of the surfaces Σ in terms of an embedding $x^\alpha(y^a)$, with

$$E_a^\alpha(y) = \partial_{y^a} x^\alpha \tag{21.139}$$

this surface deformation algebra is essentially the algebra generated by the

$$\mathcal{C}_a(y) = E_a^\alpha(y) \frac{\delta}{\delta x^\alpha(y)} \tag{21.140}$$

(which do indeed generate coordinate transformations on the hypersurface), and the

$$\mathcal{C}(y) = N^\alpha \frac{\delta}{\delta x^\alpha(y)} \tag{21.141}$$

with N^α the unit normal vector field to the hypersurfaces (generating normal deformations of the hypersurfaces).

<u>Remarks:</u>

1. From the surface deformation algebra and the requirement that evolution from a hypersurface Σ_i to a hypersurface Σ_f should be independent of how one slices / foliates the space-time between the two hypersurfaces, one can *derive* that the vanishing of the generators $\mathcal{H}$ and $\mathcal{H}_a$ of this algebra must be imposed as constraints.[48]

[48] S. Hojman, K. Kuchar, C. Teitelboim, *Geometrodynamics regained*, Ann. Phys. (NY) 96 (1976) 88-135. For a detailed discussion of this and further references, see e.g. chapters 3 and 4.1 of C. Kiefer, *Quantum Gravity* (2nd edition).

2. The momentum constraint arising in general relativity is the exact counterpart of the Gauss law constraint in the canonical formulation of Maxwell theory discussed in section 21.5.

3. The Hamiltonian constraint, on the other hand, has no counterpart in Maxwell theory, and is a characteristic general feature of generally covariant (reparametrisation invariant) systems. Indeed, in a generally covariant system time evolution is in a sense a gauge symmetry because time-translation is the coordinate transformation $t \to t + c$. As a consequence, the generator of time-evolution, i.e. the Hamiltonian H, is also a constraint, i.e. constrained to be zero $H = 0$. This is something we already saw in (2.132) for the time-reparametrisation invariant action principle for geodesics in section 2.5.

4. Taken at face value, this suggests that in a generally covariant theory there is no dynamics, or that the dynamics is "frozen", and that the only allowed observables are functions on the phase space that Poisson-commute with the Hamiltonian, i.e. that are in some sense constants of motion. This cannot be strictly correct, of course, and the problem appears in a different light once one fixes a gauge, i.e. makes a choice of coordinates. Nevertheless, this does not solve all the problems and there are endless debates in the literature about these issues. In particular, the debate over what are acceptable observables in a generally covariant (quantum) theory continues to this day.[49]

We now turn to the proof of the relation (21.132), In order to establish (21.132), it is sufficient to show that the canonical Poisson brackets (21.112),

$$\{h_{ab}(x), \pi^{cd}(y)\} = \tfrac{1}{2}(\delta_a^b\delta_c^d + \delta_a^d\delta_b^c)\delta(x,y) \quad , \tag{21.142}$$

imply that

$$\begin{aligned} \delta_\xi h_{ab} &\equiv \{h_{ab}, P[\xi]\} = L_\xi h_{ab} \\ \delta_\xi \pi^{ab} &\equiv \{\pi^{ab}, P[\xi]\} = L_\xi \pi^{ab} \quad . \end{aligned} \tag{21.143}$$

The first relation is equivalent to the identity (21.115) already derived above, since

$$L_\xi h_{ab} = \bar{\nabla}_a \xi_b + \bar{\nabla}_b \xi_a \quad . \tag{21.144}$$

The proof of the 2nd relation is a bit more complicated as it also involves the metric variation of the Christoffel symbols appearing in the covariant derivative $\bar{\nabla}_d \pi^{cd}$. Recalling that π^{cd} is a tensor density,

$$\pi^{cd} = \sqrt{h}(K^{cd} - h^{cd}K) \equiv \sqrt{h}p^{cd} \quad , \tag{21.145}$$

[49]See e.g. C. Rovelli, *Quantum Gravity* or S. Giddings, D. Marolf, J. Hartle, *Observables in effective gravity*, `arXiv:hep-th/0512200` for different points of view and discussions of these issues.

this covariant derivative is

$$\bar{\nabla}_d \pi^{cd} = \sqrt{h}\bar{\nabla}_d p^{cd} = \partial_d \pi^{cd} + \bar{\Gamma}^c{}_{ed}\pi^{ed} \quad . \tag{21.146}$$

Likewise, the Lie derivative of this tensor density is (cf. sections 9.4 and 9.6)

$$\begin{aligned} L_\xi \pi^{ab} &= (L_\xi \sqrt{h})p^{ab} + \sqrt{h}L_\xi p^{ab} \\ &= \tfrac{1}{2}\sqrt{h}(h^{cd}L_\xi h_{cd})p^{ab} + \sqrt{h}L_\xi p^{ab} \\ &= \sqrt{h}(\bar{\nabla}_c \xi^c)p^{ab} + \sqrt{h}(\xi^c\bar{\nabla}_c p^{ab} - p^{cb}\bar{\nabla}_c\xi^a - p^{ac}\bar{\nabla}_c\xi^b) \\ &= \bar{\nabla}_c(\xi^c\pi^{ab}) - \pi^{cb}\bar{\nabla}_c\xi^a - \pi^{ac}\bar{\nabla}_c\xi^b \quad . \end{aligned} \tag{21.147}$$

Now, to calculate

$$\{\pi^{ab}(x), P[\xi]\} = -2\int d^3y\ \{\pi^{ab}(x), \xi_c(y)\bar{\nabla}_d\pi^{cd}(y)\} \tag{21.148}$$

we make the h_{ab}-dependence more explicit (but suppress the y-dependence in this equation),

$$\begin{aligned} \{\pi^{ab}(x), \xi_c\bar{\nabla}_d\pi^{cd}\} &= \{\pi^{ab}(x), h_{ce}\xi^e\bar{\nabla}_d\pi^{cd}\} \\ &= \{\pi^{ab}(x), h_{ce}\}\xi^e\bar{\nabla}_d\pi^{cd} + h_{ce}\xi^e\{\pi^{ab}, \bar{\Gamma}^c{}_{df}\}\pi^{df} \quad . \end{aligned} \tag{21.149}$$

From the 1st term, one immediately obtains (from the canonical Poisson brackets, and with the factor of (-2) and the integration over y from (21.148))

$$\{\pi^{ab}(x), h_{ce}\}\xi^e\bar{\nabla}_d\pi^{cd} \quad \Rightarrow \quad \xi^b\bar{\nabla}_d\pi^{ad} + \xi^a\bar{\nabla}_d\pi^{bd} \quad . \tag{21.150}$$

For the calculation of the 2nd term, we observe that taking the Poisson bracket with π^{ab} is equivalent to taking (minus) the variation with respect to h_{ab}. We can therefore use the formula (20.14) for the variation of the Christoffel symbols under metric variations,

$$h_{ce}\delta_h\bar{\Gamma}^c{}_{df} = \tfrac{1}{2}(\bar{\nabla}_f\delta h_{ed} + \bar{\nabla}_d\delta h_{ef} - \bar{\nabla}_e\delta h_{df}) \quad . \tag{21.151}$$

Now an integration by parts (moving the derivatives off the delta-functions) shows that the 2nd term contributes

$$h_{ce}\xi^e\{\pi^{ab}, \bar{\Gamma}^c{}_{df}\}\pi^{df} \quad \Rightarrow \quad \bar{\nabla}_c(\xi^c\pi^{ab}) - \bar{\nabla}_d(\xi^a\pi^{bd}) - \bar{\nabla}_d(\xi^b\pi^{ad}) \quad . \tag{21.152}$$

Putting everything together, one finds precisely the Lie derivative (21.147),

$$\delta_\xi\pi^{ab} = \{\pi^{ab}, P[\xi]\} = \bar{\nabla}_c(\xi^c\pi^{ab}) - \pi^{cb}\bar{\nabla}_c\xi^a - \pi^{ac}\bar{\nabla}_c\xi^b = L_\xi\pi^{ab} \tag{21.153}$$

21.10 Boundary Terms in the ADM Action and Hamiltonian

So far in this section we have assumed that the spatial slices Σ have no boundary, $\partial\Sigma = \emptyset$, and we have therefore also ignored possible boundary terms that are required or generated by the presence of such a boundary. In the remainder of this section, i.e. here and in subsections 21.11 and 21.12 below, we will look at some of the issues and features that arise when one takes these into account.

To set the stage, recall that we saw in section 20.5 that differentiability of the gravitational action in the sense of variational calculus, i.e. (20.68)

$$\delta S_g[g_{\alpha\beta}] = \int \sqrt{g}\, G_{\alpha\beta}\delta g^{\alpha\beta} \tag{21.154}$$

without boundary terms on the right-hand side for Dirichlet boundary conditions, can be achieved e.g. by adding the Gibbons-Hawking-York boundary term to the Einstein-Hilbert action (20.67),

$$\begin{aligned} S_g[g_{\alpha\beta}] &= S_{EH}[g_{\alpha\beta}] + S_{GHY}[g_{\alpha\beta}] \\ &= \int \sqrt{g}\, R + 2\epsilon \oint \sqrt{h}\, K \quad . \end{aligned} \tag{21.155}$$

Moreover, we saw in section 21.2 that the Gauss-Codazzi decomposition of the Ricci scalar (21.5) ,

$$R = \bar{R} + \epsilon(K^2 - K^{\alpha\beta}K_{\alpha\beta}) - 2\nabla_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) \quad , \tag{21.156}$$

automatically takes care of the Gibbons-Hawking-York boundary term for "initial" and "final" spacelike hypersurfaces Σ_i and Σ_f which are part of the foliation of the space-time M into spacelike hypersurfaces Σ_t, with normal vector N^α, since

$$N_\alpha(N^\beta\nabla_\beta N^\alpha - N^\alpha\nabla_\beta N^\beta) = -N_\alpha N^\alpha\nabla_\beta N^\beta = K_\Sigma \quad . \tag{21.157}$$

We also noted in section 21.2 that, if in addition to initial and final spacelike boundaries Σ there is a timelike boundary $\mathcal{B}$,

$$\partial M = \{\Sigma_f\} \cup \{-\Sigma_i\} \cup \mathcal{B} \quad , \tag{21.158}$$

then additional boundary terms are required in the action (and also in the Hamiltonian). We will assume in the following that the boundary $\mathcal{B}$ is orthogonal to the spatial slices Σ in the sense that the normal N^α to Σ is orthogonal to the normal r^α to $\mathcal{B}$,

$$N^\alpha r_\alpha = 0 \quad . \tag{21.159}$$

Specifically, if the constant time hypersurfaces $\Sigma = \Sigma_t$ have (asymptotic) boundary $S_t = \partial\Sigma_t$, then the timelike boundary $\mathcal{B}$ is the union of all these surfaces S_t.[50]

[50]For a discussion of boundary terms for non-orthogonal boundaries see e.g. S. Hawking, C. Hunter, *The Gravitational Hamiltonian in the Presence of Non-Orthogonal Boundaries*, `arXiv:gr-qc/9603050`, and I. Booth, R. Mann, *Moving Observers, Non-orthogonal Boundaries, and Quasilocal Energies*, `arXiv:gr-qc/9810009`, and references thereto.

Tracing back through the various derivations in this section, one finds that there are 2 contributions to this boundary term for the action on $\mathcal{B}$ (in addition, later on we will identify a boundary term contribution to the Hamiltonian, and thus to the 1st order Hamiltonian form of the ADM action):

1. Gibbons-Hawking-York Boundary Term

   One contribution to the gravitational action not (completely) accounted for yet is the Gibbons-Hawking-York boundary term associated with the boundary $\mathcal{B}$, i.e. the term
   $$S_1 = 2\epsilon \oint_{\mathcal{B}} K_{\mathcal{B}} = 2 \oint \sqrt{h_{\mathcal{B}}}(g^{\alpha\beta} - r^\alpha r^\beta)\nabla_\alpha r_\beta \quad . \tag{21.160}$$
   Here r^α is the unit (outward-pointing) normal to $\mathcal{B}$, $r^\alpha r_\alpha = +1$, $h_{\mathcal{B}}$ is the absolute value of the determinant of the metric
   $$h_{\mathcal{B}\alpha\beta} = (g_{\alpha\beta} - r_\alpha r_\beta)|_{\mathcal{B}} \quad , \quad h_{\mathcal{B}\alpha\beta} r^\beta = 0 \tag{21.161}$$
   induced on $\mathcal{B}$. The projection term in the expression for the trace of the extrinsic curvature is not strictly speaking necessary, since $r^\alpha r_\alpha = 1$ implies
   $$(g^{\alpha\beta} - r^\alpha r^\beta)\nabla_\alpha r_\beta = g^{\alpha\beta}\nabla_\alpha r_\beta \equiv \nabla_\alpha r^\alpha \tag{21.162}$$
   but it will be instructive to keep it.

2. Gauss-Codazzi Boundary Term

   Adopting the Gauss-Codazzi decomposition of the Einstein-Hilbert Lagrangian, there is an additional boundary-term contribution at $\mathcal{B}$ from the total-derivative term, namely
   $$S_2 = -2 \oint_{\mathcal{B}} \sqrt{h_{\mathcal{B}}}\, r_\alpha (N^\beta \nabla_\beta N^\alpha - N^\alpha \nabla_\beta N^\beta) \quad . \tag{21.163}$$
   This is not equal to the standard Gibbons-Hawking-York boundary term for this boundary component (which would be expressed solely in terms of r^α, not also N^α). Using the assumption of orthogonality $r_\alpha N^\alpha = 0$, and $N^\alpha N_\alpha = -1$, this can by an integration by parts be written as
   $$S_2 = +2 \oint_{\mathcal{B}} \sqrt{h_{\mathcal{B}}}\, N^\alpha N^\beta \nabla_\alpha r_\beta \quad . \tag{21.164}$$

These two contributions combine into
$$S_1 + S_2 = 2 \oint_{\mathcal{B}} \sqrt{h_{\mathcal{B}}}\, s^{\alpha\beta}\nabla_\alpha r_\beta \quad , \tag{21.165}$$
where
$$s_{\alpha\beta} = g_{\alpha\beta} + N_\alpha N_\beta - r_\alpha r_\beta \tag{21.166}$$

with

$$s_{\alpha\beta}N^\beta = s_{\alpha\beta}r^\beta = 0 \quad . \tag{21.167}$$

Thus $s_{\alpha\beta}$ represents the metric induced on the boundary surfaces

$$S_t = \partial\Sigma_t = \Sigma_t \cap \mathcal{B} \quad . \tag{21.168}$$

As a consequence, we have

$$\sqrt{h_{\mathcal{B}}} = N\sqrt{s} \tag{21.169}$$

and

$$k_S = s^{\alpha\beta}\nabla_\alpha r_\beta \tag{21.170}$$

is the extrinsic curvature of S_t in Σ_t. Thus this new boundary term modifies the 2nd order ADM form (21.42) of the complete gravitational action to

$$S_{ADM} = \int dt \left[\int_\Sigma d^3x\,\sqrt{h}N(\bar{R} + K^{ab}K_{ab} - K^2) + 2\oint_{S_t} d^2x\sqrt{s}Nk_S\right] \quad . \tag{21.171}$$

The Legendre transform of the ADM Lagrangian to the Hamiltonian will thus, in particular, also lead to a boundary term in the ADM Hamiltonian (21.97), namely (reinserting the coupling constant)

$$\begin{aligned} & L_{ADM} \to L_{ADM} + \frac{1}{8\pi G_N}\oint_{S_t}\sqrt{s}d^2x\; Nk_S \\ \Rightarrow \quad & H_{ADM} \to H_{ADM} - \frac{1}{8\pi G_N}\oint_{S_t}\sqrt{s}d^2x\; Nk_S \quad . \end{aligned} \tag{21.172}$$

However, in performing the Legendre transformation, we obtained (rather: neglected) yet one more boundary term, namely from the integration by parts in (21.95),

$$2\pi^{ab}\bar{\nabla}_a\mathcal{N}_b = 2\bar{\nabla}_a(\pi^{ab}\mathcal{N}_b) + \mathcal{N}^a\mathcal{H}_a \quad . \tag{21.173}$$

Therefore the total Hamiltonian in the presence of timelike boundaries (or: when the spatial slices Σ have boundaries) has the form

$$\begin{aligned} H_{ADM} = & \int_\Sigma d^3x\;(N\mathcal{H} + \mathcal{N}^a\mathcal{H}_a) \\ & - \frac{1}{8\pi G_N}\oint_S\sqrt{s}d^2x\; Nk_S + \frac{1}{8\pi G_N}\oint_{S_t} d^2x\;\mathcal{N}_a\pi^{ab}r_b \quad . \end{aligned} \tag{21.174}$$

<u>Remarks:</u>

1. The necessity of these boundary terms in the Hamiltonian can also be understood form the requirement of having a differentiable Hamiltonian in the sense of variational calculus and, as we will see in section 21.11 below, this provides an alternative route to determining these boundary terms.

2. The other significance of these boundary terms lies in the fact that they give the "on-shell" value of the Hamiltonian, i.e. the value of the Hamiltonian on a solution satisfying (in particular) the Hamiltonian and Momentum constraints, namely

$$\mathcal{H} = \mathcal{H}_a = 0 \quad \Rightarrow \quad H_{ADM} = -\frac{1}{8\pi G_N}\oint_S d^2x\, \left(N\sqrt{s}k_S - \mathcal{N}_a\pi^{ab}r_b\right) \quad . \tag{21.175}$$

In particular, as such they provide a candidate definition of the "energy" of a solution. This will be briefly discussed in section 21.12 below.

21.11 Alternative Derivation of the Hamiltonian Boundary Terms

Turning to the 1st issue, recall that the Hamiltonian equations of motion are assumed to be (21.110)

$$\dot{h}_{ab} = \frac{\delta H_{ADM}}{\delta\pi^{ab}} \quad , \quad \dot{\pi}^{ab} = -\frac{\delta H_{ADM}}{\delta h_{ab}} \tag{21.176}$$

However, validity of these equations (differentiability of the Hamiltonian in the sense of variational calculus) requires that the variation of the Hamiltonian with respect to the canonical variables h_{ab} and π^{ab} has the form

$$\delta H_{ADM}[h_{ab},\pi^{ab}] = \int d^3x \left[(\ldots)^{ab}\delta h_{ab} + (\ldots)_{ab}\delta\pi^{ab}\right] \tag{21.177}$$

without any boundary terms. Analysing the bulk Hamiltonian

$$H_{ADM} \stackrel{?}{=} \int_\Sigma d^3x\, \left(N\mathcal{H} + \mathcal{N}^a\mathcal{H}_a\right) \quad , \tag{21.178}$$

with (21.102)

$$\mathcal{H} = \frac{(16\pi G_N)}{\sqrt{h}}G_{abcd}\pi^{ab}\pi^{cd} - \frac{\sqrt{h}}{(16\pi G_N)}\bar{R} \tag{21.179}$$

and

$$\mathcal{H}^a = -2\bar{\nabla}_b\pi^{ba} \tag{21.180}$$

we see that

- no boundary term arises from the variation of the 1st (kinetic) term in the Hamiltonian constraint, as it does not depend on the derivatives of the canonical variables;
- a boundary term will arise from the variation of the 2nd (potential) term $\bar{R}$, as it depends on the 2nd derivatives of h_{ab};
- a boundary term will arise from the integration by parts required to express the variation of the Momentum constraint as $(\ldots)\delta\pi^{ab}$.

The latter issue is obviously taken care of by simply reinstating the total derivative term in (21.173) and adding it to H_{ADM}. This immediately leads to the 2nd boundary term in (21.174),

$$H_{ADM} \to H_{ADM} + \frac{1}{8\pi G_N}\oint_S d^2x\; \mathcal{N}_a \pi^{ab} r_b \quad . \tag{21.181}$$

In order to resolve the issue arising from the variation of $\bar{R}$, we can observe that this is simply the 3-dimensional counterpart of the issue that arises when varying the Einstein-Hilbert action with Lagrangian R. Thus we can appeal to the discussion of the Gibbons-Hawking-York boundary term in section 20.5 to conclude that the required boundary term to be added involves the trace k_S of the extrinsic curvature of the boundary $S_t = \partial\Sigma_t$ in Σ_t. Noting that (a) the normal vector is spacelike ($\epsilon = +1$), and (b) that the Hamiltonian involves $(-N\bar{R})$ rather than $(+R)$, we deduce from (21.155), say, that this requires modifying the bulk Hamiltonian H_{ADM} according to

$$H_{ADM} \to H_{ADM} - \frac{1}{8\pi G_N}\oint_S \sqrt{s} d^2x\; N k_S \quad . \tag{21.182}$$

We thus conclude that validity of the Hamiltonian equations of motion in the presence of a spatial boundary $S = \partial\Sigma$ requires adding boundary terms to the bulk Hamiltonian according to

$$\begin{aligned} H_{ADM} = &\int_\Sigma d^3x\; (N\mathcal{H} + \mathcal{N}^a \mathcal{H}_a) \\ &- \frac{1}{8\pi G_N}\oint_S \sqrt{s} d^2x\; N k_S + \frac{1}{8\pi G_N}\oint_S d^2x\; \mathcal{N}_a \pi^{ab} r_b \quad . \end{aligned} \tag{21.183}$$

This is identical to the result (21.174) obtained before by different means.

21.12 Significance of the Hamiltonian Boundary Terms: ADM Energy

As mentioned at the end of section 21.10, the value of the Hamiltonian on a configuration (h_{ab}, π^{ab}) satisfying the Hamiltonian and Momentum constraints is

$$H_{ADM}[N, \mathcal{N}^a] = -\frac{1}{8\pi G_N}\oint_S d^2x\; \left(N\sqrt{s} k_S - \mathcal{N}_a \pi^{ab} r_b\right) \quad . \tag{21.184}$$

While in the spatially closed case, $\partial\Sigma = \emptyset$, this on-shell value of the Hamiltonian is zero, it has a significance e.g. for asymptotically flat space-times (the prototypical example being the Schwarzschild metric). While in this case there is no spatial boundary $\partial\Sigma = S$ in the strict sense, in the asymptotically flat context variations of h_{ab} should be restricted to preserve this asymptotic flatness. The boundary terms in the Hamiltonian derived above are also appropriate in this setting (as one is essentially imposing the Dirichlet condition $h_{ab} = \delta_{ab}$ "at infinity").[51]

[51] Defining and implementing the conditions for asymptotic flatness requires and merits more care. See e.g. R. Wald, *General Relativity*, chapter 11, for a careful discussion of all the issues we are glossing over in the following.

For a given configuration (h_{ab}, π^{ab}), the Hamiltonian H_{ADM} is a functional of (the asymptotic values of) the lapse N and shift vector $\mathcal{N}^a$. Recalling (21.16),

$$(\partial_t)^\alpha = NN^\alpha + E^\alpha_a \mathcal{N}^a \quad , \tag{21.185}$$

and noting that asymptotically the time-evolution of static observers in the Minkowskian geometry at infinity is orthogonal to the spatial directions, the choice $N = 1$ and $\mathcal{N}^a = 0$ (asymptotically) gives the value of the Hamiltonian associated to asymptotic time-translations. As such, it provides a candidate definition of the gravitational energy of a configuration (h_{ab}, π^{ab}), the ADM energy

$$E \stackrel{?}{=} -\frac{1}{8\pi G_N} \lim \oint_S d^2x \, \sqrt{s} k_S \quad . \tag{21.186}$$

For a boosted observer at infinity, his proper time would correspond to a non-trivial linear combination of the 2 terms in (21.185), hence to a non-trivial shift vector $\mathcal{N}^a$. The second term in (21.184), depending on $\mathcal{N}^a$ is therefore naturally associated with a linear momentum (and other choices of lapse and shift can be used analogously to define candidate notions of angular momentum etc.), but we will not explore this further here.

The above candidate expression (21.186) for the energy still requires some improvements. First of all, the limit here refers to taking the boundary 2-sphere S to infinity. This can be implemented more concretely by introducing asymptotically a Cartesian coordinate system on Σ, with an associated notion of radial distance r and considering the limit of the coordinate spheres S_R of radius $r = R$ as $R \to \infty$. Thus we can write a somewhat improved version of (21.186) as

$$E \stackrel{?}{=} -\frac{1}{8\pi G_N} \lim_{R\to\infty} \oint_{S_R} d^2x \, \sqrt{s} k_S \quad . \tag{21.187}$$

The problem with this expression is that unfortunately it diverges even for the flat metric $h^0_{ab} = \delta_{ab}$ on Σ,

$$h^0_{ab} dy^a dy^b = \delta_{ab} dy^a dy^b = dr^2 + r^2 d\Omega^2 \quad . \tag{21.188}$$

Indeed, the trace of the extrinsic curvature of a 2-sphere S_R of radius R is (18.38)

$$k^0_S = \frac{2}{R} \quad , \tag{21.189}$$

while $\sqrt{s} = R^2 \sin\theta$, so that

$$\oint_{S_R} d^2x \, \sqrt{s} k^0_S = 4\pi R^2 \, \frac{2}{R} = 8\pi R \to \infty \quad . \tag{21.190}$$

It is natural to assign the energy $E = 0$ to Minkowski space (and its flat slices), and it is therefore also reasonably natural to subtract this divergent contribution from E in (21.187). We thus finally arrive at the definition of the *ADM Energy*

$$E_{ADM} = -\frac{1}{8\pi G_N} \lim_{R\to\infty} \oint_{S_R} d^2x \, \sqrt{s}(k_S - k^0_S) \quad . \tag{21.191}$$

Here k_S^0 is defined to be the extrinsic curvature of S embedded in flat space $\mathbb{R}^3$ in such a way that the induced metric on S is the same as that induced on S by the metric h_{ab} on Σ (in particular, then, $\sqrt{s}$ is the same for both terms and therefore only appears as an overall factor in the integrand).

Note the similarity with the background-subtracted gravitational action (20.70)

$$S[g_{\alpha\beta}] = \int \sqrt{g} d^4x \, R + 2\epsilon \oint_\Sigma \sqrt{h} d^3y \, (K - K^0) \quad , \tag{21.192}$$

briefly mentioned in section 20.5, and which would have also led us to (21.191).

To see that (21.191) gives a finite and meaningful result in cases of interest, we consider the prime example of an asymptotically flat solution to the Einstein equations, namely the Schwarzschild solution describing the exterior of a spherically symmetric star (see section 24 and subsequent sections for a detailed derivation and discussion of this metric).

In the standard Schwarzschild coordinates, this metric has the form (24.37)

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} \quad , \tag{21.193}$$

where the parameter m is related to the mass M of the star by $m = G_N M$ (in units with $c = 1$). We can directly work with the (sufficiently simple) exact expression for the metric, but it will be sufficient to look at the asymptotic (large r) behaviour of the spatial metric on the slices Σ_t of constant time t. As a consequence, the following analysis applies not just to the Schwarzschild metric but to any metric of the above general form, with

$$f(r) = 1 - \frac{2m}{r} + \mathcal{O}(1/r^2) \quad . \tag{21.194}$$

To first order in an expansion in m/r, the metric on a hypersurface Σ is given by

$$ds^2|_{t=t_0} \approx (1 + 2m/r)dr^2 + r^2 d\Omega^2 \quad . \tag{21.195}$$

Alternatively, in so-called isotropic coordinates, the metric takes the form given in (24.46), and the asymptotic form of the spatial metric on the slices Σ_t of constant time t is (calling the radial coordinate r again)

$$ds^2|_{t=t_0} \approx (1 + 2m/r)(dr^2 + r^2 d\Omega^2) \quad . \tag{21.196}$$

Note that even though the (rr)-components of the metric (and hence the radial normal vector to the spheres) is the same in both coordinate systems,

$$r_\alpha = (1 + 2m/r)^{1/2} \partial_\alpha r \approx (1 + m/r)\partial_\alpha r \quad , \tag{21.197}$$

the induced metric on the spheres is different. Thus even though in both cases the flat reference metric is simply

$$(ds^0)^2 = dr^2 + r^2 d\Omega^2 \quad , \quad r_\alpha^0 = \partial_\alpha r \quad , \tag{21.198}$$

also the required isometric embedding into flat space will have to be different in the two cases, and we will now see how these things conspire to give the same (and highly reasonable) result

$$E_{ADM} = M \quad . \tag{21.199}$$

We will use that the trace of the extrinsic curvature can be written as

$$k_S = \tfrac{1}{2} s^{AB} r^\alpha \partial_\alpha (s_{AB})|_{r=R} \tag{21.200}$$

where s_{AB} is the induced metric, and that

$$r^\alpha r_\alpha = 1 \quad \Rightarrow \quad r^\alpha \partial_\alpha r \approx (1 + m/r)^{-1} \approx (1 - m/r) \quad . \tag{21.201}$$

1. Schwarzschild coordinates

   The induced metric on S_R is $R^2 d\Omega^2$ and thus the extrinsic curvature is

   $$k_S|_{r=R} \approx \frac{2}{R}(1 - \frac{m}{R}) \quad . \tag{21.202}$$

   In the flat reference metric, one obtains the same induced metric if one also chooses the radius $r = R$, and (as above)

   $$k_S^0 = \frac{2}{R} \quad , \tag{21.203}$$

   so that

   $$k_S - k_S^0 \approx -\frac{2m}{R^2} \quad . \tag{21.204}$$

   Thus the ADM energy is

   $$E_{ADM} = -\frac{1}{8\pi G_N} \lim_{R\to\infty} \oint_{S_R} d\Omega \; R^2(-2m/R^2) = m/G_N = M \quad . \tag{21.205}$$

   In particular, this is finite (and reasonable).

   We also see from this that any subleading terms in a $(1/r)$-expansion would not have contributed to the integral in the limit, so it was consistent to ignore them throughout.

2. Isotropic Coordinates

   In isotropic coordinates, the induced metric on a sphere of radius $r = R$ is

   $$(1 + 2m/R)R^2 d\Omega^2 = (R^2 + 2mR)d\Omega^2 \quad . \tag{21.206}$$

   Therefore the trace of the extrinsic curvature is

   $$k_S \approx (1 - 2m/R)(1/R^2) r^\alpha \partial_\alpha (r^2 + 2mr)|_{r=R} \approx \frac{2}{R}(1 - \frac{2m}{R}) \tag{21.207}$$

   (note that this is not the same as the corresponding expression (21.202) in Schwarzschild coordinates). The reference term k_S^0 is the extrinsic curvature of a 2-sphere embedded in flat space whose induced metric is equal to (21.206), i.e. it is the extrinsic curvature of a sphere with radius R^0 characterised by

$$(R^0)^2 = (1+2m/R)R^2 \quad \Rightarrow \quad R^0 \approx (1+m/R)R \ . \tag{21.208}$$

Thus

$$k_S^0 = \frac{2}{R^0} \approx \frac{2}{R}(1-\frac{M}{R}) \ , \tag{21.209}$$

and therefore

$$k_S - k_S^0 \approx -\frac{2m}{R^2} \ , \tag{21.210}$$

as in Schwarzschild coordinates. In the $R\to\infty$ limit, only the leading term of the induced volume element will contribute, and therefore the result is indeed identical to that obtained in Schwarzschild coordinates,

$$\begin{aligned} E_{ADM} &= -\frac{1}{8\pi G_N}\lim_{R\to\infty}\oint_{S_R} d\Omega\ (R^2+2mR)(-2m/R^2) \\ &= -\frac{1}{8\pi G_N}\lim_{R\to\infty}\oint_{S_R} d\Omega\ (R^2)(-2m/R^2) = m/G_N = M \ . \end{aligned} \tag{21.211}$$

In section 23.4 we will encounter a seemingly different expression for the ADM energy, deduced and extrapolated there not from a canonical analysis but rather from the linearised Einstein equations, namely (23.33)

$$E_{ADM} = \frac{1}{16\pi G_N}\oint_{S^2_\infty} dS_i\ (\bar{\nabla}^k h_{ik} - \bar{\nabla}_i h) \ . \tag{21.212}$$

As mentioned there, it can be shown that this agrees with the canonical expression when the induced metrics on S_∞ agree. Moreover, in section 24.8 we will evaluate this expression for the Schwarzschild metric, again both in Schwarzschild and in isotropic coordinates, and reassuringly also find $E_{ADM} = M$ in this way.

22 Energy-Momentum Tensor II: Selected Topics

22.1 Energy Conditions

When confronted with the Einstein equations

$$G_{\alpha\beta} = 8\pi G_N \, T_{\alpha\beta} \quad , \tag{22.1}$$

one can either try to find exact solutions in certain specific situations, or one can try to learn or prove something in general about solutions to the Einstein equations.

For the former, one usually starts by specifying the matter content and the energy-momentum tensor (either phenomenologically or microscopically), and then furthermore imposes some symmetry conditions, and this is how we will usually proceed in other parts of these notes, when discussing e.g. solar system physics, black holes or cosmology. In this case, one thus in particular chooses (or is at least well-advised to choose) an energy-momentum tensor with reasonable and well-motivated physical properties from the outset.

For the latter, it is clear that in order to be able to say anything of substance at all, one needs to impose some conditions on the energy-momentum tensor $T_{\alpha\beta}$. After all, any metric whatsoever can be considered to be a solution of the Einstein equations, with "energy-momentum tensor" defined by

$$T_{\alpha\beta} = \frac{1}{8\pi G_N} \, G_{\alpha\beta} \quad . \tag{22.2}$$

The problem with this approach (this is sometimes referred to as the poor man's way of solving the Einstein equations, but this is too charitable a characterisation and maligning poor men) is that generically this candidate "energy-momentum tensor" will not have any of the very general properties one would usually associate with reasonable forms of matter.

Examples of such general requirements or reasonable properties are

- positivity of energy or energy density
- causal propagation of the energy flow of matter
- giving rise to an attractive (rather than repulsive) gravitational force
- ...

It turns out that (various combinations, or variants, of) such conditions can be implemented by imposing some simple general constraints on the energy-momentum tensor known as *Energy Conditions*. The simplest and most common among these take the form of pointwise conditions on the contraction of an energy-momentum tensor with causal (i.e. timelike or null, or non-spacelike) vectors. One can also consider weaker "averaged" versions of these conditions, averaged either along geodesics or over regions of space(-time), say, but we will only consider the pointwise conditions here.

1. Weak Energy Condition (WEC)

   Given an energy-momentum tensor $T_{\alpha\beta}$, the energy density seen by an observer with timelike (and future directed) 4-velocity tangent vector t^α is $T_{\alpha\beta}t^\alpha t^\beta$. The weak energy condition is the (plausible) statement that this is non-negative for any such observer,
$$T_{\alpha\beta}t^\alpha t^\beta \geq 0 \quad \forall\, t^\alpha : \quad t^\alpha t_\alpha < 0 \ . \tag{22.3}$$
   By continutiy, this inequality is then also valid for null vectors ℓ^α, i.e. for all causal vectors v^α,
$$T_{\alpha\beta}v^\alpha v^\beta \geq 0 \quad \forall\, v^\alpha : \quad v^\alpha v_\alpha \leq 0 \ . \tag{22.4}$$

2. Null Energy Condition (NEC)

   The null energy condition imposes the previous condition (22.4) only for null vectors,
$$T_{\alpha\beta}\ell^\alpha \ell^\beta \geq 0 \quad \forall\, \ell^\alpha : \quad \ell^\alpha \ell_\alpha = 0 \ . \tag{22.5}$$
   The rationale for this condition is that the null Raychaudhuri equation describing the focussing of null geodesic congruences is (12.107)
$$\frac{d}{d\tau}\theta_\ell = -R_{\alpha\beta}\ell^\alpha \ell^\beta + \ldots \tag{22.6}$$
   so that the geometry will have a focussing (attractive) effect on null geodesics if
$$R_{\alpha\beta}\ell^\alpha \ell^\beta \geq 0 \ . \tag{22.7}$$
   By the Einstein equations
$$R_{\alpha\beta} = 8\pi G_N (T_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} T) \equiv 8\pi G_N\, \bar{T}_{\alpha\beta} \tag{22.8}$$
   and because ℓ^α is null, this translates into the null energy condition (22.5).

3. Strong Energy Condition (SEC)

   Analogously, a term $R_{\alpha\beta}t^\alpha t^\beta$ appears in the Raychaudhuri equation (12.36) for timelike geodesic congruences,
$$\frac{d}{d\tau}\theta_t = -R_{\alpha\beta}t^\alpha t^\beta + \ldots \tag{22.9}$$
   Thus the geometry will have a focussing effect on timelike geodesic congruences (families of freely falling particles) if
$$R_{\alpha\beta}t^\alpha t^\beta \geq 0 \ . \tag{22.10}$$

By the Einstein equations, this can be rewritten as the condition

$$(T_{\alpha\beta} - \tfrac{1}{2}g_{\alpha\beta}T)t^\alpha t^\beta \equiv \bar{T}_{\alpha\beta}t^\alpha t^\beta \geq 0 \quad \forall\, t^\alpha : \quad t^\alpha t_\alpha < 0 \ . \tag{22.11}$$

Again by continuity this is then also true (or required to be satisfied) for all causal vectors v^α,

$$\bar{T}_{\alpha\beta}v^\alpha v^\beta \geq 0 \quad \forall\, v^\alpha : \quad v^\alpha v_\alpha \leq 0 \ . \tag{22.12}$$

This is known as the strong energy condition (this terminology is standard but confusing because the strong energy condition does not imply the weak energy condition - more on the relations among the various energy conditions below).

4. Dominant Energy Condition (DEC)

   Given the energy-momentum tensor $T_{\alpha\beta}$ and an observer with timelike (and future directed) 4-velocity t^α, the current

$$P^\alpha = -T^\alpha_\beta t^\beta \tag{22.13}$$

   represents the energy-momentum current density seen by that observer. The physically eminently reasonable dominant energy condition is then the statement that the speed of the flow of energy should not exceed the speed of light, i.e. that P^α should be causal (and future-directed),

$$P^\alpha = -T^\alpha_\beta t^\beta \quad \left\{ \begin{array}{c} \text{causal and future directed} \\ \text{for all timelike and future directed } \ t^\alpha \end{array} \right. \tag{22.14}$$

   Since t^α is itself timelike and future directed, this is equivalent to the 2 conditions

$$P^\alpha t_\alpha \leq 0 \quad \text{and} \quad P^\alpha P_\alpha \leq 0 \ . \tag{22.15}$$

   The 1st of these is more explicitly

$$P^\alpha t_\alpha \leq 0 \quad \Leftrightarrow \quad T_{\alpha\beta}t^\alpha t^\beta \geq 0 \tag{22.16}$$

   so that the dominant energy condition implies the weak energy condition, but requires additionally $P^\alpha P_\alpha \leq 0$.

It is clear from the above definitions that one has the implications

$$\text{(DEC)} \quad \Rightarrow \quad \text{(WEC)} \quad \Rightarrow \quad \text{(NEC)} \tag{22.17}$$

and

$$\text{(SEC)} \quad \Rightarrow \quad \text{(NEC)} \ , \tag{22.18}$$

and thus the NEC is the weakest of these energy conditions. However, neither does the SEC imply the WEC nor is there a simple relation between the SEC and the DEC.

It is good to keep in mind that none of these energy conditions are sacrosanct. While they are all satisfied in simple models (like that of a massless scalar field below), it is easy to construct reasonable physical models that violate any one of these energy conditions, either classically or at the quantum level (where e.g. negative Casimir vacuum energy densities can arise). Thus any result that is obtained on the basis of one of these energy conditions comes with a built-in caveat that it only applies to matter satisfying that energy condition. How plausible the assumption of a particular energy condition is depends on the specific context.

To see what these energy conditions require concretely, it is useful to look at some examples:

1. Free Massless Scalar Field

   In this case, the energy-momentum tensor is

   $$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}g_{\alpha\beta}(\partial\phi)^2 \quad , \tag{22.19}$$

   with

   $$(\partial\phi)^2 \equiv g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi \tag{22.20}$$

   and

   $$\bar{T}_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi \quad . \tag{22.21}$$

   - We see that the NEC is automatically satisfied,

     $$T_{\alpha\beta}\ell^\alpha\ell^\beta = (\ell^\alpha\partial_\alpha\phi)^2 \equiv (\ell.\partial\phi)^2 \geq 0 \quad . \tag{22.22}$$

   - For the WEC, we need to look at $T_{\alpha\beta}t^\alpha t^\beta$. Normalising t^α to $t^\alpha t_\alpha = -1$, we thus need to check non-negativity of

     $$T_{\alpha\beta}t^\alpha t^\beta = (t.\partial\phi)^2 + \tfrac{1}{2}(\partial\phi)^2 \quad . \tag{22.23}$$

     Here the 1st term is manifestly non-negative but the 2nd term is not. In order to disentangle this, consider the covector

     $$s_\alpha = \partial_\alpha\phi + t_\alpha(t.\partial\phi) \quad . \tag{22.24}$$

     It has the property

     $$t^\alpha s_\alpha = t.\partial\phi - t.\partial\phi = 0 \quad , \tag{22.25}$$

     and thus s_α is the projection of $\partial_\alpha\phi$ into the directions orthogonal to t^α. Since t^α is timelike, s^α is spacelike,

     $$s^\alpha s_\alpha > 0 \tag{22.26}$$

     (unless $s^\alpha = 0$). Moreover, explicitly the norm is

     $$s^\alpha s_\alpha = (\partial\phi)^2 + (t.\partial\phi)^2 \quad , \tag{22.27}$$

and therefore we can write $T_{\alpha\beta}t^\alpha t^\beta$ as a sum of non-negative terms,

$$T_{\alpha\beta}t^\alpha t^\beta = \tfrac{1}{2}(t.\partial\phi)^2 + \tfrac{1}{2}s^\alpha s_\alpha \geq 0 \ . \tag{22.28}$$

Therefore the WEC is satisfied.

- The SEC is automatically satisfied as well,

$$\bar{T}_{\alpha\beta}v^\alpha v^\beta = (v.\partial\phi)^2 \geq 0 \ . \tag{22.29}$$

- Finally, for the DEC we need to consider

$$P_\alpha = -\partial_\alpha\phi t.\partial\phi + \tfrac{1}{2}t_\alpha(\partial\phi)^2 \ . \tag{22.30}$$

We already checked the WEC part

$$-P_\alpha t^\alpha = (t.\partial\phi)^2 + \tfrac{1}{2}(\partial\phi)^2 \geq 0 \ , \tag{22.31}$$

and it remains to determine the norm of P^α, which is easily seen to be

$$P^\alpha P_\alpha = -\tfrac{1}{4}((\partial\phi)^2)^2 \leq 0 \ . \tag{22.32}$$

Therefore the DEC is satisfied.

Thus reassuringly a free massless scalar field satisfies all the 4 energy conditions.

2. Interacting Scalar Field

When we add a potential $V(\phi)$, the energy momentum tensor is modified to

$$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}g_{\alpha\beta}(\partial\phi)^2 - g_{\alpha\beta}V(\phi) \ , \tag{22.33}$$

and

$$\bar{T}_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi + g_{\alpha\beta}V(\phi) \ . \tag{22.34}$$

The NEC is clearly unaffected by the addition of this potential term to the energy-momentum tensor. For the remaining energy conditions, there is no need to go through the detailed analysis again. It is clear that a potential that is too negative can lead to a negative energy density and can thus threaten or violate the WEC and the DEC, while a potential that is too positive can threaten the SEC. To see e.g. the latter, note that

$$\bar{T}_{\alpha\beta}t^\alpha t^\beta = (t.\partial\phi)^2 - V(\phi) \ , \tag{22.35}$$

so that any static field configuration (in the sense of $t.\partial\phi = 0$) with a positive potential will violate the SEC. Thus even though the SEC is occasionally used, it needs to be kept in mind that even quite ordinary matter can violate this particular energy condition.

3. Cosmological Constant

   The effect of a positive versus a negative potential can be seen very explicitly by looking at such a term in isolation, e.g. in the form of a cosmological constant contribution to the energy-momentum tensor (19.51),

$$T^{\Lambda}_{\alpha\beta} = -\rho_\Lambda g_{\alpha\beta} \quad . \tag{22.36}$$

   with (19.50)

$$\rho_\Lambda = -p_\Lambda = \frac{\Lambda}{8\pi G_N} \tag{22.37}$$

   and

$$\bar{T}^{\Lambda}_{\alpha\beta} = +\rho_\Lambda g_{\alpha\beta} \quad . \tag{22.38}$$

   - For either sign of the cosmological constant this will (marginally) satisfy the NEC.
   - For the WEC we need to look at

$$T^{\Lambda}_{\alpha\beta} t^\alpha t^\beta = +\rho_\Lambda \quad . \tag{22.39}$$

     Thus a positive cosmological constant will satisfy the WEC, while a negative cosmological constant will violate it.
   - For the DEC, we see that

$$P_\alpha = -T^{\Lambda}_{\alpha\beta} t^\beta = +\rho_\Lambda t_\alpha \tag{22.40}$$

     is manifestly timelike, and will be future oriented iff $\rho_\Lambda > 0$, i.e. for a positive cosmological constant. A negative cosmological constant, on the other hand, violates both DEC conditions.
   - Finally, for the SEC the signs are reversed with respect to the WEC,

$$\bar{T}^{\Lambda}_{\alpha\beta} t^\alpha t^\beta = -\rho_\Lambda \quad , \tag{22.41}$$

     and therefore a positive cosmological constant violates the SEC while a negative cosmological constant satisfies it.

4. Perfect Fluid

   Another useful and instructive example to look at is the energy-momentum tensor of a perfect fluid (7.70),

$$T_{\alpha\beta} = (\rho + p) u_\alpha u_\beta + p g_{\alpha\beta} \quad , \tag{22.42}$$

   with $g_{\alpha\beta} u^\alpha u^\beta = -1$, and with

$$T^\alpha_{\ \alpha} = -\rho + 3p \quad , \tag{22.43}$$

so that

$$\bar{T}_{\alpha\beta} = (\rho + p)u_\alpha u_\beta + \tfrac{1}{2}(\rho - p)g_{\alpha\beta} \quad . \tag{22.44}$$

This energy-momentum tensor depends on the timelike velocity field u^α, $u^\alpha u_\alpha = -1$, not to be confused with the arbitrary timelike (and future oriented) vector field t^α (without loss of generality also with $t^\alpha t_\alpha = -1$) featuring in the energy conditions. In particular, for the WEC, say, it is not sufficient to just look at

$$T_{\alpha\beta}u^\alpha u^\beta = \rho \quad . \tag{22.45}$$

Rather, one needs to look at $T_{\alpha\beta}t^\alpha t^\beta$ for all timelike future directed t^α. In order to implement this concretely, and nevertheless make use of the presence of u^α in the energy-momentum tensor, let us assume without loss of generality that at the space-time point of interest (or by choosing a comoving coordinate system) u^α has the form

$$u^\beta = (1, 0, 0, 0) \quad . \tag{22.46}$$

We can then boost this vector with rapidity α to another timelike vector

$$t^\beta = (\cosh\alpha, \sinh\alpha, 0, 0) \tag{22.47}$$

or more generally to

$$t^\beta = \cosh\alpha u^\beta + \sinh\alpha n^\beta \tag{22.48}$$

where n^β is any spacelike unit vector orthogonal to u^β. Any future directed timelike unit vector can be written in this way (i.e. can be boosted back to u^β). The only quantity that will play a role in the following is the scalar product

$$u.t = u^\beta t_\beta = -\cosh\alpha < 0 \tag{22.49}$$

and its square

$$(u.t)^2 = \cosh^2\alpha \in [1, \infty) \quad , \tag{22.50}$$

with $\alpha = 0$ ($\cosh\alpha = 1$) corresponding to $t^\beta = u^\beta$, and $\alpha \to \infty$ corresponding to an infinite boost of u^α to some lightlike vector ℓ^α.

Equipped with this, we can now again analyse the energy conditions in turn:

- The NEC (for which none of the above gymnastics are required) requires

  $$T_{\alpha\beta}\ell^\alpha\ell^\beta = (\rho + p)(u.\ell)^2 \geq 0 \quad , \tag{22.51}$$

  and therefore

  $$\rho + p \geq 0 \quad . \tag{22.52}$$

  In particular, the energy-density ρ need not be positive provided that it is compensated by a sufficiently large and positive pressure.

- For the WEC we need to look at

$$T_{\alpha\beta}t^\alpha t^\beta = (\rho+p)(u.t)^2 - p \quad . \tag{22.53}$$

This expression is linear in $(u.t)^2$. Thus to check (or impose) positivity (or non-negativity, to be precise), we just need to check (or impose) it at the 2 endpoints of the $(u.t)^2$-interval $[1,\infty)$. For $(u.t)^2 = 1$ and $(u.t)^2 \to \infty$ we obtain respectively

$$\rho + p - p = \rho \geq 0 \quad , \quad \rho + p \geq 0 \quad . \tag{22.54}$$

When these conditions are satisfied, the WEC is also satisfied for any other choice of t. Therefore the WEC is equivalent to the NEC with the additional requirement that $\rho \geq 0$.

- For the DEC, in addition to the WEC we need to impose the condition that

$$P^\alpha = -T^\alpha_{\ \beta}t^\beta = -(\rho+p)(u.t)u^\alpha - pt^\alpha \tag{22.55}$$

is non-spacelike. Calculating the norm of P^α, one finds, using $u^2 = t^2 = -1$,

$$P^\alpha P_\alpha = -(\rho+p)^2(u.t)^2 + 2p(\rho+p)(u.t)^2 - p^2 = (u.t)^2(p^2-\rho^2) - p^2 \quad . \tag{22.56}$$

Again this is linear in $(u.t)^2$, so it is sufficient to look at the condition $P^\alpha P_\alpha \leq 0$ in the 2 limits. For $(u.t)^2 = 1$ one finds the empty condition $\rho^2 \geq 0$, while for $(u.t)^2 \to \infty$ one finds

$$\rho^2 > p^2 \quad \Leftrightarrow \quad |\rho| > |p| \quad . \tag{22.57}$$

Together with the conditions arising from the WEC, this can simply be written as the single requirement

$$\rho \geq |p| \quad . \tag{22.58}$$

- Finally, for the SEC we need to look at

$$\bar{T}_{\alpha\beta}t^\alpha t^\beta = (\rho+p)(u.t)^2 - \tfrac{1}{2}(\rho - p) \quad . \tag{22.59}$$

By the same argument as above, the SEC is then equivalent to

$$\rho + p \geq 0 \quad , \quad \rho + p - \tfrac{1}{2}(\rho - p) \geq 0 \tag{22.60}$$

or

$$\rho + p \geq 0 \quad , \quad \rho + 3p \geq 0 \quad . \tag{22.61}$$

We will make use of these results in the discussion of cosmology later on, in particular in sections 35 and 36.

22.2 Canonical vs Covariant Energy-Momentum Tensor

We will generalise the analysis of section 20.6 and look at some of the implications of general covariance of the matter action when taking into account boundary terms. This turns out so be surprisingly rewarding - surprising because of the lack of obvious significance (not to be confused with “obvious lack of significance”, their significance is explained and illustrated e.g. in the references given in footnote 39 of section 20.6) of the identically conserved Noether currents obtained from the gravitational Einstein-Hilbert action in this case.

As we will see, this will lead us to a relation between the covariant energy-momentum tensor (as defined here, via the metric-variation of the action) and the canonical energy-momentum tensor (deduced from the translation-invariance of Poincaré-invariant field theories in Minkowski space via Noether’s theorem).

In order to keep things notationally manageable (and, admitttedly, in order to be able to side-step a few technical fine-points), we will at first just consider the usual minimally coupled action for a scalar field. I will then discuss the general case (without, however, going through the entire calculation).

We thus have the minimally coupled matter action

$$S_M[\phi, g_{\alpha\beta}] = \int \sqrt{g} d^4x \; L_M(\phi, \partial_\nu\phi; g_{\alpha\beta}) \quad . \tag{22.62}$$

On general grounds we know that, under the transformation $\delta_\xi = L_\xi$ the action transforms as (9.69)

$$\delta_\xi \int \sqrt{g} d^4x \; L_M = \int \sqrt{g} d^4x \; \nabla_\mu(\xi^\mu L_M) \quad . \tag{22.63}$$

On the other hand, explicitly performing the variation, one finds

$$\delta_\xi S_M = \int \sqrt{g} d^4x \; \left[\frac{\delta_\xi\sqrt{g}}{\sqrt{g}} L_M + \frac{\partial L_M}{\partial g^{\alpha\beta}}\delta_\xi g^{\alpha\beta} + \frac{\partial L_M}{\partial\phi}\delta_\xi\phi + \frac{\partial L_M}{\partial(\partial_\nu\phi)}\delta_\xi(\partial_\nu\phi)\right] \tag{22.64}$$

Using the formula (9.30) for the Lie derivative of a covector, one sees that

$$\begin{aligned} \delta_\xi\partial_\nu\phi &= L_\xi\partial_\nu\phi = \xi^\mu\partial_\mu\partial_\nu\phi + (\partial_\nu\xi^\mu)\partial_\mu\phi \\ &= \partial_\nu(\xi^\mu\partial_\mu\phi) = \partial_\nu L_\xi\phi = \partial_\nu\delta_\xi\phi \end{aligned} \tag{22.65}$$

(this is one of the simplifications brought about by considering only scalars). Then, performing the usual integration by parts, keeping track of the boundary term, and combining (22.63) and (22.64), one finds

$$\int \sqrt{g} d^4x \; T_{\alpha\beta}\nabla^\alpha\xi^\beta + \int \sqrt{g} d^4x \; \frac{\delta L_M}{\delta\phi}\delta_\xi\phi = \int \sqrt{g} d^4x \; \nabla_\alpha(\Theta^\alpha_{\;\beta}\xi^\beta) \quad , \tag{22.66}$$

where

$$\Theta^\alpha_{\;\beta} = \delta^\alpha_{\;\beta} L_M - \frac{\partial L_M}{\partial(\partial_\alpha\phi)}\partial_\beta\phi \tag{22.67}$$

is the usual canonical Noether energy-momentum tensor, and

$$\frac{\delta L_M}{\delta\phi} = \frac{\partial L_M}{\partial\phi} - \nabla_\mu \frac{\partial L_M}{\partial(\partial_\mu\phi)} \tag{22.68}$$

is the Euler-Lagrange variational derivative. Since this has to hold for all ξ^α, we deduce

$$T_{\alpha\beta}\nabla^\alpha\xi^\beta + \frac{\delta L_M}{\delta\phi}(\partial_\beta\phi)\,\xi^\beta = \nabla_\alpha(\Theta^\alpha_{\ \beta}\xi^\beta) \ \ , \tag{22.69}$$

or

$$(T_{\alpha\beta} - \Theta_{\alpha\beta})\nabla^\alpha\xi^\beta = (\nabla_\alpha\Theta^\alpha_{\ \beta} - \frac{\delta L_M}{\delta\phi}\partial_\beta\phi)\xi^\beta \ \ . \tag{22.70}$$

If this is to be valid for arbitrary ξ^β, the coefficients of ξ^β and $\nabla^\alpha\xi^\beta$ have to vanish separately (at the origin of an inertial coordinate system this would be the coefficients of the obviously functionally independent functions ξ^β and $\partial_\alpha\xi^\beta$). This has the following implications:

1. First of all, we learn that (in the case of scalar fields) the canonical and covariant energy-momentum tensors agree (off-shell),

$$T_{\alpha\beta} = \Theta_{\alpha\beta} \ \ . \tag{22.71}$$

   This is of course something that we already checked explicitly in section 7.6, but it is reassuring to see this drop out of the general formalism as well.

2. We also learn that the latter (and hence the former) is conserved if the equations of motion of the matter fields are satisfied,

$$\frac{\delta L_M}{\delta\phi} = 0 \quad \Rightarrow \quad \nabla_\alpha\Theta^\alpha_{\ \beta} = \nabla_\alpha T^\alpha_{\ \beta} = 0 \ \ . \tag{22.72}$$

   This could have also alternatively been deduced from integrating (22.69) over a domain (with the ξ^α chosen to vanish on the boundary), and integrating by parts - this is just a repetition of the argument that previously led us to (20.99).

3. We can also see directly from (22.69) that for ξ^α a Killing vector, the Noether current

$$J^\alpha_N(\xi) = \Theta^\alpha_{\ \beta}\xi^\beta \tag{22.73}$$

   is on-shell conserved,

$$\nabla^{(\alpha}\xi^{\beta)} = 0 \quad \text{and} \quad \frac{\delta L_M}{\delta\phi} = 0 \quad \Rightarrow \quad \nabla_\alpha J^\alpha_N(\xi) = 0 \ \ . \tag{22.74}$$

   These are the conserved currents previously discussed in section 10.1.

As mentioned above, the case of scalar fields has some simplifying and non-generic features (exemplified e.g. already by the fact that no Belinfante improvement is required in this case). In general, i.e. when one does not restrict to scalar fields, the above chain of reasoning leads to the (on-shell) identification of the covariant energy-momentum tensor $T_{\alpha\beta}$ with a suitable covariantisation $\hat{\Theta}_{\alpha\beta}$ of the Belinfante symmetric improvement $\hat{\Theta}_{ab}$ of the canonical Noether energy-momentum tensor Θ_{ab}.

That the above argument naturally gives rise to the improved energy-momentum tensor, and hence to the on-shell identification of the covariant and improved tensors, can be seen quite explicitly in the case of Maxwell theory (and this can easily be extended to other theories).

In general, for any matter action L_M we have (we now remove all the, really quite unnecessary, integrals from the previous argument)

$$L_\xi(\sqrt{g}L_M) = \sqrt{g}\nabla_\alpha(\xi^\alpha L_M) = \sqrt{g}\nabla_\alpha[(\delta^a_{\ \beta} L_M)\xi^\beta)] \quad , \tag{22.75}$$

because $\sqrt{g}L_M$ is a scalar density. On the other hand, by explicitly acting with the Lie derivative on the metric and the other fields the Lagrangian density depends on, and splitting

$$L_\xi = (L_\xi)_g + (L_\xi)_M \tag{22.76}$$

into its action on the metric and the matter fields, we can write this as

$$\begin{aligned} L_\xi(\sqrt{g}L_M) &= (L_\xi)_g(\sqrt{g}L_M) + \sqrt{g}(L_\xi)_M L_M \\ &= -\tfrac{1}{2}\sqrt{g}T_{\alpha\beta}L_\xi g^{\alpha\beta} + \sqrt{g}(L_\xi)_M L_M \\ &= \sqrt{g}T_{\alpha\beta}\nabla^\alpha\xi^\beta + \sqrt{g}(L_\xi)_M L_M \quad . \end{aligned} \tag{22.77}$$

Now let us specialise to Maxwell theory. In that case, we have

$$L_M = -\tfrac{1}{4}F^{\alpha\beta}F_{\alpha\beta} \quad \Rightarrow \quad (L_\xi)_M L_M = -\tfrac{1}{2}F^{\alpha\beta}L_\xi F_{\alpha\beta} \quad . \tag{22.78}$$

Using the explicit covariant expression (9.31) for the Lie derivative on a $(0,2)$-tensor, we can write this as

$$(L_\xi)_M L_M = -\tfrac{1}{2}F^{\alpha\beta}[\xi^\gamma\nabla_\gamma F_{\alpha\beta} + 2(\nabla_\alpha\xi^\gamma)F_{\gamma\beta}] \quad . \tag{22.79}$$

"Integrating by parts" the 2nd term, this can be written as

$$\begin{aligned} (L_\xi)_M L_M = &-\tfrac{1}{2}F^{\alpha\beta}\xi^\gamma[\nabla_\gamma F_{\alpha\beta} + \nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha}] \\ &+ (\nabla_\alpha F^{\alpha\beta})F_{\gamma\beta}\xi^\gamma - \nabla_\alpha(F^{\alpha\gamma}F_{\beta\gamma}\xi^\beta) \quad . \end{aligned} \tag{22.80}$$

Simply by transferring the last term to the right-hand side of (22.75), we arrive at

$$\begin{aligned} &T_{\alpha\beta}\nabla^\alpha\xi^\beta + \tfrac{1}{2}F^{\alpha\beta}\xi^\gamma[\nabla_\gamma F_{\alpha\beta} + \nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha}] + (\nabla_\alpha F^{\alpha\beta})F_{\gamma\beta}\xi^\gamma \\ =&\nabla_\alpha[F^{\alpha\gamma}F_{\beta\gamma} - \tfrac{1}{4}\delta^\alpha_\beta F^2)\xi^\beta] \quad . \end{aligned} \tag{22.81}$$

This is the Maxwell counterpart of (22.69) and we see that where we had the covariantised canonical energy-momentum tensor $\Theta^\alpha_{\ \beta}$ we now have automatically obtained the improved gauge-invariant and symmetric expression

$$T_{\alpha\beta} = F^{\alpha\gamma}F_{\beta\gamma} - \tfrac{1}{4}\delta^\alpha_\beta F^2 \quad . \tag{22.82}$$

We also see very explicitly from this result that conservation of $T_{\alpha\beta}$ requires both sets of Maxwell equations (in particular also the Bianchi idenitity, something one would not have anticipated from (22.69)).

If, instead, one had written the last term in (22.80) as

$$\nabla_\alpha(F^{\alpha\gamma}F_{\beta\gamma}\xi^\beta) = \nabla_\alpha(F^{\alpha\gamma}(\nabla_\beta A_\gamma)\xi^\beta) - \nabla_\alpha(F^{\alpha\gamma}(\nabla_\gamma A_\beta)\xi^\beta) \tag{22.83}$$

and had just moved the first of these over to the other side, then that term would have combined with $\delta^\alpha_\beta L_M$ to give the covariantised (but not gauge-invariant) canonical energy-momentum tensor

$$\Theta^\alpha_\beta = F^{\alpha\gamma}\nabla_\beta A_\gamma - \tfrac{1}{4}\delta^\alpha_\beta F^2 \quad , \tag{22.84}$$

which, as already discussed in section 7.5 is not covariantly conserved. All the other terms, including some coming from the Bianchi, identities would then usually be attributed as covariantisations of Belinfante improvement terms (and other cosmetic correction terms arising e.g. from the fact that covariant and Lie derivatives do not commute, $[L_\xi, \nabla_\alpha]A_\beta \neq 0$ etc.).

However, it is clear from the above derivation that this is an unnatural way of splitting up a perfectly reasonable and gauge-invariant equation like (22.80).

Moreover, the rationale underlying the entire Belinfante symmetrisation procedure is not particularly compelling or natural in the case of non-trivial curved space-times (because of its origins in the Poincaré invariance and angular momentum conservation of the Minkowskian field theory, as summarised in section 7.4), and in general one encounters the following features (or, rather, bugs):

- The covariantisation of the canonical Noether energy-momentum tensor, say something like

$$\Theta^\mu{}_\nu = -\frac{\partial L_M}{\partial(\nabla_\mu\phi)}\nabla_\nu\phi + \delta^\mu_\nu L_M \quad . \tag{22.85}$$

  is typically not covariantly on-shell conserved all by itself,

$$\nabla_\mu\Theta^\mu{}_\nu \neq 0 \quad \text{on-shell} \tag{22.86}$$

  because of curvature terms, i.e. because one would have to be able to commute covariant derivatives on the fields in order to establish that the tensor is covariantly conserved, and this is not possible for field of spin higher than zero. Indeed, by explicitly calculating the covariant divergence, one finds

$$\nabla_\mu\Theta^\mu{}_\nu = \frac{\delta L_M}{\delta\phi}(\nabla_\nu\phi) + \frac{\partial L_M}{\partial(\nabla_\mu\phi)}[\nabla_\nu, \nabla_\mu]\phi \quad . \tag{22.87}$$

- Likewise the covariantisation of the improvement term, in particular with

$$\partial_\lambda\Psi^{\lambda\mu\nu} \quad \rightarrow \quad \nabla_\lambda\Psi^{\lambda\mu\nu} \quad \text{with} \quad \Psi^{\lambda\mu\nu} = -\Psi^{\mu\lambda\nu} \quad , \tag{22.88}$$

is not identically conserved,

$$\nabla_\mu \nabla_\lambda \Psi^{\lambda\mu\nu} = \tfrac{1}{2}[\nabla_\mu, \nabla_\lambda]\Psi^{\lambda\mu\nu} = \tfrac{1}{2} R^\nu{}_{\rho\mu\lambda} \Psi^{\lambda\mu\rho} \neq 0 \quad . \tag{22.89}$$

- Thus there seems to be no good reason in the first place to add this term to the non-conserved $\Theta_{\mu\nu}$. Nevertheless, it turns out that the sum of these two contributions,

$$\hat{\Theta}^{\mu\nu} = \Theta^{\mu\nu} + \nabla_\lambda \Psi^{\lambda\mu\nu} \quad , \tag{22.90}$$

is indeed (somewhat miraculously) on-shell conserved and symmetric,

$$\hat{\Theta}^{\mu\nu} = \hat{\Theta}^{\nu\mu} \quad \text{and} \quad \nabla_\mu \hat{\Theta}^{\mu\nu} = 0 \quad . \tag{22.91}$$

- This can be demystified somewhat by deriving this result from Noether's theorem applied to coordinate transformations (generalising the calculation done for scalar fields at the beginning of this section by taking into account the non-trivial transformation behaviour of higher-rank tensor fields under coordinate transformations and/or extending the usual Belinfante argument from Lorentz transformations to general coordinate transformations).[52]

- Ultimately, however, what this proof shows is that these properties hold for $\hat{\Theta}_{\mu\nu}$ because this tensor is equal to the covariant energy-momentum tensor on-shell,

$$\hat{\Theta}_{\mu\nu} = T_{\mu\nu} \quad \text{on-shell} \quad , \tag{22.92}$$

the latter being (on-shell) covariantly conserved and (off-shell) symmetric by contruction, and requiring absolutely no Belinfante-like gymnastics and improvement terms for its construction.

- In particular, upon restriction to a Poincaré-invariant field theory in Minkowski space, by undoing the minimal coupling, i.e. by sending $x^\alpha \to \xi^a$, $g_{\alpha\beta} \to \eta_{ab}$ etc. at the end of the calculation, one deduces that the resulting energy-momentum tensor T_{ab} (7.113)

$$T_{ab} := (T_{\alpha\beta})|_{x^\alpha \to \xi^a, g_{\alpha\beta} \to \eta_{ab}} \quad . \tag{22.93}$$

agress with the Belinfante-improved canonical Noether energy-momentum tensor $\hat{\Theta}_{ab}$ (7.114),

$$\hat{\Theta}_{ab} = T_{ab} \quad \text{on-shell} \quad . \tag{22.94}$$

[52] The literature on the Noether theorem and general covariance is littered with confused or at best confusing articles. Recent reasonably clear and to the point discussions of these and / or closely related issues can be found in R. Gamboa Saravi, *On the energy-momentum tensor*, `arXiv:math-ph/0306020`, J. Pons, *Noether symmetries, energy-momentum tensors and conformal invariance in classical field theory*, `arXiv:0902.4871 [hep-th]`, and M. Holman, *Generalized Noether Theorems for Field Theories Formulated in Minkowski Spacetime*, `arXiv:1009.1803 [gr-qc]`.

> The energy-momentum tensor T_{ab} defined in this way is automatically off-shell symmetric, on-shell conserved (and, in particular, automatically takes into account the orbital and spin contributions to the total conserved angular momentum of a Poincaré-invariant field theory).

For me, the upshot of this long discussion is (but please draw your own conclusions) that the covariant definition of the energy-momentum tensor appears to be superior, both conceptually and calculationally, to other definitions, because it is more general, more concise and to the point, and easier to work with. It should therefore also be the definition of choice even if one is just interested in Poincaré-invariant field theories in Minkowski space, in particular as it completely side-steps the issue of having to contruct the Belinfante improvement terms to the canonical energy-momentum tensor.

Moreover, as already emphasised in section 7.6, regardless of this entire discussion, if one derives the Einstein equations from a variational principle then it is unavoidably in any case this covariant energy-momentum tensor that is the source term in the Einstein equations.

22.3 Energy-Momentum Tensor of a Conformally Coupled Scalar Field

In section 7.7 we had discussed matter actions that are invariant under Weyl transformations

$$g_{\mu\nu}(x) \to \bar{g}_{\mu\nu}(x) = \Omega(x)^2 g_{\mu\nu}(x) \tag{22.95}$$

of the metric alone (i.e. without also transforming the matter fields) and had shown that such actions lead to an (off-shell) traceless covariant energy-momentum tensor. It was also evident from that discussion that one would obtain an on-shell traceless energy-momentum tensor in theories that are invariant under a joint non-trivial Weyl rescaling of the metric and the matter fields. We have now accumulated everything that we need to discuss an example of this kind, namely the so-called *conformally coupled scalar field*.

We start off by considering the minimally coupled action of a free massless scalar field in D space-time dimensions,

$$S_0[\phi, g_{\alpha\beta}] = -\tfrac{1}{2} \int \sqrt{g} d^D x \; g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \tag{22.96}$$

This action is invariant under joint constant rescalings

$$g_{\mu\nu} \to \bar{g}_{\mu\nu} = \mathrm{e}^{2\lambda} g_{\mu\nu} \quad , \quad \phi \to \bar{\phi} = \mathrm{e}^{w_\phi \lambda} \phi \tag{22.97}$$

of the metric and the scalar field, provided that one chooses the *scaling weight* or *Weyl weight* of the scalar field ϕ to be

$$w_\phi = \frac{2-D}{2} \quad . \tag{22.98}$$

To see this, note that under this rescaling one has, for the gravitational part,

$$\sqrt{g}g^{\mu\nu} \to \mathrm{e}^{(D-2)\lambda}\sqrt{g}g^{\mu\nu} \quad , \tag{22.99}$$

while for the scalar field part one evidently has

$$\partial_\mu\phi\partial_\nu\phi \to \mathrm{e}^{2\lambda w_\phi}\partial_\mu\phi\partial_\nu\phi = \mathrm{e}^{(2-D)\lambda}\partial_\mu\phi\partial_\nu\phi \quad , \tag{22.100}$$

which establishes the invariance of the massless free action under this scaling for the choice of weight (22.98). We will return to this scale invariance, and its relation to the, perhaps more familiar, dilatation invariance of relativistic field theories, below (and just note for now, with apologies, that for a scalar field the scaling weight w_ϕ is related to what is called the scale dimension d_ϕ of a field ϕ in relativistic field theories by $w_\phi = -d_\phi$).

While the action (22.96) is invariant under constant rescalings with the above weights, as it stands the action is clearly not invariant under Weyl rescalings, i.e. space-time dependent rescalings of the form

$$\begin{aligned} g_{\mu\nu}(x) \to \bar{g}_{\mu\nu}(x) &= \Omega(x)^2 g_{\mu\nu}(x) \\ \phi(x) \to \bar{\phi}(x) &= \Omega(x)^{w_\phi}\phi(x) = \Omega(x)^{(2-D)/2}\phi(x) \quad , \end{aligned} \tag{22.101}$$

because one will invariably pick up derivatives of $\Omega(x)$ that are not cancelled by anything else (unless $w_\phi = 0$, i.e. $D = 2$, which brings us back to the case already discussed in section 7.7).

It is a remarkable fact, however, that for $D > 2$ invariance under (22.101) can be achieved by adding a non-minimally coupled mass term of the form $R\phi^2$ to the Lagrangian. Indeed, the action (cf. (8.129))

$$S_\xi[\phi, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g}d^Dx \; \left(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \xi R\phi^2\right) \tag{22.102}$$

turns out to be invariant under (22.101) for a specific choice of ξ, namely

$$\xi = \frac{D-2}{4(D-1)} \quad \Rightarrow \quad S_\xi[\Omega(x)^{(2-D)/2}\phi, \Omega(x)^2 g_{\alpha\beta}] = S_\xi[\phi, g_{\alpha\beta}] \quad . \tag{22.103}$$

In particular, for $D = 4$ one requires the peculiar value

$$D = 4: \qquad \xi = 1/6 \quad . \tag{22.104}$$

At first sight this invariance seems to be not only unlikely but also somewhat unpleasant to try to prove (or disprove), since it appears that one would have to work out how the curvature scalar of the Weyl-rescaled metric $\bar{g}_{\alpha\beta} = \Omega^2 g_{\alpha\beta}$ is related to that of $g_{\alpha\beta}$. While this is something that can be worked out with a steady hand, by first determining how the Christoffel symbols are related (cf. (28.7)), and then working out the relation between the curvature tensors etc. from there, this is no fun. Since I am not aware of a particularly efficient way to short-cut this calculation, and the final result is not even particularly illuminating, we will forego this here.[53] Just for the record, and for the sake of illustration, here are the resulting expressions for the Ricci tensor and the Ricci scalar:

$$
\begin{aligned}
\bar{R}_{\alpha\beta} &= R_{\alpha\beta} + (D-2)\left(2\Omega^{-2}\nabla_\alpha\Omega\nabla_\beta\Omega - \Omega^{-1}\nabla_\alpha\nabla_\beta\Omega\right) \\
&\qquad - g_{\alpha\beta}\left(\Omega^{-1}\Box_g\Omega + (D-3)\Omega^{-2}g^{\gamma\delta}\nabla_\gamma\Omega\nabla_\delta\Omega\right) \\
\bar{R} &= \Omega^{-2}R - 2(D-1)\Omega^{-3}\Box_g\Omega - (D-1)(D-4)\Omega^{-4}g^{\alpha\beta}\nabla_\alpha\Omega\nabla_\beta\Omega
\end{aligned}
\tag{22.105}
$$

The quantities on the left-hand side are calculated with respect to the metric $\bar{g}_{\alpha\beta}$. In particular,

$$
\bar{R} = \bar{g}^{\alpha\beta}\bar{R}_{\alpha\beta} = \Omega^{-2}g^{\alpha\beta}\bar{R}_{\alpha\beta} \quad , \tag{22.106}
$$

while those on the right-hand side are calculated with respect to the metric $g_{\alpha\beta}$ (and I have indicated this explicitly in the notation $\Box_g$ for the covariant Laplacian to make this slightly more manifest).

Since the Weyl tensor (the trace-free part of the Riemann tensor) is invariant under Weyl rescalings (cf. section 11.4),

$$
\bar{C}^{\alpha}_{\ \beta\gamma\delta} = C^{\alpha}_{\ \beta\gamma\delta} \quad \Leftrightarrow \quad \bar{C}_{\alpha\beta\gamma\delta} = \Omega^2 C_{\alpha\beta\gamma\delta} \quad , \tag{22.107}
$$

this is enough to reconstruct the transformation behaviour of the entire Riemann tensor.

However, since there are no global subtleties hiding in Weyl transformations, we will not need to make use of any of these equations. It is completely sufficient to check invariance of the action under infinitesimal Weyl transformations, i.e. transformations with Ω of the form

$$
\Omega(x) = 1 + \omega(x) \tag{22.108}
$$

and $\omega(x)$ infinitesimal, namely

$$
\delta g_{\mu\nu}(x) = 2\omega(x)g_{\mu\nu}(x) \quad , \quad \delta\phi(x) = \frac{2-D}{2}\omega(x)\phi(x) \quad , \tag{22.109}
$$

and this turns out to be much simpler. Before turning to this, note that the tranformation of the metric also implies that the volume element $\sqrt{g}$ transforms as

$$
\delta\sqrt{g} = D\omega\sqrt{g} \quad . \tag{22.110}
$$

[53] If, for whatever reason, you need these formulae and you do not want to work them out yourself, you can look them up e.g. in Appendix D of R. Wald, *General Relativity* or Appendix G of S. Carroll, *Spacetime and Geometry*, or Appendix E of T. Ortin, *Gravity and Strings* (see how nobody wants to clutter the body of the text with these formulae?), which, I am happy to report, all agree once differences in conventions have been accounted for.

In section 20.2 we had already worked out a formula for the variation of the curvature scalar under an *arbitrary* infinitesimal variation of the metric, namely (20.19)

$$\begin{aligned}\delta R &= (\delta g^{\mu\nu})R_{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} \\ &= (\delta g^{\mu\nu})R_{\mu\nu} + (\nabla^\mu\nabla^\nu - g^{\mu\nu}\Box)\delta g_{\mu\nu} \ ,\end{aligned} \tag{22.111}$$

so all we need to do is plug $\delta g_{\mu\nu} = 2\omega g_{\mu\nu}$ into this general equation, leading to

$$\begin{aligned}\delta R &= -2\omega R + (\nabla^\mu\nabla^\nu - g^{\mu\nu}\Box)2\omega g_{\mu\nu} \\ &= -2\omega R + (2-2D)\Box\omega \ .\end{aligned} \tag{22.112}$$

Note that this also follows from the expression for $\bar{R}$ in (22.105). Therefore, using

$$\delta(\sqrt{g}g^{\mu\nu}) = (D-2)\omega\sqrt{g}g^{\mu\nu} \tag{22.113}$$

the variation of the $\int R\phi^2$ term in the action (22.102) is (suppressing the d^Dx here and in the subsequent equations)

$$\begin{aligned}\delta\int\sqrt{g}\phi^2 R &= \int\sqrt{g}[2\omega\phi^2 R + \phi^2\delta R] \\ &= (2-2D)\int\sqrt{g}\;\phi^2\;\Box\omega \ .\end{aligned} \tag{22.114}$$

It is even more straightforward to work out how the first (standard kinetic) term of the action transforms under (22.109). Ignoring boundary terms arising from integrations by part one finds

$$\begin{aligned}\delta\int\sqrt{g}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi &= \int\sqrt{g}g^{\mu\nu}(D-2)[\omega\partial_\mu\phi\partial_\nu\phi - \partial_\mu\phi\partial_\nu(\omega\phi)] \\ &= \int\sqrt{g}g^{\mu\nu}(2-D)(\nabla_\nu\omega)(\nabla_\mu\phi)\phi \\ &= \frac{D-2}{2}\int\sqrt{g}\;\phi^2\;\Box\omega\end{aligned} \tag{22.115}$$

It is now evident that for a judicious choice of coupling constant ξ one can get (22.114) to cancel against (22.115). Specifically, one finds that

$$\delta\int\sqrt{g}(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \xi R\phi^2) = \left[\frac{D-2}{2} + \xi(2-2D)\right]\int\sqrt{g}\;\phi^2\;\Box\omega \ , \tag{22.116}$$

so that invariance of the action S_ξ (22.102) is achieved for

$$\frac{D-2}{2} + \xi(2-2D) = 0 \quad\Leftrightarrow\quad \xi = \frac{D-2}{4(D-1)} \ , \tag{22.117}$$

as claimed.

Remarks:

1. We can also write ξ in a somewhat more informative manner in terms of the Weyl (or scaling) weight w_ϕ of the scalar field (given in (22.109) or (22.98)) as

$$\xi = -\frac{w_\phi}{2(D-1)} \ . \tag{22.118}$$

2. Under the Weyl transformation (22.101) of the scalar field ϕ and the metric $g_{\mu\nu}$, a monomial of the scalar field ϕ^p transforms as

$$\phi^p \to \Omega^{p(2-D)/2}\phi^p \quad , \quad \sqrt{g}\phi^p \to \Omega^{D+p(2-D)/2}\sqrt{g}\phi^p \ . \tag{22.119}$$

Therefore this term ϕ^p can be added to the action as a potential $V(\phi)$ while preserving Weyl invariance if the condition

$$D + p(2-D)/2 = 0 \quad \Leftrightarrow \quad p = \frac{2D}{D-2} \tag{22.120}$$

is satisfied. Thus the action

$$S[\phi, g_{\alpha\beta}] = -\tfrac{1}{2}\int \sqrt{g}d^Dx\ \left(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \tfrac{D-2}{4(D-1)}R\phi^2 + \lambda\phi^{2D/(D-2)}\right) \tag{22.121}$$

is Weyl invariant. In particular, for $D=4$ one can add a quartic interaction term,

$$D = 4 \quad \Rightarrow \quad V(\phi) \sim \phi^4 \ , \tag{22.122}$$

and the two other integer solutions for p are $p=6$ for $D=3$ and $p=3$ for $D=6$ (cf. also the discussion around (22.154) below).

3. If one considers gravity coupled to a single scalar field in this conformally coupled Weyl-invariant manner, then the local gauge symmetry (22.101) is essentially enough to gauge fix the scalar field $\phi(x)$ to a constant (away from zeros of $\phi(x)$). Then the action (22.121) essentially reduces to the Einstein-Hilbert gravitational action (albeit with the wrong sign, the constant value of the scalar field being related to Newton's constant or the Planck length/scale) plus a cosmological constant term. One can therefore also attempt to interpret the Einstein-Hilbert + cosmological constant action as arising from the spontaneous breaking of the Weyl invariance of the original Weyl-invariant action (22.121) through the eppearance of this gravitational scale.[54]

Returning to more down-to-earth things, in order to complete this discussion we should now check that, in accordance with the remarks at the end of section 7.7, the covariant energy-momentum tensor derived from (22.102) for this value of ξ is traceless *on-shell*, where on-shell means that ϕ satisfies the equation of motion

$$(\Box - \xi R)\phi = 0 \tag{22.123}$$

following from the action (22.102). Not only is this a useful exercise and good to check; as we will see below, the explicit expression for the energy-momentum tensor is also quite interesting in its own right.

[54]See e.g. R. Kallosh, A. Linde, *Hidden Superconformal Symmetry of the Cosmological Evolution*, `arXiv:1311.3326 [hep-th]` for an exploration of these ideas in the context of cosmology and inflation.

Thus we need to determine the variation of the action with respect to the metric, for which we can again use (22.111). Integrating the term $\phi^2(\nabla^\mu\nabla^\nu - g^{\mu\nu}\Box)\delta g_{\mu\nu}$ by parts (twice) and expressing this in terms of $\delta g^{\mu\nu}$ rather than $\delta g_{\mu\nu}$ (which just leads to a change of sign), one finds that the energy-momentum tensor associated with the action S_ξ is

$$\begin{aligned} T^{\xi}_{\mu\nu} &= \partial_\mu\phi\partial_\nu\phi - \tfrac{1}{2}g_{\mu\nu}(g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi) + \xi\left((G_{\mu\nu} + g_{\mu\nu}\Box - \nabla_\mu\nabla_\nu)\phi^2\right) \\ &= T_{\mu\nu} + \xi\left((G_{\mu\nu} + g_{\mu\nu}\Box - \nabla_\mu\nabla_\nu)\phi^2\right) \end{aligned} \tag{22.124}$$

,

where the first (ξ-independent) part is just the standard (Noether $\Theta_{\mu\nu}$ = covariant $T_{\mu\nu}$) energy-momentum tensor of a massless scalar field and $G_{\mu\nu}$ in the second part is the Einstein tensor (evidently one of the terms arising from the metric variation of $\sqrt{g}R\phi^2$).

This expression is valid for any ξ, not just the conformally invariant value. When referring specifically to the conformally invariant theory with ξ given by (22.117), I will use the superscript (c) on $T_{\mu\nu}$, i.e.

$$T^{(c)}_{\mu\nu} = T^{\xi=(D-2)/4(D-1)}_{\mu\nu} \tag{22.125}$$

With

$$\Box\phi^2 = 2g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + 2\phi\Box\phi \tag{22.126}$$

one finds that the trace of the energy-momentum tensor can be written as

$$T^{\xi\mu}{}_{\mu} = 2(D-1)\left[\left(\xi - \frac{D-2}{4(D-1)}\right)g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + \xi\phi\left(\Box\phi - \frac{D-2}{4(D-1)}R\phi\right)\right] \quad . \tag{22.127}$$

Note that we only have a single parameter ξ to play with but that cooperatively both terms vanish (the first off-shell, the second on-shell) precisely when ξ is chosen to have the value (22.117),

$$\xi = \frac{D-2}{4(D-1)} \quad \Rightarrow \quad T^{(c)\mu}{}_{\mu} = 0 \quad \text{on-shell.} \tag{22.128}$$

While without higher knowledge this may appear to be a minor miracle at this point, we know on general grounds that this had to work out.

Thus we have been able to construct a symmetric, on-shell conserved and on-shell traceless energy-momentum tensor for the action of a massless scalar field in any dimension $D \geq 2$. As a consequence of this and the considerations in section 10.2, associated to any (conformal) Killing vector

$$\nabla_\mu C_\nu(x) + \nabla_\nu C_\mu(x) = 2\omega(x)g_{\mu\nu}(x) \tag{22.129}$$

we have a conserved current

$$J^\mu_C = T^{(c)\mu}{}_{\nu}C^\nu : \quad \nabla_\mu J^\mu_C = 0 \quad \text{on-shell} \quad . \tag{22.130}$$

In particular, the conformal invariance of this model in Minkowski space (in this sense) is implied by the Weyl invariance of the action in curved space.

To conclude this section, I just want to point out that the Einstein equations

$$G_{\mu\nu} = \kappa T^{\xi}_{\mu\nu} \qquad (\kappa = 8\pi G_N) \tag{22.131}$$

for such a scalar field non-minimally coupled to the scalar curvature can be (and occasionally are) written in a different way, involving a field-dependent "effective" coupling constant $\kappa^{eff} = \kappa^{eff}(\xi, \phi)$, and a corresponding "effective" energy-momentum tensor $T^{(eff)}_{\alpha\beta}$ which is different from, and should not be confused with, the energy-momentum tensor $T^{\xi}_{\mu\nu}$ given in (22.124). This comes about as follows:

1. Equations of Motion

   At the level of the equations of motion, one sees from the explicit expression (22.124) that the Einstein tensor appears both on the left- and on the right-hand sides of the Einstein equation (22.131),

$$\begin{aligned} G_{\mu\nu} &= \kappa \left[T_{\mu\nu} + \xi \left((G_{\mu\nu} + g_{\mu\nu}\Box - \nabla_\mu \nabla_\nu)\phi^2\right)\right] \\ &\equiv \kappa\xi\phi^2 G_{\mu\nu} + \kappa T^{eff}_{\mu\nu} \quad . \end{aligned} \tag{22.132}$$

   Combining the $G_{\mu\nu}$-terms, this equation can be written equivalently as

$$G_{\mu\nu} = \kappa^{eff} T^{eff}_{\mu\nu} \tag{22.133}$$

   with the effective field-dependent gravitational coupling constant

$$\kappa^{eff}(\xi, \phi) = \frac{\kappa}{1 - \kappa\xi\phi^2} \equiv 8\pi G^{eff}_N \quad . \tag{22.134}$$

2. Action

   At the level of the action, this can be understood (or could have been anticipated) by noting that the non-minimal coupling term $\sim \xi R\phi^2$ in the scalar field action (22.102) can also be regarded as modifying the gravitational coupling constant appearing in front of the Einstein-Hilbert action,

$$\frac{1}{2\kappa}\int \sqrt{g}R - \frac{1}{2}\int \sqrt{g}\xi\phi^2 R = \int \sqrt{g}\frac{1}{2\kappa^{eff}} R \quad . \tag{22.135}$$

   However, variation of this action with respect to the metric does *not* give rise to $G_{\mu\nu}/\kappa^{eff}$, and variation of the rest of the matter action (22.102), which is just the standard action for a scalar field, does *not* give rise to $T^{eff}_{\mu\nu}$ (which depends on ξ), so care is required when using these effective quantities. In particular, $T^{eff}_{\mu\nu}$ is *not* traceless for a conformally coupled scalar field and can therefore not be used in the construction of additional conserved currents.

One interesting property of the energy-momentum tensor (22.124) is that, in spite of the fact that its difference from the canonical energy-momentum tensor is due to the $R\phi^2$-term in the action (which thus vanishes in Minkowski space), it actually differs from the canonical energy-momentum tensor $\Theta_{\mu\nu}$ even when one specialises it to Minkowski space,

$$T^{(c)}_{ab} \equiv T^{(c)}_{\mu\nu}|_{g_{\mu\nu}\to\eta_{ab}} = \Theta_{ab} + \xi\,(\eta_{ab}\Box_\eta - \partial_a\partial_b)\,\phi^2 \quad , \tag{22.136}$$

the second term being the remnant of the (linearised) Ricci scalar. We note that by construction this energy-momentum tensor is symmetric, and on-shell traceless and conserved,

$$\partial^a T^{(c)}_{ab} = T^{(c)a}{}_a = 0 \qquad \text{on-shell} \quad . \tag{22.137}$$

Now, what is the significance of this energy-momentum tensor and the fact that it differs from the canonical energy-momentum tensor of a scalar field?

First of all, we oberve that the term proportional to ξ is an *improvement term* in the sense of (7.62) and (7.63), i.e. the energy-momentum tensor can be written as

$$T^{(c)}_{ab} = \Theta_{ab} + \partial^c\Psi_{cab} \quad , \tag{22.138}$$

where the second term is identically conserved,

$$\Psi_{cab} = -\Psi_{acb} = \xi(\eta_{ab}\partial_c - \eta_{cb}\partial_a)\phi^2 \quad \Rightarrow \quad \partial^a\partial^c\Psi_{cab} = 0 \quad . \tag{22.139}$$

While this is reassuring and legitimises the energy-momentum tensor as an energy-momentum tensor for the free massless scalar field, it does also raise the *cui bono* question, namely what is this good for and what exactly does this term actually improve? After all, not every "improvement term" in the technical sense above is necessarily also an improvement in the more colloquial sense of the word. In order to explain this, I need to digress a bit (but I will try to keep this short, relegating the details to an appendix to this section).

As I mentioned before, what I referred to as scale invariance above (the special case of Weyl invariance for a constant rescaling of the metric) is closely related to what is known as *dilatation invariance* (or also as scale invariance) for relativistic field theories in Minkowski space. There one considers scalings not of the (fixed) Minkowski metric but instead of the coordinates (this is the *dilatation*), accompanied by an appropriate rescaling of the fields. Thus one considers transformations of the form (x^a denote inertial Minkowski coordinates usually denoted ξ^a elsewhere in these notes)

$$\bar{x}^a = \mathrm{e}^{-\lambda}x^a \quad , \quad \bar{\phi}(\bar{x}) = \mathrm{e}^{d_\phi\lambda}\phi(x) = \mathrm{e}^{d_\phi\lambda}\phi(\mathrm{e}^{\lambda}\bar{x}) \quad , \tag{22.140}$$

where $d_\phi = d(\phi)$ is the *scale dimension* (or *dilatation weight*) of the field ϕ, while the constant Weyl rescaling operation (22.97) is of the form

$$g_{\mu\nu} \to \bar{g}_{\mu\nu} = \mathrm{e}^{2\lambda}g_{\mu\nu} \quad , \quad \phi \to \bar{\phi} = \mathrm{e}^{w_\phi\lambda}\phi \quad . \tag{22.141}$$

Remarks:

1. In accordance with the standard practice in field theory, the dilatation action on the coordinates is written as $\bar{x} = \exp(-\lambda)x$, so that the coordinates have dimension (-1) and the corresponding scale dimensions d_ϕ count mass dimensions. By contrast, what I have called the scaling (or Weyl) weight $w_\phi = w(\phi)$ of a field ϕ in (22.141) counts length dimensions.

2. The latter appears to me to be the natural thing to do in general relativity, while in field theory and particle physics it is equally natural to instinctively convert derivatives and inverse lengths into masses or energy (using $\hbar$), and the definition (22.140) respects that.

3. For a scalar field one finds that a necessary condition for dilatation invariance is that $d_\phi = -w_\phi$, with w_ϕ given in (22.98), but the relation between the Weyl weight and the scale dimension is different for derivatives of fields and/or higher rank tensors. We will see examples of this below.

4. Since the operation (22.140) is not particularly meaningful in a general curved space (as it depends on a choice of coordinates), in that context it is preferable to consider scalings of the metric rather than of the coordinates. In this sense (22.141) is the appropriate extension of (22.140) to curved spaces.

Now let us consider a (Poincaré-invariant) action of the general form

$$S = \int d^D x \; L(x) \quad . \tag{22.142}$$

Under dilatations the integration measure transforms with dimension $-D$. Thererfore, dilatation-invariance of the action is simply the requirement that the Lagrangian have dimension D,

$$\text{dilatation invariance:} \quad d(L) = D \quad . \tag{22.143}$$

When one has such a dilatation invariant action, by the standard procedure the Noether theorem provides one with an on-shell conserved dilatation Noether current j^a_N,

$$\partial_a j^a_N = 0 \qquad \text{on-shell} \quad . \tag{22.144}$$

On the other hand, we know on general grounds from the discussion in section 10 that if we have a traceless symmetric conserved energy-momentum tensor T_{ab}, like our $T^{(c)}_{ab}$ in (22.136), we can construct a conserved current J^a_D (10.18) associated to the Killing vector $D = x^a\partial_a$ generating dilatations in Minkowski space, namely

$$J^a_D = T^a_{\ b} D^b = T^a_{\ b} x^b \quad \Rightarrow \quad \partial_a J^a_D = T^a_{\ a} = 0 \qquad \text{on-shell} \quad . \tag{22.145}$$

In passing we note that symmetry of T_{ab} is not strictly necessary for this conclusion since $\partial_b D_a = \eta_{ab}$ is symmetric (or $D_a = \partial_a(x^2)/2$ is a gradient vector). However, if we have a symmetric, on-shell tracefree and on-shell conserved energy-momentum tensor we can construct not only a conserved dilatation current but in fact conserved currents for the entire menagerie of generators of the conformal group (see section 10.3), in particular therefore also for the generators $C^{(m)}$ of special conformal transformations.

Thus the question if the dilatation current can be written in the form (22.145) is strictly related to the question (of interest in field theory) if dilatation invariance extends to full-fledged conformal invariance.

Now in the example at hand, of a free scalar field, it turns out that the Noether current is not of the form (22.145) for $D > 2$ with respect to the canonical Noether energy-momentum tensor Θ_{ab},

$$D > 2 : \quad j_N^a \neq \Theta^a{}_b x^b \tag{22.146}$$

(the precise form of j_N^a is derived below - see (22.171)). Indeed, this could hardly be otherwise as the Noether current is conserved by construction while we know from section 7.7 that Θ_{ab} is not traceless unless $D = 2$.

We can now (finally) return to the issue raised at the beginning of this section regarding the significance of the "improvement term" in the energy-momentum tensor (22.136)

$$T^{(c)}_{ab} = \Theta_{ab} + \frac{d_\phi}{2(D-1)} \left(\eta_{ab}\Box_\eta - \partial_a\partial_b\right)\phi^2 \tag{22.147}$$

for a free masless scalar field in D dimensions. Namely, it turns out that using this improved energy-momentum tensor the Noether current *can* be written in the form (22.145), modulo an identically conserved total derivative term (that does not contribute to charge integrals and may as well be neglected). Concretely one finds that (off-shell)

$$J_D^a - j_N^a = T^{(c)a}{}_b x^b - j_N^a = \partial_b \Delta^{ab} \quad , \tag{22.148}$$

where

$$\Delta^{ab} = -\Delta^{ba} = \frac{d_\phi}{2(D-1)}(x^a\partial^b - x^b\partial^a)\phi^2 \quad . \tag{22.149}$$

A proof of this assertion will be provided at the end of this section.

Thus what we have seen is that this improved energy-momentum tensor and its associated dilatation and conformal currents arise directly from the covariant energy-momentum tensor of a conformally coupled (i.e. Weyl invariant) scalar field in curved space-time. It is pleasing to see that also this improvement can be understood (and derived) directly from the gravitationally coupled matter action.

This gravitational perspective provides

- a more systematic way of finding the required improvement term,
- an understanding of why the improvement terms has the precise form it has (in the present case it can be traced back to the structure of the Ricci scalar)

- and a different perspective on the relation between scale and conformal invariance.

More generally, the question under which conditions one can construct an improvement term for the energy-momentum tensor such that the dilatation Noether current has the form (22.145) has been analysed in detail first by Callan, Coleman and Jackiw, and the improved energy-momentum tensor (when it exists) is known as the *Callan - Coleman - Jackiw* or CCJ tensor.[55]

In the remainder of this section, for the sake of completeness I will give a self-contained proof of the assertion (22.148), which really boils down to determining the Noether current for dilatations. It is useful, however, to start with a slightly more detailed discussion of two prototypical examples of non-trivially scale-invariant theories.

As examples of dilatation-invariant theories we consider (of course) a free massless scalar field (this is, after all, the example we want to construct the Noether current for), but we will also briefly consider Maxwell theory in D dimensions.

<u>EXAMPLES:</u>

1. Free Massless Scalar Field in D Dimensions

   The scalar field action is

$$S_0[\phi] = -\tfrac{1}{2}\int d^D x\, \eta^{ab}\partial_a\phi\partial_b\phi \ \ , \tag{22.150}$$

   In order to achieve dilatation invariance of the scalar field action, we need the Lagrangian to have dimension D. Thus we need to associate dimension $D/2$ to $\partial_a\phi$. Evidently, if ϕ has dimension $d(\phi)$, then its derivative $\partial_a\phi$ has dimension

$$d(\partial_a\phi) = d(\phi) + 1 \ \ . \tag{22.151}$$

   By contrast, the Weyl weight of a field under constant rescalings of the metric has nothing to do with the coordinates, and therefore

$$w(\partial_\alpha\phi) = w(\phi) \ \ . \tag{22.152}$$

   In any case, from the requirement $d(\partial_a\phi) = D/2$ we deduce

$$d(\partial_a\phi) = D/2 \quad \Rightarrow \quad d(\phi) = D/2 - 1 = (D-2)/2 = -w(\phi) \ \ . \tag{22.153}$$

[55] C. Callan, S. Coleman, R. Jackiw, *A new improved energy-momentum tensor*, Ann. Phys. 59 (1970) 42. See e.g. S. Coleman, R. Jackiw, *Why dilatation generators do not generate dilatations*, Ann. Phys. 67 (1971) 552 for a detailed analysis, and section 2.4 of T. Ortin, *Gravity and Strings* for a modern and gravitational prespective (but note that his improved energy-momentum tensor (2.74) is in general *not* symmetric, so that the construction of conserved conformal generators fails).

Thus, by postulating that the scalar field has this dimension, one arrives at a dilatation-invariant action.

REMARKS:

(a) In particular, for $D = 4$ this results in the "canonical" dimension $d(\phi) = 1$ of a scalar field (and, more generally, as we will see below, of any bosonic field with a standard action quadratic in first derivatives of the fields).

(b) It is evident that adding a mass term $\sim \phi^2$ or a generic potential $V(\phi)$ of the field to the action (22.96) or (22.150) violates this scale (or dilatation) invariance. Indeed, for a potential terms $\sim \phi^p$ to be scale-invariant one needs $p = 2D/(D-2)$ (by analogous reasoning to that leading to (22.120), but now employing dilatations rather than Weyl transformations), i.e.

$$V(\phi) \sim \phi^{2D/(D-2)} \quad . \tag{22.154}$$

Only for $D = 3, 4, 6$ is this an integer power of ϕ, namely

$$\begin{aligned} D = 3 \quad &\Rightarrow \quad V(\phi) \sim \phi^6 \\ D = 4 \quad &\Rightarrow \quad V(\phi) \sim \phi^4 \\ D = 6 \quad &\Rightarrow \quad V(\phi) \sim \phi^3 \quad . \end{aligned} \tag{22.155}$$

Thus e.g. a quartic coupling would be allowed in $D = 4$ and a sextic coupling in $D = 3$.

Precisely such interaction terms actually do appear (and are required) e.g. in certain superconformal gauge theories in these dimensions, but this is not even remotely our concern here.

2. Maxwell Theory in D Dimensions

The action of Maxwell theory in any dimension D is

$$S_0[A_a] = -\tfrac{1}{4} \int d^D x\, \eta^{ac}\eta^{bd} F_{ab} F_{cd} \quad . \tag{22.156}$$

In this case for dilatation invariance we need $d(F_{ab}) = D/2$, and thus

$$d(F_{ab}) = D/2 \quad \Rightarrow \quad d(A_a) = (D-2)/2 \quad , \tag{22.157}$$

With this dimension-assignment for the gauge field, the D-dimensional Maxwell theory is dilatation invariant.

REMARKS:

(a) The result $d = (D-2)/2$ evidently only relies on the fact that one has a kinetic term which is quadratic in first derivatives, and therefore is valid for the standard action of a bosonic field af any spin.

(b) By the same reasoning, for actions of the fermionic type that are first-order in derivatives dilatation-invariance requires that the dimension of the spinorial field ψ is

$$d(\psi) = (D-1)/2 \quad , \tag{22.158}$$

reducing to the "canonical" dimension 3/2 of a spinor field for $D = 4$.

(c) In order to determine how to achieve scaling invariance in the sense of (22.97), we need to loook at the minimally coupled action

$$S_0[A_\alpha, g_{\alpha\beta}] = -\tfrac{1}{4}\int d^Dx\sqrt{g}\; g^{\alpha\gamma}g^{\beta\delta}F_{\alpha\beta}F_{\gamma\delta} \quad . \tag{22.159}$$

Since the Weyl weight of $\sqrt{g}g^{\alpha\gamma}g^{\beta\delta}$ is

$$w(\sqrt{g}g^{\alpha\gamma}g^{\beta\delta}) = D-4 \quad , \tag{22.160}$$

D arising from $\sqrt{g}$ and (-4) from the two inverse metrics, scaling invariance requires that the Weyl weight of A_α and $F_{\alpha\beta}$ should be

$$w(A_\alpha) = w(F_{\alpha\beta}) = (4-D)/2 \quad . \tag{22.161}$$

We see that in this case this is not the same as minus the dimension, but it should be clear from these two examples how to translate in general Weyl weights into scale dimensions and vice-versa.

(d) For $D = 4$ we have $w(A_\alpha) = 0$, and this is precisely the example already discussed in section 7.7. Since A_α does not transform, the action is then actually invariant under arbitrary Weyl transformations of the metric, not just constant rescalings.

(e) When $D \neq 4$, Maxwell theory has invariance under constant scalings but not under Weyl transformations. Since the Weyl transformation of the Maxwell Lagrangian is not gauge invariant, there is no obvious way to fix this in analogy with what we did in the scalar field case in section 22.3. In particular it seems that there is no way to construct a symmetric and traceless energy-momentum tensor for this theory, strongly suggesting (from this gravitational perspective) that the restriction of the theory to Minkowski space is an example of a theory that is scale- but not conformally invariant.[56]

Now let us return to the dilatations (22.140) for theories involving only scalar fields,

$$\bar{x}^a = \mathrm{e}^{-\lambda}x^a \quad , \quad \bar{\phi}(\bar{x}) = \mathrm{e}^{d_\phi\lambda}\phi(x) = \mathrm{e}^{d_\phi\lambda}\phi(\mathrm{e}^{\lambda}\bar{x}) \tag{22.162}$$

[56] See e.g. R. Jackiw, S.-Y. Pi, *Tutorial on Scale and Conformal Symmetries in Diverse Dimensions*, `arXiv:1101.4886 [math-ph]` and S. El-Showk, Y. Nakayama, S. Rychkov, *What Maxwell Theory in $d \neq 4$ teaches us about scale and conformal invariance*, `arXiv:1101.5385 [hep-th]` and references therein and thereto for discussions of this and related issues from a more field-theoretic perspective.

(for higher spin fields one would need to also invoke the Belinfante improvement procedure in the discussion below, but since this is a tangential issue for our purposes, the principal aim here being to prove (22.148), we will not consider this).

For Noether's theorem we need the corresponding infinitesimal transformations (at the same point x), namely

$$\delta x^a = -x^a \quad , \quad \delta\phi(x) = (d_\phi + x^b\partial_b)\phi(x) \ . \tag{22.163}$$

When ϕ has dimension d_ϕ, $\partial_a\phi$ has dimension $d_\phi + 1$. Therefore one also has

$$\delta(\partial_a\phi) = (d_\phi + 1 + x^b\partial_b)\partial_a\phi = \partial_a\delta\phi \tag{22.164}$$

(and likewise for the higher derivatives).

Now consider a Lagrangian $L = L(\phi, \partial_\alpha\phi)$ depending on the scalar field (or fields) ϕ and its derivatives. L has a well-defined dimension $d_L = d(L)$ if all the terms (summands) in L have the same dimension. In this case L satisfies

$$d_\phi\phi\frac{\partial L}{\partial\phi} + (d_\phi + 1)\partial_\alpha\phi\frac{\partial L}{\partial(\partial_\alpha\phi)} = d_L L \ . \tag{22.165}$$

Then the infinitesimal variation of L under dilatations is

$$\delta L = (d_L + x^b\partial_b)L \ . \tag{22.166}$$

We can also write this as

$$\delta L = \partial_a(x^a L) + (d_L - D)L \ . \tag{22.167}$$

Thus, when L has dimension $d_L = D$, the Lagrangian changes by a total derivative,

$$d_L = D \quad \Rightarrow \quad \delta L = \partial_a(x^a L) \ , \tag{22.168}$$

and we have a symmetry. This reproduces the result (22.143).

Since we have a continuous symmetry, by the usual Noether argument there will be a corresponding on-shell conserved current

$$j_N^a = x^a L - \Pi^a\delta\phi \ , \tag{22.169}$$

where

$$\Pi^a = \frac{\partial L}{\partial(\partial_a\phi)} \tag{22.170}$$

is the field momentum conjugate to ϕ. Using (22.163) for $\delta\phi$, one finds explicitly that the current is

$$j_N^a = \Theta^a_{\ b}x^b - d_\phi\Pi^a\phi \tag{22.171}$$

where

$$\Theta^a_{\ b} = \delta^a_b L - \Pi^a\partial_b\phi \tag{22.172}$$

is the canonical conserved Noether energy-momentum tensor (which is automatically symmetric for a scalar field). This already proves the claim in (22.146) that $j_N^a \neq \Theta^a_{\ b} x^b$ for $D \neq 2$. In this context the second term $\sim \Pi\phi$ in the Noether current is known as the *virial current*.

Using the Lagrangian from (22.150),

$$L = -\tfrac{1}{2}\eta^{ab}\partial_a\phi\partial_b\phi \quad \Rightarrow \quad \Pi^a = -\partial^a\phi \ , \tag{22.173}$$

and the explicit form of $T^{(c)}_{ab}$ (22.147), one finds that

$$\begin{aligned} J_D^a - j_N^a &= \frac{d_\phi}{2(D-1)} x^b \left(\delta^a_b \Box_\eta - \partial^a\partial_b\right)\phi^2 - d_\phi \phi \partial^a\phi \\ &= \frac{d_\phi}{2(D-1)}\left(x^a \Box\phi^2 - x^b\partial_b\partial^a\phi^2 - (D-1)\partial^a\phi^2\right) \\ &= \frac{d_\phi}{2(D-1)}\partial_b(x^a\partial^b - x^b\partial^a)\phi^2 \ , \end{aligned} \tag{22.174}$$

which establishes (22.148), with Δ^{ab} as given in (22.149).

All the above relations are valid off-shell. Calculating the divergence of (22.171) and using that on-shell one has

$$\partial_a\Theta^a_{\ b} = 0 \quad , \quad \partial_a\Pi^a = \frac{\partial L}{\partial\phi} \ , \tag{22.175}$$

one finds

$$\partial_a j_N^a = DL - \left[d_\phi\phi\frac{\partial L}{\partial\phi} + (d_\phi+1)\partial_\alpha\phi\frac{\partial L}{\partial(\partial_\alpha\phi)}\right] = (D - d_L)L \ , \tag{22.176}$$

thus confirming that dilatation invariance requires (and is equivalent to) $d_L = D$.

22.5 Energy-Momentum Tensor and (quasi-)Topological Couplings

The key result and insight of section 22.2, expressed in equations (22.94) and (22.92), is that the improved canonical Noether energy-momentum tensor $\hat{\Theta}_{\mu\nu}$ is actually equal on-shell to the (on-shell) covariantly conserved and (off-shell) symmetric covariant energy-momentum tensor $T_{\mu\nu}$.

At first sight this equality appears to be threatened by the quasi-topological couplings discussed in section 6.7, i.e. terms in the action that are independent of the metric even after minimal coupling to the gravitational field: as such terms contribute to the canonical energy-momentum tensor but not to the covariant energy-momentum tensor, it is not clear a priori how the two can end up being equal. While, as discussed in section 22.2, one can prove this equality in general, it is instructive to see how this comes about concretely in a non-trivial situation such as the present one.

We look at this issue in turn for the two examples of section 6.7, working in Minkowski space. Thus the relevant energy-momentum tensor is the restriction of the covariant energy-momentum tensor $T_{\alpha\beta}$ to Minkowski space, as in (7.113) or (22.93).

1. Axionic Coupling in (3+1) Dimensions

   The action (6.56) consists of the sum of a standard scalar field action, the standard Maxwell action and the axionic quasi-topological coupling term (6.57).

   We will need to know the equations of motion which are

$$\begin{aligned}\frac{\delta L_s}{\delta \phi} &= \tfrac{1}{4} f'(\phi) \tilde{F}^{\alpha\beta} F_{\alpha\beta} \\ \frac{\delta L_m}{\delta A_\beta} &= -\partial_\alpha (f \tilde{F}^{\alpha\beta}) \ ,\end{aligned} \tag{22.177}$$

   the current appearing on the right-hand side of the gauge field equations being identically conserved,

$$\partial_\beta \partial_\alpha (f \tilde{F}^{\alpha\beta}) \equiv 0 \ , \tag{22.178}$$

   as it should be.

   The canonical energy-momentum tensor of this model has the form

$$\Theta^\alpha_{\ \beta} = \Theta^\alpha_{s\ \beta} + \Theta^\alpha_{m\ \beta} + \Theta^\alpha_{a\ \beta} \tag{22.179}$$

   (the sum of the contributions from the scalar, Maxwell and axion action), while the covariant energy-momentum tensor has the form

$$T^\alpha_{\ \beta} = T^\alpha_{s\ \beta} + T^\alpha_{m\ \beta} \tag{22.180}$$

   because the axionic term does not couple to the metric. The scalar energy-momentum tensors are equal on the nose (i.e. off-shell, without improvement terms),

$$\Theta^\alpha_{s\ \beta} = T^\alpha_{s\ \beta} \ , \tag{22.181}$$

   so I will not discuss them further. The canonical Maxwell energy-momentum tensor is

$$\Theta^\alpha_{m\ \beta} = F^{\alpha\gamma} \partial_\beta A_\gamma - \tfrac{1}{4} \delta^\alpha_\beta F^{\gamma\delta} F_{\gamma\delta} \ , \tag{22.182}$$

   and the canonical axion energy-momentum tensor is

$$\Theta^\alpha_{a\ \beta} = f(\phi) \tilde{F}^{\alpha\gamma} \partial_\beta A_\gamma - \tfrac{1}{4} \delta^\alpha_\beta f(\phi) \tilde{F}^{\gamma\delta} F_{\gamma\delta} \ . \tag{22.183}$$

   As we know, for pure Maxwell theory one would proceed and argue as follows: one writes $\Theta^\alpha_{m\ \beta}$ as

$$\begin{aligned}\Theta^\alpha_{m\ \beta} &= F^{\alpha\gamma} F_{\beta\gamma} - \tfrac{1}{4} \delta^\alpha_\beta F^{\gamma\delta} F_{\gamma\delta} + F^{\alpha\gamma} \partial_\gamma A_\beta \\ &= T^\alpha_{m\ \beta} + \partial_\gamma (F^{\alpha\gamma} A_\beta) - (\partial_\gamma F^{\alpha\gamma}) A_\beta \ .\end{aligned} \tag{22.184}$$

The 1st term on the right-hand side is the covariant (symmetric, gauge invariant, metric) energy-momentum tensor, the 2nd is identically conserved, and the 3rd is zero on-shell, so that on-shell the canonical and covariant energy-momentum tensors agree modulo identically conserved terms that do not contribute to surface integrals etc. In the present (non-vacuum) case, one needs to take into account the non-trivial equation of motion

$$\partial_\alpha (F^{\alpha\beta} + f(\phi)\tilde{F}^{\alpha\beta}) = 0 \ , \tag{22.185}$$

and the axionic energy-momentum tensor $\Theta^\alpha_{a\ \beta}$. Doing this one finds

$$\begin{aligned}
\Theta^\alpha_{m\ \beta} + \Theta^\alpha_{a\ \beta} = T^\alpha_{m\ \beta} & & \text{(the term one expects)} \\
+ \partial_\gamma [(F^{\alpha\gamma} + f(\phi)\tilde{F}^{\alpha\gamma})A_\beta] & & \text{(identically conserved)} \\
+ [\partial_\gamma (F^{\gamma\alpha} + f(\phi)\tilde{F}^{\gamma\alpha})]A_\beta & & \text{(on-shell zero)} \\
+ f(\phi)\left[\tilde{F}^{\alpha\gamma}F_{\beta\gamma} - \tfrac{1}{4}\delta^\alpha_\beta \tilde{F}^{\gamma\delta}F_{\gamma\delta}\right] & & \text{(???)}
\end{aligned} \tag{22.186}$$

At first sight the last terms appears to spoil the claimed relation between the canoncial and covariant energy-momentum tensors. However, either by explicit calculation in components or from a more covariant argument given below one finds that the term in square barackest in the last line,

$$\tilde{T}^\alpha_{m\ \beta} := \tilde{F}^{\alpha\gamma}F_{\beta\gamma} - \tfrac{1}{4}\delta^\alpha_\beta \tilde{F}^{\gamma\delta}F_{\gamma\delta} \tag{22.187}$$

is actually cooperatively identically zero (for any anti-symmetric $F_{\alpha\beta}$ and its dual),

$$\tilde{T}^\alpha_{m\ \beta} \equiv 0 \ . \tag{22.188}$$

Therefore one concludes that, as it should be, the canonical and covariant energy-momentum tensors are indeed equal on-shell modulo identically conserved terms. Here is a general proof that $\tilde{T}^\alpha_{m\ \beta} = 0$ identically:

- The first term of $\tilde{T}^\alpha_{m\ \beta}$ can be written as

$$\begin{aligned}
\tilde{F}^{\alpha\gamma}F_{\beta\gamma} &= \tfrac{1}{2} \in^{\alpha\gamma\mu\nu} F_{\mu\nu}F_{\beta\gamma} = -\tfrac{1}{2} \in^{\alpha\gamma\mu\nu} F_{\mu\nu}F_{\gamma\beta} \\
&= -\tfrac{1}{2} \in^{\alpha\gamma\mu\nu} F_{[\mu\nu}F_{\gamma]\beta} = -\tfrac{1}{2} \in^{\alpha\gamma\mu\nu} F_{[\mu\nu}F_{\gamma\beta]}
\end{aligned} \tag{22.189}$$

  because anti-symmetry of $F_{\gamma\beta}$ implies that $F_{[\mu\nu}F_{\gamma]\beta}$ is already totally anti-symmetric in all its 4 indices,

$$F_{[\mu\nu}F_{\gamma]\beta} = F_{[\mu\nu}F_{\gamma\beta]} \ . \tag{22.190}$$

- As a consequence, it must be proportional to $\in_{\mu\nu\gamma\beta}$,

$$F_{[\mu\nu}F_{\gamma]\beta} = F_{[\mu\nu}F_{\gamma\beta]} = g \in_{\mu\nu\gamma\beta} \ , \tag{22.191}$$

  with g determined by

$$\in^{\mu\nu\gamma\beta}\in_{\mu\nu\gamma\beta} = -24 \quad \Rightarrow \quad g = -\tfrac{1}{12}\tilde{F}^{\gamma\delta}F_{\gamma\delta} \ . \tag{22.192}$$

- It follows that the first term of $\tilde{T}^{\alpha}_{m\ \beta}$ is

$$\begin{aligned}\tilde{F}^{\alpha\gamma}F_{\beta\gamma} &= -\tfrac{1}{2}\in^{\alpha\gamma\mu\nu} F_{[\mu\nu}F_{\gamma\beta]} = \tfrac{1}{24}\ \in^{\alpha\gamma\mu\nu}\in_{\mu\nu\gamma\beta}\ \ \tilde{F}^{\gamma\delta}F_{\gamma\delta} \\ &= \tfrac{1}{24}\ 6\delta^{\alpha}_{\beta}\ \tilde{F}^{\gamma\delta}F_{\gamma\delta} = \tfrac{1}{4}\delta^{\alpha}_{\beta}\ \tilde{F}^{\gamma\delta}F_{\gamma\delta}\end{aligned} \tag{22.193}$$

  which is equivalent to the statement that $\tilde{T}^{\alpha}_{m\ \beta} \equiv 0$. (qed)

2. Maxwell - Chern-Simons Theory in (2+1) Dimensions

   The action (the Lagrangian is given in (6.63)) is the sum of the standard Maxwell action and a quasi-topological Chern-Simons term. The canonical energy-momentum tensor is

$$\Theta^{\alpha}_{\ \beta} = \Theta^{\alpha}_{m\ \beta} + \Theta^{\alpha}_{cs\,\beta} \tag{22.194}$$

   with

$$\Theta^{\alpha}_{m\ \beta} = T^{\alpha}_{m\ \beta} + \partial_{\gamma}(F^{\alpha\gamma}A_{\beta}) - (\partial_{\gamma}F^{\alpha\gamma})A_{\beta}\ \ , \tag{22.195}$$

   as usual, and

$$\Theta^{\alpha}_{cs\,\beta} = k\ \delta^{\alpha}_{\beta}L_{cs} + k\in^{\alpha\delta\gamma}\ A_{\delta}\partial_{\beta}A_{\gamma}\ \ . \tag{22.196}$$

   The covariant energy-momentum tensor, on the other hand, is just the Maxwell $T^{\alpha}_{m\ \beta}$,

$$T^{\alpha}_{\ \beta} = T^{\alpha}_{m\ \beta}\ \ , \tag{22.197}$$

   because the CS term is metric independent.

   After some rearrangement the CS contribution to the canonical energy-momentum tensor can be written as

$$\begin{aligned}\Theta^{\alpha}_{cs\,\beta} &= k\ \delta^{\alpha}_{\beta}L_{cs} + \tfrac{1}{2}k\in^{\delta\alpha\gamma}(A_{\gamma}F_{\beta\delta} + A_{\delta}F_{\gamma\beta} + A_{\beta}F_{\delta\gamma}) \\ &\quad + k\ \in^{\delta\alpha\gamma}\ A_{\beta}F_{\gamma\delta} - k\ \partial_{\gamma}(\in^{\delta\alpha\gamma}\ A_{\delta}A_{\beta})\end{aligned} \tag{22.198}$$

   Thus for the sum of the Maxwell and CS contributions one has the result

$$\begin{aligned}\Theta^{\alpha}_{m\ \beta} + \Theta^{\alpha}_{cs\,\beta} = T^{\alpha}_{m\ \beta}& &&\text{(the term one expects)} \\ + \partial_{\gamma}[(F^{\alpha\gamma} - k\ \in^{\delta\alpha\gamma}\ A_{\delta})A_{\beta}]& &&\text{(identically conserved)} \\ + [\partial_{\gamma}F^{\gamma\alpha} + k\ \in^{\alpha\gamma\delta}\ F_{\gamma\delta}]A_{\beta}& &&\text{(on-shell zero)} \\ + \tilde{T}^{\alpha}_{\ \beta}\ \ ,&\end{aligned} \tag{22.199}$$

   with

$$\tilde{T}^{\alpha}_{\ \beta} = k\ \delta^{\alpha}_{\beta}L_{cs} + \tfrac{1}{2}k\in^{\delta\alpha\gamma}(A_{\gamma}F_{\beta\delta} + A_{\delta}F_{\gamma\beta} + A_{\beta}F_{\delta\gamma})\ \ . \tag{22.200}$$

   By an argument analogous to that in the previous example, one can show that

$$\tilde{T}^{\alpha}_{\ \beta} \equiv 0 \tag{22.201}$$

   identically:

- The term in brackets is totally anti-symmetric,

$$A_\gamma F_{\beta\delta} + A_\delta F_{\gamma\beta} + A_\beta F_{\delta\gamma} = 3A_{[\gamma}F_{\beta\delta]} \ ; \tag{22.202}$$

- one has

$$A_{[\gamma}F_{\beta\delta]} = -\tfrac{1}{6}(\in^{\lambda\mu\nu} A_\lambda F_{\mu\nu}) \in_{\gamma\beta\delta} = -\tfrac{1}{3}L_{cs} \in_{\gamma\beta\delta} \ ; \tag{22.203}$$

- thus

$$\begin{aligned} \tilde{T}^\alpha_\beta &= k\ \delta^\alpha_\beta L_{cs} + \tfrac{3}{2}k\ \ \in^{\delta\alpha\gamma} A_{[\gamma}F_{\beta\delta]} \\ &= k\ \delta^\alpha_\beta L_{cs} - \tfrac{1}{2}k\ \ \in^{\delta\alpha\gamma}\in_{\gamma\beta\delta}\ \ L_{cs} = 0 \ . \end{aligned} \tag{22.204}$$

These two examples evidently show a common and general pattern, and it is now easy to generalise this to topological or quasi-topological interactions in higher dimensions, but this is not necessary - after all, one can prove in complete generality that the canonical and covariant energy-momentum tensors are equal on-shell modulo identically conserved terms, and the present examples just serve to illustrate how this works in a non-trivial situation.

22.6 Comments on Gravitational Energy

It may not have escaped your attention that in the entire discussion of energy and energy-momentum tensors of fields in a gravitiational field the notion of the energy of the gravitational field itself has not appeared so far, even though clearly there can be an exchange of energy between matter and gravitational fields and one should not expect one to be conserved without taking into account the other. This is evidently a major omission, and I will try to rectify this now, but as you will perhaps understand from the discussion below there is a good reason why I have so far tried to avoid this issue.

To get us started, let us return to the covariant conservation law $\nabla_\mu T^{\mu\nu} = 0$ for the matter energy-momentum tensor which played such a key role above, and which we now write more explicitly as the "non-conservation law" (cf. (7.144))

$$\nabla_\mu T^{\mu\nu} = 0 \quad \Leftrightarrow \quad \partial_\mu(\sqrt{g}\, T^{\mu\nu}) = -\sqrt{g}\, \Gamma^\nu_{\mu\lambda} T^{\mu\lambda} \ . \tag{22.205}$$

In the spirit of the interpretation of G^ν in $\partial_\mu T^{\mu\nu} = G^\nu$ (7.142) as the external force density acting on the system, one might like to interpret the term on the right-hand side as the gravitational external force density representing the exchange of energy between matter (represented by $T_{\mu\nu}$) and the gravitational field (represented by $\Gamma^\lambda_{\mu\nu}$). The external matter force-density G^ν can be made to disappear from the right-hand side of the (non-)conservation equation (7.142) by including and taking into account also the energy-momentum tensor of the source (one has no right to expect to obtain a true conservation law otherwise).

This suggests that also in the present gravitational case it should be possible to define a conserved total energy-momentum tensor by taking into account not only the matter energy-momentum tensor but also the energy-momentum of the gravitational field itself. However, this is easier said then done, and attempts to make sense of this and make this well-defined are an important and controversial part of research in general relativity from the earliest days of general relativity to today. For example, Noether's fundamental and famous work *Invariante Variationsprobleme* on symmetries and variational problems was prompted by questions of Hilbert regarding the apparent failure of what he referred to as the "energy theorem" in general relativity.[57]

I will just make some short, scattered and introductory remarks on this subject, but must precede these with a *caveat*:

This is treacherous territory and I cannot guarantee that even these elementary remarks are widely considered to be uncontroversial (in fact, the number of uncontroversial statements one *can* make about this subject is probably quite small - but is likely to include this parenthetical remark ...).

1. Covariant Gravitational Energy-Momentum Tensor?

   One's first thought may be that, in precise analogy with the definition of the covariant matter energy-momentum tensor, the gravitational energy-momentum tensor should be defined in terms of the variation of the gravitational (Einstein-Hilbert) action with respect to the metric, or, equivalently, that the total energy-momentum tensor should be defined as the variation of the total (Einstein-Hilbert + matter) action with respect to the metric. While this seems to be the logical thing to do,

$$\begin{aligned} T^{\text{tot}}_{\mu\nu} &\stackrel{?}{=} -\frac{2}{\sqrt{g}}\frac{\delta}{\delta g^{\mu\nu}}\left(\frac{1}{16\pi G_N}S_{EH}[g_{\alpha\beta}] + S_M[\phi, g_{\alpha\beta}]\right) \\ &= -\frac{1}{8\pi G_N}G_{\mu\nu} + T_{\mu\nu} \quad , \end{aligned} \tag{22.206}$$

   it is evidently not useful because by the variational principle for general relativity this total energy-momentum tensor is identically zero for any solution to the Einstein equations. This reinterpretation of the Einstein tensor as the energy-momentum tensor of the gravitational field was first suggested by Lorentz (in 1916) and Levi-Civita (in 1917), but immediately dismissed by Einstein (1918).[58]

   Whatever thus the poetic or philosophical virtues of such a definition might be ('everything from nothing', 'the total energy of the universe is zero'), it is clearly

[57] See e.g. N. Byers, *E. Noether's Discovery of the Deep Connection Between Symmetries and Conservation Laws*, `arXiv:physics/9807044 [physics.hist-ph]` for an account of the historical circumstances of these discoveries.

[58] For a discussion and references cf. section 61 of W. Pauli's famous and amazing (written at the age of 21!) 1921 encyclopaedia article *Relativitätstheorie* (or any of its translations).

deficient in many other respects and does not provide a whole lot of insight. In particular, with this definition one would assign zero gravitational energy to any source-free region of space-time, i.e. to vacuum gravitational fields (like gravitational waves, the exterior of a star etc.). Evidently, therefore, this definition fails to capture some essential aspects of and contributions to what one would commonly refer to as gravitational potential energy.

2. Non-Existence of a Covariant Gravitational Energy Density?

   Let us therefore return to (22.205). Note that the would-be gravitational force density term on the right-hand side of (22.205) is non-tensorial. It is often argued that this non-tensoriality reflects the equivalence principle and is therefore also a necessary, profound, and characteristic feature of any potential definition of the energy density and/or the energy-momentum 'tensor' of the gravitational field.

   The argument for this runs along the lines of

   - "by the equivalence principle one can always go to an inertial, i.e. freely falling, reference frame in which, along the worldline of an observer (Fermi coordinates), the effects of gravity are absent"
   - "in such a frame the energy-momentum 'tensor' (or energy density) of the gravitational field should therefore be zero"
   - "if the gravitational energy-momentum 'tensor' were a true tensor, it would then have to be zero in all reference frames"
   - "thus a non-trivial gravitational energy-momentum 'tensor' cannot be tensorial"

   This argument, as compelling as it may appear at first sight, is not watertight, however, in particular because the statement "the effects of gravity are absent" refers only to the metric and its first derivatives, the deviation from the flat metric showing up in the second derivatives of the metric, i.e. the Riemann curvature tensor. Thus it is possible that an appropriate notion of gravitational energy is just sufficiently 'non-local' (by depending on the second derivatives of the metric) that it cannot be eliminated by going to a freely falling reference frame along a single worldline. Nevertheless, this provides a first indication that the notion of energy density of the gravitational field is perhaps somewhat more subtle than expected.

3. Comparison with the Newtonian Situation

   The above line of reasoning can be strengthened and illuminated somewhat by looking at the Newtonian limit. In the Newtonian theory, there is a (more or less, see the comment below) well-defined notion of gravitational potential energy density, and it is proportional to -$(\vec{\nabla}\phi)^2$, minus the square of the gradient of the gravitational potential ϕ (the strictly negative gravitational binding energy).

With the metric $g_{\mu\nu}$ playing the role of ϕ in general relativity, one might therefore expect to be able to write the gravitational energy density as an expression quadratic in the first derivatives of the metric or, equivalently, quadratic in the Christoffel symbols. However, there is no scalar or tensor that one can build only from the metric and its first derivatives and thus, the reasoning goes, whatever generalises the Newtonian notion of gravitational energy density cannot be tensorial.

It is a curious and apparently under-appreciated fact that already the Newtonian theory exhibits an analogous behaviour, i.e. that even in the Newtonian theory the notion of gravitiational energy density at a point is strictly speaking ambiguous.[59] This is essentially a consequence of the equality of gravitational and inertial mass, the precursor of the equivalence principle, which allows one to replace

$$\vec{\nabla}\phi \to \vec{\nabla}\phi + \vec{a} = \vec{\nabla}(\phi + \vec{a}.\vec{x}) \tag{22.207}$$

for any constant acceleration vector $\vec{a}$. One can thus always make the the force and the potential energy density vanish at a point x_0 by choosing $\vec{a} = -\vec{\nabla}\phi(x_0)$ (i.e. by going to a freely falling reference frame, as in our first discussion of the equivalence principle in section 1.1). This local ambiguity of the gravitational potential energy density can be (and is usually implicitly) fixed by invoking (non-locally) the required or expected asymptotic behaviour of the gravitational field (e.g. that it goes to zero asymptotically).

It is this local ambiguity that sets the gravitational potential energy density apart from (formally similar) quantities like the electrostatic potential energy density. This suggests that perhaps this non-localisability or non-tensoriality of the gravitational potential energy density is an intrinsic property of this quantity, thus simply an inevitable feature and not a bug.

Again this argument may sound compelling but is not without loopholes. For example, an integration by parts would put the Newtonian gravitational potential energy density into the form $\phi\Delta\phi$, and following the argument in section 19.3 (extended to quadratic order in the fluctuations of the metric) it is certainly conceivable that one come up with some tensorial generalisations of this expression (involving second derivatives of the metric, i.e. contractions of the Riemann tensor). However, then one needs to investigate whether such a candidate expression has other desired or desirable properties, beyond having the "right" Newtonian limit, in order to qualify as a candidate for the gravitational energy density or energy-momentum tensor.

[59] For a lucid recent discussion of this issue and its ramifications in general relativity, see J. Frauendiener, L. Szabados, *A note on the post-Newtonian limit of quasi-local energy expressions*, `arXiv:1102.1867 [gr-qc]`.

In some way, this situation reflects the tension of having to choose between working with

- either a covariant action, but involving second derivatives of the metric (the Einstein-Hilbert action (20.2)),
- or a non-covariant action, but nicely quadratic in the first derivatives of the metric (the Einstein action (20.45)).

4. Gravitational Pseudotensors?

If among the desirable features for a potential gravitational energy-momentum tensor one wants to include the statement that, whatever the energy-momentum 'tensor' $t_G^{\mu\nu}$ of the gravitational field might be, it should be such that the total energy-momentum tensor consisting of the sum of the matter and gravitational energy momentum tensors is conserved (in the ordinary sense of providing conserved currents), then this conflicts directly with the requirement of tensoriality. Indeed, then the total energy-momentum tensor would have to satisfy some ordinary local conservation law of the type

$$\partial_\mu \sqrt{g}(T^{\mu\nu} + t_G^{\mu\nu}) = 0 \tag{22.208}$$

(or perhaps with some other power of $\sqrt{g}$). However, since $\partial_\mu(\sqrt{g}T^{\mu\nu})$ is not tensorial by itself (it is only the first part of the covariant derivative), it is clear that $t_G^{\mu\nu}$ is then also invariably not tensorial. And if the sum $T^{\mu\nu} + t_G^{\mu\nu}$ were covariantly conserved instead, then we would again face the same problem as in (22.205).

One simple way to construct (non-tensorial) objects $t_G^{\mu\nu}$ satisfying an equation like (22.208), known as *energy-momentum pseudotensors*, and illustrating the high degree of arbitrariness in such a construction uses the method of *superpotentials*:

- Construct an object
$$U^{\mu\lambda\nu} = U^{[\mu\lambda]\nu} \tag{22.209}$$
out of the metric and its first derivatives (say), so that in particular its divergence $\partial_\lambda(\sqrt{g}U^{\mu\lambda\nu})$ is identically conserved,
$$\partial_\mu\partial_\lambda(\sqrt{g}U^{\mu\lambda\nu}) \equiv 0 \ . \tag{22.210}$$
- Use this to split off a total derivative (divergence) term from the Einstein tensor to *define* a corresponding candidate gravitational pseudo-tensor $t_U^{\mu\mu}$ by
$$16\pi G_N\sqrt{g}t_U^{\mu\nu} := -2\sqrt{g}G^{\mu\nu} + \partial_\lambda(\sqrt{g}U^{\mu\lambda\nu}) \ . \tag{22.211}$$
- Using the Einstein equations $G^{\mu\nu} = 8\pi G_N T^{\mu\nu}$, this can be written as
$$16\pi G_N\sqrt{g}(T^{\mu\nu} + t_U^{\mu\nu}) = \partial_\lambda(\sqrt{g}U^{\mu\lambda\nu}) \ , \tag{22.212}$$

and shows that

$$\partial_\mu \sqrt{g}(T^{\mu\nu} + t_U^{\mu\nu}) = 0 \ , \tag{22.213}$$

as desired.

A multitude of such (and related) objects have been constructed, e.g. by Einstein himself, or by Landau and Lifshitz, the pseudotensor of the latter at least having the attractive feature that it is symmetric, by Weinberg, by Møller, etc. These are typically indeed quadratic in the first derivatives of the metric (and hence, as we already discussed, necessarily non-tensorial but with a potentially viable Newtonian limit). They suffer from at least 2 ambiguities, however, one being the choice of the superpotential, and the second (since we are dealing with non-tenorial quantities) the choice of coordinates / reference frame. Their overall usefulness for providing a local expression for the energy-density of the gravitational field is therefore somewhat (or severely) limited.[60]

5. Looking for the Right Answer to the Wrong Question?

   Such obervations and considerations (and a dislike of pseudo-tensorial objects) have led to the realisation that the notion of local energy density of the gravitational field at a point may not be particularly useful or meaningful in general relativity, as already summarised so eloquently by Misner, Thorne and Wheeler in the early 1970s:

   > Anybody who looks for a magic formula for "local gravitational energy-momentum" is looking for the right answer to the wrong question. Unhappily, enormous time and effort were devoted in the past to trying to "answer this question" before investigators realised the futility of the enterprise.[61]

6. Total (Gravitational + Matter) Energy

   There appears to be a general consensus, though, that it is possible (and meaningful) to unambiguously define, at the other extreme, the *total* energy, momentum and angular momentum of an isolated (asymptotically flat, say) gravitating system in general relativity, and that these quantities can be calculated by boundary surface integrals over a 2-sphere 'at infinity' (e.g. the ADM charges discussed in sections 21.12 and 23.4). Problems with the potential non-tensoriality of the boundary integrand can in those cases be avoided e.g. by appealing to an inertial Minkowskian observer at infinity and only requiring a tensorial behaviour under Lorentz transformations at infinity.

[60] For modern accounts and further references see e.g. the brief discussions in the textbooks N. Straumann, *General Relativity* or T. Ortin, *Gravity and Strings*, the review article by L. Szabados (footnote 63) and C. Chen, J. Nester, R. Tung, *Gravitational Energy for GR and Poincaré gauge theories: a covariant Hamiltonian approach*, `arXiv:1507.07300`.

[61] C. Misner, K. Thorne, J. Wheeler, *Gravitation*, section 20.4.

7. Quasi-Local Gravitational Energy?

   Perhaps the most interesting question is therefore if one can do a bit better than that and talk in a meaningful way not only about the total energy contained in the total spatial volume of a space-time, but also somewhat more locally (quasi-locally) about the total energy (including the effects of gravity) in some finite localised volume of space-time.[62] In spite of the problems with strictly local expressions for the energy-density, numerous candidates for such *quasi-local* expressions for the energy have indeed been proposed and debated in the literature, and have been argued to provide reasonable definitions of gravitational energy in specific circumstances.[63]

   The final outcome may be that eventually a single "best" and universally accepted and agreed upon definition of quasi-local gravitational energy emerges from these investigations. However, at present it seems at least equally (if not far more) likely that no such universal "best" definition exists and that the "right" answer depends on the specific context and question one is asking.

[62] See P. McGrath, R. Epp, R. Mann, *Quasilocal Conservation Laws: Why We Need Them*, `arXiv:1206.6512 [gr-qc]` for a particularly illuminating discussion of the shortcomings of strictly local definitions to account for the effects of accelerating reference frames in special relativity (and thus, by the equivalence principle, of gravity) and how these can be overcome by adopting a quasi-local perspective.

[63] For an excellent detailed review and critical assessment of the attempts to define a notion of local or quasi-local energy density of the gravitational field, see L. Szabados, *Quasi-Local Energy-Momentum and Angular Momentum in General Relativity*, Living Rev. Relativity 12, (2009), 4; `http://www.livingreviews.org/lrr-2009-4`.

23 Linearised Gravity and Gravitational Waves

23.1 Preliminary Remarks

While it is evidently of interest to apply general relativity to situations where the gravitational field is so strong that the Newtonian approximation fails (and we will consider such situations later on), in most ordinary situations, the gravitational field is weak, very weak, and then it is legitimate to work with a linearisation of the Einstein equations.

When we first derived the Einstein equations we checked that we were doing the right thing by deriving the Newtonian theory in the limit where

1. the gravitational field is weak
2. the gravitational field is static
3. matter is non-relativistic

The fact that the 2nd condition had to be imposed in order to recover the Newtonian theory is interesting in its own right, as it suggests that there are novel features in general relativity, even for weak fields, when the 2nd condition is not imposed. Indeed, as we will see, in this more general setting one discovers that general relativity also predicts the existence of *gravitational waves*, i.e. linearised perturbations of the gravitational field propagating like ordinary waves on the (Minkowski) background geometry. Our principal aim in this section will be to derive these linearised equations, show that they are wave equations, and study some of their more elementary consequences.

An important next step would be to study and understand how or under which circumstances gravitational waves are created and how they can be detected. These, unfortunately, are rather complicated questions in general and I will not enter into this. The things I will cover in the following are much more elementary, both technically and conceptually, than other applications of general relativity discussed elsewhere in these notes.

23.2 Linearised Einstein Equations

We express the weakness of the gravitational field by the condition that the metric be 'close' to that of Minkowski space, i.e. that

$$g_{\mu\nu} = g^{(1)}_{\mu\nu} \equiv \eta_{\mu\nu} + h_{\mu\nu} \tag{23.1}$$

with $|h_{\mu\nu}| \ll 1$. This means that we will drop terms which are quadratic or of higher power in $h_{\mu\nu}$. Here and in the following the superscript $^{(1)}$ indicates that we keep only up to linear (first order) terms in $h_{\mu\nu}$. In particular, the inverse metric is

$$g^{(1)\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \tag{23.2}$$

where indices are raised with $\eta^{\mu\nu}$.

As one has thus essentially chosen a background metric, the Minkowski metric, one can think of the linearised version of the Einstein equations (which are field equations for $h_{\mu\nu}$) as a Lorentz-invariant theory of a symmetric tensor field propagating in Minkowski space-time. I won't dwell on this but it is good to keep this in mind. It gives rise to the field theorist's picture of gravity as the theory of an interacting spin-2 field (which is useful for many purposes but which I do not subscribe to unconditionally because it is an inherently perturbative and background dependent picture).

It is straightforward to work out the Christoffel symbols and curvature tensors in this approximation. The terms quadratic in the Christoffel symbols do not contribute to the Riemann curvature tensor and one finds

$$\begin{aligned} \Gamma^{(1)\mu}_{\nu\lambda} &= \eta^{\mu\rho}\tfrac{1}{2}(\partial_\lambda h_{\rho\nu} + \partial_\nu h_{\rho\lambda} - \partial_\rho h_{\nu\lambda}) \\ R^{(1)}_{\mu\nu\rho\sigma} &= \tfrac{1}{2}(\partial_\rho\partial_\nu h_{\mu\sigma} + \partial_\mu\partial_\sigma h_{\rho\nu} - \partial_\rho\partial_\mu h_{\nu\sigma} - \partial_\nu\partial_\sigma h_{\rho\mu}) \end{aligned} \tag{23.3}$$

(this result for the Riemann tensor can also be inferred directly from the expression (8.20) of the Riemann tensor at the origin of an inertial coordinate system). Hence

- the linearised Ricci tensor is

$$R^{(1)}_{\mu\nu} = \tfrac{1}{2}(\partial_\sigma\partial_\nu h^\sigma_{\ \mu} + \partial_\sigma\partial_\mu h^\sigma_{\ \nu} - \partial_\mu\partial_\nu h - \Box h_{\mu\nu}) \ , \tag{23.4}$$

  where

$$h \equiv h^\mu_{\ \mu} = \eta^{\mu\nu}h_{\mu\nu} = -h_{00} + \delta^{ik}h_{ik} \tag{23.5}$$

  is the trace of $h_{\mu\nu}$ and

$$\Box = \partial^\mu\partial_\mu = -(\partial_0)^2 + \Delta \tag{23.6}$$

  is the Minkowski wave operator;

- the linearised Ricci scalar is

$$R^{(1)} = \eta^{\mu\nu}R^{(1)}_{\mu\nu} = \partial_\mu\partial_\nu h^{\mu\nu} - \Box h \ ; \tag{23.7}$$

- the linearised Einstein tensor is

$$\begin{aligned} G^{(1)}_{\mu\nu} &= R^{(1)}_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}R^{(1)} \\ &= \tfrac{1}{2}(\partial_\sigma\partial_\nu h^\sigma_{\ \mu} + \partial_\sigma\partial_\mu h^\sigma_{\ \nu} - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial_\rho\partial_\sigma h^{\rho\sigma} + \eta_{\mu\nu}\Box h) \ ; \end{aligned} \tag{23.8}$$

- and the linearised vacuum Einstein equations

$$R^{(1)}_{\mu\nu} = 0 \quad \Leftrightarrow \quad G^{(1)}_{\mu\nu} = 0 \tag{23.9}$$

  are thus

$$R^{(1)}_{\mu\nu} = 0 \quad \Leftrightarrow \quad \partial_\sigma\partial_\nu h^\sigma_{\ \mu} + \partial_\sigma\partial_\mu h^\sigma_{\ \nu} - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} = 0 \ , \tag{23.10}$$

  or, equivalently (in the form $G^{(1)}_{\mu\nu} = 0$)

$$\partial_\sigma\partial_\nu h^\sigma_{\ \mu} + \partial_\sigma\partial_\mu h^\sigma_{\ \nu} - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial_\rho\partial_\sigma h^{\rho\sigma} + \eta_{\mu\nu}\Box h = 0 \ . \tag{23.11}$$

The latter can be derived from the quadratic Minkowski space (Poincaré invariant) field theory action (the *Fierz-Pauli action* (1939))

$$
\begin{aligned}
S[h_{\mu\nu}] &= \int d^4x\, \mathcal{L}(h_{\mu\nu}, \partial_\sigma h_{\mu\nu}) \\
\mathcal{L}(h_{\mu\nu}, \partial_\sigma h_{\mu\nu}) &= -\tfrac{1}{4} h_{\mu\nu,\sigma} h^{\mu\nu,\sigma} + \tfrac{1}{2} h_{\mu\nu,\sigma} h^{\sigma\mu,\nu} + \tfrac{1}{4} h^{,\sigma} h_{,\sigma} - \tfrac{1}{2} h_{,\sigma} h^{\mu\sigma}{}_{,\mu}
\end{aligned}
\tag{23.12}
$$

for a free massless spin-2 field described by the Lorentz tensor $h_{\mu\nu}$. Indeed, variation of the action with respect to the $h_{\mu\nu}$ (and the usual integration by parts) leads to

$$
h_{\mu\nu} \to h_{\mu\nu} + \delta h_{\mu\nu} \quad \Rightarrow \quad \delta S[h_{\mu\nu}] = -\int d^4x\, G^{(1)}_{\mu\nu} \delta h^{\mu\nu} \quad . \tag{23.13}
$$

Remarks:

1. Signs 1: note that from (23.2) one has

$$
g^{(1)\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \quad \Rightarrow \quad \delta g^{(1)\mu\nu} = -\delta h^{\mu\nu} \quad , \tag{23.14}
$$

   so that the sign in (23.13) is the standard one if expressed in terms of the variation of $g^{\mu\nu}$,

$$
\delta S[h_{\mu\nu}] = +\int d^4x\, G^{(1)}_{\mu\nu} \delta g^{(1)\mu\nu} \quad . \tag{23.15}
$$

2. Signs 2: the sign and overall normalisation of the Lagrangian are also such that one obtains the canonically normalised kinetic term

$$
\mathcal{L} = +\tfrac{1}{4} (\partial_t h_{ik})^2 + \ldots \tag{23.16}
$$

   for a 2-index field.

3. Note that the Lagrangian is *not* the linearised Einstein-Hilbert Lagrangian, i.e. the linearised Ricci scalar (23.7) (as the latter is, by construction, linear in $h_{\mu\nu}$). Rather,

$$
\mathcal{L} = \tfrac{1}{2} h^{\mu\nu} G^{(1)}_{\mu\nu} \quad + \quad \text{total derivative} \quad , \tag{23.17}
$$

   where the total derivative contributions serve to turn the terms of the form $h\partial^2 h$ arising from $h^{\mu\nu} G^{(1)}_{\mu\nu}$ into the standard $(\partial h)^2$-form, so that $\mathcal{L}$ is of a standard (quadratic) form in the derivatives of $h_{\mu\nu}$, as it should be.

4. Note that we can indeed take the integration measure to be the flat Minkowski measure d^4x, as $\mathcal{L}$ is already quadratic in $h_{\mu\nu}$ and its derivatives, so that any contribution of $h_{\mu\nu}$ to the measure would give a subleading contribution.

5. This theory, just like its spin-1 Maxwell counterpart, has a local gauge invariance which we will return to and make use of below.

6. In the presence of matter the linearised Einstein equations are

$$G^{(1)}_{\mu\nu} = 8\pi G_N T^{(0)}_{\mu\nu} \quad . \tag{23.18}$$

Note that only the zero'th order term in the h-expansion appears on the right hand side of this equation. This is due to the fact that $T_{\mu\nu}$ must itself already be small in order for the linearised approximation to be valid, i.e. $T^{(0)}_{\mu\nu}$ should be of order $h_{\mu\nu}$. Therefore, any terms in $T_{\mu\nu}$ depending on $h_{\mu\nu}$ would already be of order $(h_{\mu\nu})^2$ and can be dropped.

7. The conservation law for the energy-momentum tensor is therefore just

$$\partial_\mu T^{(0)\mu\nu} = 0 \quad , \tag{23.19}$$

and this is indeed compatible with the linearised Bianchi identity

$$\partial_\mu G^{(1)\mu\nu} = 0 \quad , \tag{23.20}$$

which can easily be verified, and which reflects the invariance of the theory under the linearised coordinate transformations to be discussed below.

23.3 Newtonian Limit Revisited

In section 19.3 we verified that the Newtonian (weak field, static, non-relativistic matter) limit of the Einstein equations reduces to the Newtonian (Poisson) equation $\Delta\phi = 4\pi G_N \rho$. This is a special case of the above general weak-field limit equations (23.18), with $G^{(1)}_{\mu\nu}$ given in (23.8), and it will be instructive to redo the calculation of section 19.3 from this more general perspective.

We thus assume an energy-momentum tensor whose only non-negligible component is the energy density $T_{00} = \rho$, with ρ static and $G_N\rho \ll 1$. Then we can assume that the deviations of the space-time geometry from the Minkowski metric are small and time-independent,

$$g_{\mu\nu} = g^{(1)}_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad , \quad \partial_0 h_{\mu\nu} = 0 \quad . \tag{23.21}$$

Then the (00)-component of the Einstein tensor reduces (after some cancellations) to

$$G^{(1)}_{00} = -\tfrac{1}{2}\Delta(\delta^{ik}h_{ik}) + \tfrac{1}{2}\partial_i\partial_k h^{ik} \quad . \tag{23.22}$$

In particular, h_{00} and its derivatives have dropped out of this expression, which appears to be at odds with the desired $G_{00} = -\Delta g_{00}$ for Newtonian fields (section 19.3). However, we have not yet used at all the condition that $T_{ik} = 0 \to G_{ik} = 0$. In particular, for static perturbations we find from (23.8) that the trace of the spatial components of the Einstein tensor is

$$\delta^{ik}G^{(1)}_{ik} = \tfrac{1}{2}\Delta(\delta^{ik}h_{ik}) - \tfrac{1}{2}\partial_i\partial_k h^{ik} - \Delta h_{00} \tag{23.23}$$

so that

$$G^{(1)}_{ik} = 0 \quad \Rightarrow \quad G^{(1)}_{00} = -\Delta h_{00} \ . \tag{23.24}$$

This is precisely the relation required (and verified) in the analysis of section 19, in order to have the correct Newtonian limit.

23.4 ADM and Komar Energies of an Isolated System

As we are in the realm of a Poincaré-invariant classical field theory, we can define the total energy of the system as the integral of the energy-density $T_{00} = \rho$ over a spatial constant time $x^0 =$ const. slice Σ,

$$E = \int_\Sigma d^3x \; T_{00} \tag{23.25}$$

(another way to see that we can take the integration measure to be the flat Euclidean measure d^3x is to recall that T_{00} is small by assumption). If the linearised Einstein equations are satisfied, we can express the total energy in terms of the Einstein tensor,

$$E = \frac{1}{8\pi G_N} \int_\Sigma d^3x \; G^{(1)}_{00} \ . \tag{23.26}$$

Now we have two distinct expressions for $G^{(1)}_{00}$ in terms of the derivatives of the metric and, interestingly, both of them are spatial total derivatives, namely

$$G^{(1)}_{00} = -\tfrac{1}{2}\Delta(\delta^{ik}h_{ik}) + \tfrac{1}{2}\partial_i\partial_k h^{ik} = \partial_i(-\tfrac{1}{2}\partial^i h^k_{\ k} + \tfrac{1}{2}\partial_k h^{ik}) \tag{23.27}$$

and

$$G^{(1)}_{00} = -\Delta h_{00} = -\partial_i(\partial^i h_{00}) \ . \tag{23.28}$$

This allows us to write the total energy E of the system as a boundary integral over the boundary

$$\partial\Sigma = S^2_\infty \tag{23.29}$$

of the spatial slice "at infinity" in two different ways, namely as

$$E = \frac{1}{16\pi G_N} \oint_{S^2_\infty} dS_i \; (\partial_k h^{ik} - \partial^i h^k_{\ k}) \tag{23.30}$$

or as

$$E = -\frac{1}{8\pi G_N} \oint_{S^2_\infty} dS_i \; \partial^i h_{00} \ . \tag{23.31}$$

Note that these final expressions only depend on the asymptotics of the gravitational field at spatial infinity (the assumption that the gravitational field is weak there translating into the statement that the metric is asymptotically flat). It is thus tempting to adopt E as the definition of the total energy of an isolated (i.e. asymptotically flat) system in general, even when the field in the interior is not weak:

1. Thus in general the *ADM Energy* of such a system is defined by

$$E_{ADM} = \frac{1}{16\pi G_N} \oint_{S^2_\infty} dS_i \, (\partial_k h_{ik} - \partial_i h_{kk}) \quad , \tag{23.32}$$

this expression relying on a Cartesian coordinate system at infinity (which is why we can permit ourselves to be careless with the positioning of the indices - repeated indices are to be summed, as usual). Somewhat more covariantly one can write this as

$$E_{ADM} = \frac{1}{16\pi G_N} \oint_{S^2_\infty} dS_i \, (\bar{\nabla}^k h_{ik} - \bar{\nabla}_i h) \quad , \tag{23.33}$$

where $\bar{\nabla}$ is the covariant derivative with respect to a background metric $\bar{g}_{ik}$,

$$h_{ik} = g_{ik} - \bar{g}_{ik} \quad , \tag{23.34}$$

and

$$h = \bar{g}^{ik} h_{ik} \tag{23.35}$$

is the trace of h_{ik} with respect to the background metric. If the background metric is the Minkowski metric, *and* one works in Cartesian coordinates, then (23.33) reduces to (23.32).

The justification for the identification of this quantity with the energy is substantially strengthened by the canonical (Hamiltonian) ADM analysis of general relativity discussed in section 21, in particular in section 21.12. Indeed, in section 21.12 we deduced another candidate expression for the energy from the boundary term in the ADM Hamiltonian, namely (21.191)

$$E_{ADM} = -\frac{1}{8\pi G_N} \oint_{S_\infty} d^2x \, \sqrt{s}(k_S - k^0_S) \tag{23.36}$$

At first sight, beyond the fact that both are expressed as integrals over a 2-sphere at infinity and that both depend in some way on a choice of reference (background) metric, these two expresssions appear to have little in common. However, it is possible to show that (23.33) is in fact equal to (23.36) provided that the metrics on S_∞ induced by the metrics $\bar{g}_{ik}$ (appearing in (23.33)) and g^0_{ik} (implicitly appearing in (23.36)) are the same. The proof of this statement I am aware of relies on a convenient choice of coordinate system and is therefore not particularly insightful *per se* and will not be given here.[64]

2. The second expression is actually a special case of the Komar charges briefly mentioned in section 13.7, here applied to the (asymptotic) timelike Killing vector ∂_t (the generator of time-translations, and thus naturally associated with the

[64]For this proof, see e.g. S. Hawking, G. Horowitz, *The Gravitational Hamiltonian, Action, Entropy, and Surface Terms*, `arXiv:gr-qc/9501014`, or Exercise 4.5.7 of E. Poisson, *A Relativist's Toolkit* (modelled on the cited article).

energy or Hamiltonian). Indeed, recall that for any Killing vector K^μ we had the conserved current $J^\mu = R^\mu_{\ \nu} K^\nu$ (13.49), which was itself a divergence of an anti-symmetric tensor, $J^\mu = \nabla_\nu A^{\mu\nu}$, with $A^{\mu\nu} = -A^{\nu\mu} = \nabla^\mu K^\nu$. Thus the corresponding conserved charge $Q_K(V)$ contained in a volume V can be written as a surface integral of $\nabla^\mu K^\nu$ over the boundary ∂V of the volume V,

$$Q_K(V) \sim \oint_{\partial V} dS_{\mu\nu} \nabla^\mu K^\nu \quad . \tag{23.37}$$

Applying this to $V = \Sigma$, $\partial V = S^2_\infty$, and $K = \partial_t$, we find (asymptotically)

$$\begin{aligned}
\nabla^i K^0 &= \nabla_i K^0 = \Gamma^0_{i\alpha} K^\alpha = \Gamma^0_{i0} = -\Gamma_{0i0} = -\tfrac{1}{2}\partial_i h_{00} \\
\nabla^0 K^i &= -\nabla_0 K^i = -\Gamma^i_{0\alpha} K^\alpha = -\Gamma^i_{00} = -\Gamma_{i00} = +\tfrac{1}{2}\partial_i h_{00} \quad .
\end{aligned} \tag{23.38}$$

Thus the second expression (23.31) can be written as

$$\begin{aligned}
E &= -\frac{1}{8\pi G_N} \oint_{S^2_\infty} dS_i\ \partial_i h_{00} = -\frac{1}{4\pi G_N} \oint_{S^2_\infty} dS_{0i}\ \nabla^0 K^i \\
&= -\frac{1}{8\pi G_N} \oint_{S^2_\infty} dS_{\mu\nu}\ \nabla^\mu K^\nu \quad .
\end{aligned} \tag{23.39}$$

This fixes the normalisation of the Komar charge $Q_K(V)$ (23.37) in this case,

$$E_{\text{Komar}}(V) \equiv Q_{K=\partial_t}(V) = -\frac{1}{8\pi G_N} \oint_{\partial V} dS_{\mu\nu}\ \nabla^\mu K^\nu \tag{23.40}$$

and we can identify (23.31) with the total Komar energy ($V = \Sigma$) associated to the Killing vector $K = \partial_t$,

$$E_{\text{Komar}}(\Sigma) = -\frac{1}{8\pi G_N} \oint_{S^2_\infty} dS_i\ \partial_i h_{00} = -\frac{1}{8\pi G_N} \oint_{S^2_\infty} dS_{\mu\nu}\ \nabla^\mu K^\nu \quad . \tag{23.41}$$

As our primary litmus test, in section 24.8 we will apply these expressions to the Schwarzschild metric, the unique spherically symmetric asymptotically flat solution of the vacuum Einstein equations. In particular, it therefore describes the gravitational field in the exterior of a star, and it depends on a single parameter m, related to the mass of the star by $m = G_N M/c^2$, and it is then of interest to see what the ADM and Komar energies have to say about this. Suffice it to say here that reassuringly one indeed finds $E_{ADM} = E_{\text{Komar}} = M$.

In the same way one can also introduce various notions of momentum or angular momentum of an isolated system (the latter for instance being non-zero for the Kerr metric (30.3) describing a rotating star or black hole), but we will not pursue this here.

23.5 Wave Equations and Gauge Conditions in Maxwell Theory

We will now abandon the assumption of static non-relativistic fields and return to the general weak-field linearised Einstein equations (23.18). In order to understand how to proceed from there, in this section we will briefly recall in a condensed way the analogous steps in the case of Maxwell theory.

1. The Maxwell equations (in a Minkowski background) are already linear (no need to linearise) and read
$$\partial^\alpha F_{\alpha\beta} = -J_\beta \ . \tag{23.42}$$
In terms of the potentials A_α these equations are
$$\Box A_\beta - \partial_\beta(\partial_\alpha A^\alpha) = -J_\beta \ . \tag{23.43}$$
These are to be thought of as the analogue of the linearised Einstein equations (23.18), with the linearised Einstein tensor given in (23.8). These are wave equations (with a source) for the components A_α, but due to the second term on the left-hand side the individual components of the potential are not yet completely decoupled in this equation.

2. Maxwell theory has the gauge invariance
$$A_\alpha \to A_\alpha + \partial_\alpha V \tag{23.44}$$
(in particular, the field strenghts $F_{\alpha\beta}$ are invariant) which allows one to choose for instance the Loren(t)z gauge condition[65]
$$\partial_\alpha A^\alpha = 0 \ . \tag{23.45}$$
Indeed, if V is chosen to satisfy the differential equation
$$\Box V = -\partial_\alpha A^\alpha \ , \tag{23.46}$$
then
$$\partial^\alpha(A_\alpha + \partial_\alpha V) = 0 \ . \tag{23.47}$$
With this choice of gauge the Maxwell equations become decoupled wave equations,
$$\Box A_\beta = -J_\beta \ , \tag{23.48}$$
and can now be straightforwardly solved in terms of Green functions etc. One particular solution to the inhomogeneous equation is the retarded potential
$$A_\alpha(t,\vec{x}) = (4\pi)^{-1}\int d^3x' \, \frac{J_\alpha(t-|\vec{x}-\vec{x}'|,\vec{x}')}{|\vec{x}-\vec{x}'|} \ . \tag{23.49}$$
Note that, as a consequence of $\partial_\alpha J^\alpha = 0$, this solution is automatically in the Lorenz gauge. As usual, the general solution is then a sum of this particular solution of the inhomogeneous equation and the general solution of the homogeneous equation.

[65] Credit for this should really go the Danish physicist Ludvig V. Lorenz (who used this condition in its Lorent(!)z non-invariant form already in 1867, just 3 years after the publicaton of Maxwell's treatise) rather than the more famous Dutch physicist Hendrik A. Lorentz. See J. Jackson's *Electrodynamics* (note at the end of chapter 6).

3. The homogeneous equation, i.e. the vacuum Maxwell equations, can now be solved in terms of plane waves,

$$A_\alpha = \epsilon_\alpha \mathrm{e}^{ik_\beta x^\beta} \quad \text{with} \quad k_\beta k^\beta = 0 \tag{23.50}$$

(or wave packets constructed from them), and the Lorenz gauge constrains the *polarisation vector* ϵ_α by

$$k^\alpha \epsilon_\alpha = 0 \ . \tag{23.51}$$

For a wave travelling in the x^3-direction, this implies for the polarisation vector

$$k^\alpha = (\omega, 0, 0, k^3 = \omega) \quad \Rightarrow \quad \epsilon_3 = -\epsilon_0 \ . \tag{23.52}$$

4. One still has the residual gauge invariance

$$A'_\alpha = A_\alpha + \partial_\alpha \xi \quad , \quad \Box \xi = 0 \tag{23.53}$$

leaving the Lorenz gauge condition invariant. Choosing a particular solution of this homegenous wave equation, say

$$\xi = \xi_0 \mathrm{e}^{ik_\beta x^\beta} \tag{23.54}$$

now completely fixes this residual gauge invariance. Under this residual gauge transformation the polarisation vector transforms as

$$\epsilon'_\alpha = \epsilon_\alpha + i\xi_0 k_\alpha \ , \tag{23.55}$$

in particular

$$\epsilon'_0 = \epsilon_0 - i\xi_0 \omega \ . \tag{23.56}$$

Thus choosing $\xi_0 = \epsilon_0 / i\omega$, one has $\epsilon'_0 = 0$, implying $\epsilon'_3 = 0$, and one is left with the polarisation vector

$$\epsilon^\alpha = (0, \epsilon^1, \epsilon^2, 0) \tag{23.57}$$

displaying the two physical transverse polarisations of an electromagnetic wave.

23.6 Linearised Gravity: Gauge Invariance and Coordinate Choices

We now proceed analogously in the case of linearised gravity. First of all we need to understand the gauge invariance (or the counterpart of gauge invariance) in the present case.

The original non-linear Einstein equations have a local invariance consisting of general coordinate transformations. What remains of general coordinate invariance in the linearised approximation are, naturally, linearised general coordinate transformations. Indeed, $h_{\mu\nu}$ and

$$h'_{\mu\nu} = h_{\mu\nu} + L_V \eta_{\mu\nu} \tag{23.58}$$

represent the same physical perturbation because $\eta_{\mu\nu} + L_V \eta_{\mu\nu}$ is just an infinitesimal coordinate transform of the Minkowski metric $\eta_{\mu\nu}$. Therefore linearised gravity has the gauge freedom

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu V_\nu + \partial_\nu V_\mu \ . \tag{23.59}$$

For example, the linearised Riemann tensor $R^{(1)}_{\mu\nu\rho\sigma}$ is, rather obviously, invariant under this transformation. As a consequence, in particular also the Einstein tensor, and hence the vacuum linearised Einstein equations, have this local gauge invariance. This can be understood without any calculation by noting that under the variation

$$\delta_V h_{\mu\nu} = L_V \eta_{\mu\nu} \tag{23.60}$$

the linearised Riemann tensor transforms into the Riemann tensor for the Minkowski metric,

$$\delta_V R^{(1)}_{\mu\nu\rho\sigma} = L_V R_{\mu\nu\rho\sigma}(\eta_{\alpha\beta}) = 0 \ . \tag{23.61}$$

However, in contrast to the Lagrangian of Maxwell theory, say, the Lagrangian (23.12) for linearised gravity is *not* strictly gauge invariant, but only invariant up to a total derivative.

Given this gauge invariance of the linearised theory, for explicit calculations it is useful to make a particular gauge choice and, as in Maxwell theory, a good choice of gauge can simplify things considerably (and a bad choice of gauge can have the opposite effect).

It turns out that, in general, a very useful gauge condition is

$$g^{\mu\nu}\Gamma^\rho_{\ \mu\nu} = 0 \ . \tag{23.62}$$

It is called the *harmonic gauge condition*, or Fock, or de Donder gauge condition (even though harmonic coordinates were used extensively by Einstein himself until just before the discovery of the final, truly generally covariant, formulation of general relativity). The name harmonic derives from the fact that in this gauge the coordinate functions x^μ are harmonic:

$$\Box x^\mu \equiv g^{\nu\rho}\nabla_\rho \partial_\nu x^\mu = -g^{\nu\rho}\Gamma^\mu_{\ \nu\rho} \ , \tag{23.63}$$

and thus

$$\Box x^\mu = 0 \quad \Leftrightarrow \quad g^{\nu\rho}\Gamma^\mu_{\ \nu\rho} = 0 \ . \tag{23.64}$$

It is the analogue of the Lorenz gauge $\partial_\mu A^\mu = 0$ in Maxwell theory. Moreover, in flat space Cartesian coordinates are obviously harmonic, and in general harmonic coordinates are (like geodesic coordinates) a nice and useful curved space counterpart of Cartesian coordinates.

In the linearised theory, this gauge condition becomes

$$\partial_\mu h^\mu_{\ \lambda} - \tfrac{1}{2}\partial_\lambda h = 0 \ . \tag{23.65}$$

The gauge parameter V_μ which will achieve this is the solution to the equation

$$\Box V_\lambda = -(\partial_\mu h^\mu_{\ \lambda} - \tfrac{1}{2}\partial_\lambda h) \ . \tag{23.66}$$

Indeed, with this choice one has

$$\partial_\mu(h^\mu_{\ \lambda} + \partial^\mu V_\lambda + \partial_\lambda V^\mu) - \tfrac{1}{2}\partial_\lambda(h + 2\partial^\mu V_\mu) = 0 \ . \tag{23.67}$$

Note for later that, as in Maxwell theory, this gauge choice does not necessarily fix the gauge completely. Any transformation $x^\mu \to x^\mu + \xi^\mu$ with $\Box\xi^\mu = 0$ will leave the harmonic gauge condition invariant.

23.7 Wave Equation

Now let us use this gauge condition in the linearised Einstein equations. In this gauge they simplify somewhat to

$$\Box h_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\Box h = -16\pi G_N T^{(0)}_{\mu\nu} \ . \tag{23.68}$$

In particular, the vacuum equations (or the equations in a source-free region of space-time) are just

$$T^{(0)}_{\mu\nu} = 0 \ \ \Rightarrow \ \ \Box h_{\mu\nu} = 0 \ , \tag{23.69}$$

which is the standarad relativistic wave equation. Together, the equations

$$\begin{aligned} \Box h_{\mu\nu} &= 0 \\ \partial_\mu h^\mu_{\ \lambda} - \tfrac{1}{2}\partial_\lambda h &= 0 \end{aligned} \tag{23.70}$$

determine the evolution of a disturbance in a gravitational field in vacuum in the harmonic gauge.

It is often convenient to consider the linear combination

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}h \ . \tag{23.71}$$

This is also commonly known as the *trace reversed perturbation*, because in 4 space-time dimensions (but only there) one has

$$\bar{h}^\mu_{\ \mu} = -h^\mu_{\ \mu} \ . \tag{23.72}$$

Note, as an aside, that with this notation and terminology the Einstein tensor (again in 4 space-time dimensions only) is the trace reversed Ricci tensor,

$$\bar{R}_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = G_{\mu\nu} \ . \tag{23.73}$$

In terms of $\bar{h}_{\mu\nu}$, the linearised Einstein equations and the harmonic gauge condition (in any dimension) are just

$$
\begin{aligned}
\Box \bar{h}_{\mu\nu} &= -16\pi G_N T^{(0)}_{\mu\nu} \\
\partial_\mu \bar{h}^\mu_{\ \nu} &= 0 \ .
\end{aligned}
\tag{23.74}
$$

This way of writing the linearised Einstein equations sharpens the analogy with the Maxwell equations in the Lorenz gauge:

- the Lorenz gauge $\partial_\mu A^\mu = 0$ decouples the Maxwell equations for A_μ which in this gauge read $\Box A_\mu = -j_\mu$
- the gauge condition $\partial_\mu \bar{h}^\mu_{\ \nu} = 0$ decouples the linearised Einstein equations for the variables $\bar{h}_{\mu\nu}$ which in this gauge read $\Box \bar{h}_{\mu\nu} = -16\pi G_N T^{(0)}_{\mu\nu}$.

One solution is, of course, the retarded solution

$$
\bar{h}_{\mu\nu}(t,\vec{x}) = 4G_N \int d^3x' \, \frac{T^{(0)}_{\mu\nu}(t - |\vec{x}-\vec{x}'|, \vec{x}')}{|\vec{x}-\vec{x}'|} \quad . \tag{23.75}
$$

Note that, as a consequence of $\partial_\mu T^{(0)\mu\nu} = 0$, this solution is automatically in the harmonic gauge. As usual, the general solution is then a sum of this particular solution of the inhomogeneous equation and the general solution of the homogeneous equation.

23.8 Polarisation Tensor and the Metric of a Gravitational Wave

The homogeneous equation is the linearised vacuum Einstein equation in the harmonic gauge,

$$
\Box \bar{h}_{\mu\nu} = 0 \ . \tag{23.76}
$$

This is clearly solved by

$$
\bar{h}_{\mu\nu} = \epsilon_{\mu\nu} \mathrm{e}^{ik_\alpha x^\alpha} \ , \tag{23.77}
$$

where $\epsilon_{\mu\nu}$ is a constant, symmetric *polarisation tensor* and k^α is a constant *wave vector*, provided that k^α is null, $k^\alpha k_\alpha = 0$. In order to obtain real metrics from this one should of course use real linear combinations of such solutions in the end.

Thus plane waves are solutions to the linearised equations of motion and the Einstein equations predict the existence of gravitational waves travelling along null geodesics (at the speed of light). The timelike component of the wave vector is often referred to as the *frequency* ω of the wave, and we can write $k^\mu = (\omega, k^i)$. Plane waves are of course not the most general solutions to the wave equations but any solution can be written as a superposition of plane wave solutions (wave packets).

So far, we have ten parameters $\epsilon_{\mu\nu}$ and four parameters k^μ to specify the wave, but many of these are spurious, i.e. can be eliminated by using the freedom to perform linearised coordinate transformations and Lorentz rotations.

First of all, the harmonic gauge condition implies that

$$\partial_\mu \bar{h}^\mu_{\ \nu} = 0 \quad \Rightarrow \quad k^\mu \epsilon_{\mu\nu} = 0 \ . \tag{23.78}$$

Now we can make use of the residual gauge freedom $x^\mu \to x^\mu + \xi^\mu$ with $\Box \xi^\mu = 0$ to impose further conditions on the polarisation tensor. Since this is a wave equation for ξ^μ, once we have specified a solution for ξ^μ we will have fixed the gauge completely. Alternatively, note that under $h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu V_\nu + \partial_\nu V_\mu$ one has

$$\begin{aligned} &\bar{h}_{\mu\nu} \to \bar{h}_{\mu\nu} + \partial_\mu V_\nu + \partial_\nu V_\mu - \eta_{\mu\nu}\partial_\lambda V^\lambda \\ \Rightarrow \quad &\partial^\mu \bar{h}_{\mu\nu} \to \partial^\mu \bar{h}_{\mu\nu} + \Box V_\nu \end{aligned} \tag{23.79}$$

so that the gauge condition is invariant precisely under linearised coordinate transformations with $\Box V_\mu = 0$. Taking the solution of this equation to be of the form

$$V_\mu = v_\mu \mathrm{e}^{ik_\alpha x^\alpha} \ , \tag{23.80}$$

the polarisation tensor transforms according to

$$\epsilon_{\alpha\beta} \to \epsilon_{\alpha\beta} + i(k_\alpha v_\beta + k_\beta v_\alpha) - i\eta_{\alpha\beta} k^\gamma v_\gamma \ . \tag{23.81}$$

One can now choose the v_μ in such a way (see the example below) that the new polarisation tensor satisfies $k^\mu \epsilon_{\mu\nu} = 0$ (as before) as well as

$$\epsilon_{\mu 0} = \epsilon^\mu_{\ \mu} = 0 \ . \tag{23.82}$$

All in all, we appear to have nine conditions on the polarisation tensor $\epsilon_{\mu\nu}$ but as both (23.78) and the first of (23.82) imply $k^\mu \epsilon_{\mu 0} = 0$, only eight of these are independent. Therefore, there are two independent polarisations for a gravitational wave.

Together with $k^\mu \epsilon_{\mu\nu} = 0$, these gauge conditions thus impose the conditions

$$k^\mu \bar{h}_{\mu\nu} = 0 \quad , \quad \bar{h}_{\mu 0} = 0 \quad , \quad \bar{h}^\mu_{\ \mu} = 0 \ . \tag{23.83}$$

This is known as the *transverse traceless gauge*, and a field satisfying this gauge is frequently denoted by $\bar{h}^{TT}_{\mu\nu}$.

For example, let us consider a wave travelling in the x^3-direction,

$$k^\mu = (\omega, 0, 0, k^3) = (\omega, 0, 0, \omega) \ . \tag{23.84}$$

Then

- the condition $k^\mu \epsilon_{\mu\nu} = 0$ becomes $\epsilon_{3\nu} = -\epsilon_{0\nu}$

- the condition $\epsilon_{00} = 0$ determines $v_0 - v_3$, and also implies $\epsilon_{30} = 0$
- the other linear combination $v_0 + v_3 \sim k^\alpha v_\alpha$ can be used to achieve $\epsilon^\mu_{\ \mu} = 0$
- the conditions $\epsilon_{10} = \epsilon_{20} = 0$ determine v_1 and v_2

Therefore, the only independent components are ϵ_{ab} with $a, b = 1, 2$. As ϵ_{ab} is symmetric and traceless, this wave is completely characterised by $\epsilon_{11} = -\epsilon_{22}$, $\epsilon_{12} = \epsilon_{21}$ and the frequency ω.

Now we should not forget that, when talking about the polarisation tensor of a gravitational wave, we are actually talking about the space-time metric itself. Namely, since for a traceless perturbation we have

$$\epsilon^\mu_{\ \mu} = 0 \quad \Rightarrow \quad \bar{h}_{\alpha\beta} = h_{\alpha\beta} \quad , \tag{23.85}$$

we have deduced that the metric describing a gravitational wave travelling in the x^3-direction can always be put into the form

$$ds^2 = -dt^2 + (\delta_{ab} + h_{ab})dx^a dx^b + (dx^3)^2 \quad , \tag{23.86}$$

with $h_{ab} = h_{ab}(t \mp x^3)$ symmetric and traceless. This neatly encodes and describes the distortion of the space-time geometry in the directions transverse to the gravitational wave.

<u>Remarks:</u>

1. Introducing null coordinates

$$u = t - x^3 \quad , \quad v = t + x^3 \quad , \tag{23.87}$$

the gravitational wave metric can be written as

$$ds^2 = -du\, dv + (\delta_{ab} + h_{ab}(u))dx^a dx^b \quad . \tag{23.88}$$

Explicitly, with $x^a = (x, y)$, the linearised perturbation can be written as

$$h_{ab}(u)dx^a dx^b = h_+(u)(dx^2 - dy^2) + 2h_\times(u)dxdy \quad , \tag{23.89}$$

with two arbitrary functions $h_{11}(u) \equiv h_+(u)$ and $h_{12}(u) \equiv h_\times(u)$ (see the next section 23.9 for an explanation of this notation).

2. This analysis was rather evidently independent of the dimension. In D dimensions the polarisation states of a graviton are described by a symmetric, transverse, and traceless tensor h_{ab} where $a, b = 1, \ldots, D-2$. Thus the number of physical polarisation states of a graviton in D dimensions are

$$\#[h_{ab}] = \frac{(D-2)(D-1)}{2} - 1 = \frac{D(D-3)}{2} \quad . \tag{23.90}$$

Note that this gives zero for $D = 3$, in agreement with the fact, noted before, that in 3 space-time dimensions there is no gravitational vacuum dynamics since vanishing of the Ricci tensor is equivalent to vanishing of the Riemann tensor.

3. By abandoning the assumption that $h_{ab}(u)$ be small, one can look for (and easily find) solutions of the full non-linear Einstein equations of the form

$$ds^2 = -du\, dv + g_{ab}(u)dx^a dx^b \quad . \tag{23.91}$$

These are known as exact gravitational plane waves, and are discussed in some detail in section 43.

23.9 Physical Effects of Gravitational Waves

To determine the physical effect of a gravitational wave racing by, we cannot just look at the gravitational field (23.86) at a point (by the equivalence principle), i.e. we cannot detect the presence of such a wave (ultra-)locally. However, we can consider its influence on the relative motion of nearby particles. In other words, we look at the geodesic deviation equation (8.45).

Consider a family of nearby particles described by the *velocity field* $u^\mu(x)$ and separation (deviation) vector $S^\mu(x)$. Then the change of the deviation vector along the flow lines of the velocity field is determined by

$$(D_\tau)^2 S^\mu = R^\mu_{\ \nu\rho\sigma} u^\nu u^\rho S^\sigma \quad . \tag{23.92}$$

We consider the situation where the test particles are initially, in the absence of the gravitational wave, at rest, $u^\mu = (1, 0, 0, 0)$. Then the gravitational wave will lead, to lowest order in the perturbation $h_{\mu\nu}$, to a 4-velocity

$$u^\mu = (1, 0, 0, 0) + \mathcal{O}(h) \quad . \tag{23.93}$$

However, because the Riemann tensor is already of order h, to lowest order the right hand side of the geodesic deviation equation reduces to

$$R^{(1)}_{\mu 00\sigma} = \tfrac{1}{2}\partial_0\partial_0 h_{\mu\sigma} \tag{23.94}$$

(because $h_{0\mu} = 0$). On the other hand, to lowest order the left hand side is just the ordinary time derivative. Thus the geodesic deviation equation becomes (an overdot denoting a t-derivative)

$$\ddot{S}^\mu = \tfrac{1}{2}\ddot{h}^\mu_{\ \sigma}\, S^\sigma \quad . \tag{23.95}$$

In particular, we see immediately that the gravitational wave is *transversally polarised*, i.e. the component S^3 of S^μ in the longitudinal direction of the wave is unaffected and

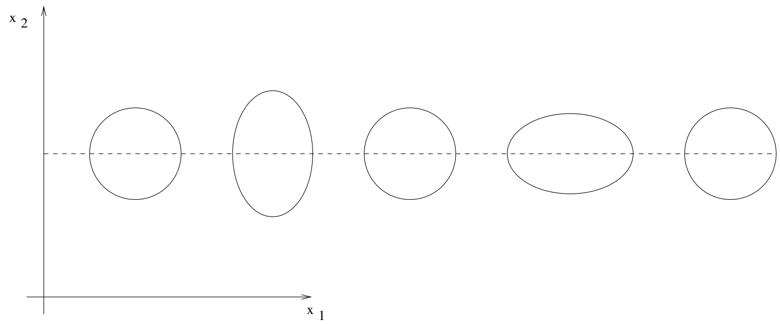


Figure 9: Effect of a gravitational wave with polarisation ϵ_{11} moving in the x^3-direction, on a ring of test particles in the $x^1 - x^2$-plane.

the particles are only disturbed in directions perpendicular to the wave. The movement of the particles in the 1-2 plane is then governed by

$$\ddot{S}^a = \tfrac{1}{2}\ddot{h}^a_{\ b}\, S^b \equiv -(\Omega^2)^a_{\ b}\, S^b \quad , \tag{23.96}$$

which is the equation of a 2-dimensional time-dependent harmonic oscillator, giving rise to characteristic oscillating movements of the test particles in the 1-2 plane. With

$$h^a_{\ b} = \mathrm{e}^{-i\omega(t - x^3)}\epsilon^a_{\ b} \quad , \tag{23.97}$$

as above, one has

$$(\Omega^2)^a_{\ b} = \tfrac{1}{2}\omega^2 h^a_{\ b} \quad , \tag{23.98}$$

and we can consider separately the two cases (1) $\epsilon_{12} = 0$ and (2) $\epsilon_{11} = -\epsilon_{22} = 0$.

1. For $\epsilon_{12} = 0$ one has

$$\begin{aligned} \ddot{S}^1(t) &= -\tfrac{1}{2}\epsilon_{11}\omega^2 \mathrm{e}^{-i\omega t} S^1(t)\, \mathrm{e}^{i\omega x^3} \\ \ddot{S}^2(t) &= +\tfrac{1}{2}\epsilon_{11}\omega^2 \mathrm{e}^{-i\omega t} S^2(t)\, \mathrm{e}^{i\omega x^3} \end{aligned} \tag{23.99}$$

   Recalling that ϵ_{ab} is small, the solution to lowest order is simply

$$\begin{aligned} S^1(t) &= (1 + \tfrac{1}{2}\epsilon_{11}\mathrm{e}^{-i\omega(t - x^3)})S^1(0) \\ S^2(t) &= (1 - \tfrac{1}{2}\epsilon_{11}\mathrm{e}^{-i\omega(t - x^3)})S^2(0) \end{aligned} \tag{23.100}$$

   Given the interpretation of S^μ as a separation vector, this means that particles originally separated in the x^1-direction will oscillate back and forth in the x^1-direction and likewise for x^2. A nice (and classical) way to visualise this (see Figure 9) is to start off with a ring of particles in the $1 - 2$ plane. As the wave passes by the particles will start bouncing in such a way that the ring bounces in the shape of a cross $+$. For this reason, ϵ_{11} is also frequently written as ϵ_+.

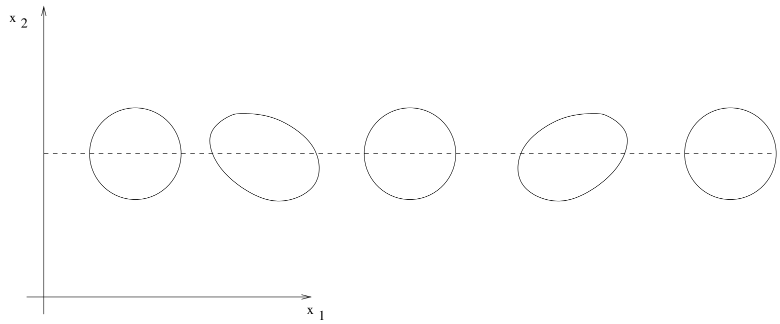


Figure 10: Effect of a gravitational wave with polarisation ϵ_{12} moving in the x^3-direction, on a ring of test particles in the $x^1 - x^2$-plane.

2. If, on the other hand, $\epsilon_{11} = 0$ but $\epsilon_{12} = \epsilon_{21} \neq 0$, then the lowest order solution is

$$\begin{aligned} S^1(t) &= S^1(0) + \tfrac{1}{2}\epsilon_{12}\mathrm{e}^{-i\omega(t-x^3)} S^2(0) \\ S^2(t) &= S^2(0) + \tfrac{1}{2}\epsilon_{12}\mathrm{e}^{-i\omega(t-x^3)} S^1(0) \ . \end{aligned} \tag{23.101}$$

   This time the deplacement in the x^1-direction is governed by the original deplacement in the x^2-direction and vice-versa, and the ring of particles will bounce in the shape of a $\times$ ($\epsilon_{12} = \epsilon_\times$) - see Figure 10.

3. Of course, one can also construct circularly polarised waves by using

$$\epsilon_{R,L} = \frac{1}{\sqrt{2}}(\epsilon_{11} \pm i\epsilon_{12}) \ . \tag{23.102}$$

These solutions display the characteristic behaviour of *quadrupole radiation*, and this is something that we might have anticipated on general grounds. First of all, we know from Birkhoff's theorem (see section 24.6) that a spherically symmetric vacuum solution of the Einstein field equations is necessarily static. Thus there can be no radial oscillations, and thus no monopole (s-wave) radiation. Moreover, dipole radiation is due to oscillations of the center of charge. While this is certainly possible for electric charges, an oscillation of the center of mass would violate momentum conservation and is therefore ruled out. Thus the lowest possible mode of gravitational radiation is quadrupole radiation, just as we have found.

23.10 Brief Comments on Production and Energy of Gravitational Waves

Now that we have found the solutions to the vacuum equations, we should include sources and study the production of gravitational waves, characterise the type of radiation that is emitted, estimate the radiated energy etc. However, this is quite a delicate and both technically and conceptually quite challenging subject, and I will just develop this to the extent that the quadrupole property of the radiation becomes plausible.[66]

In order to study the production of gravitational waves, we need to include sources, i.e. we need to go back to the retarded solution (23.75)

$$\bar{h}_{\mu\nu}(t,\vec{x}) = 4G_N \int d^3y \, \frac{T^{(0)}_{\mu\nu}(t - |\vec{x}-\vec{y}|, \vec{y})}{|\vec{x}-\vec{y}|} \quad . \tag{23.103}$$

At large distances, and if the source does not oscillate too rapidly (the wavelength should be much larger than the size of the source), one can approximate this by

$$\bar{h}_{\mu\nu}(t,\vec{x}) \approx \frac{4G_N}{r} \int d^3y \, T^{ret}_{\mu\nu}(t,\vec{y}) \quad , \tag{23.104}$$

where $r = |\vec{x}|$ and the retarded source is

$$T^{ret}_{\mu\nu}(t,\vec{y}) = T^{(0)}_{\mu\nu}(t-r,\vec{y}) \quad . \tag{23.105}$$

This is the gravitational analogue of the dipole approximation to the multipole expansion in electrodynamics (and, as we will see, here this turns out to be a quadrupole approximation).

Next, since $T^{(0)}_{\mu\nu}$ is conserved, also

$$\bar{h}_{\mu 0} \sim \int d^3y \, T^{(ret)}_{\mu 0} \tag{23.106}$$

is conserved, i.e. time-independent. Therefore, in this approximation the leading $(1/r)$-part of $\bar{h}_{\mu 0}$ will not lead to gravitational waves. We can thus concentrate on the spatial components

$$\bar{h}_{ik}(t,\vec{x}) \approx \frac{4G_N}{r} \int d^3y \, T^{ret}_{ik}(t,\vec{y}) \quad , \tag{23.107}$$

Using the Laue theorem (7.53),

$$\int d^3x \, T^{ik} = \tfrac{1}{2}(\partial_0)^2 \int d^3x \, T_{00} x^i x^k \equiv \tfrac{1}{2} \ddot{Q}^{ik} \quad , \tag{23.108}$$

we thus have

$$\bar{h}_{ik}(t,\vec{x}) \approx \frac{2G_N}{r} \ddot{Q}^{ret}_{ik} \quad , \tag{23.109}$$

where (7.54)

$$Q^{ret}_{ik}(t) = \int d^3x \, \rho^{ret} x_i x_k \tag{23.110}$$

is the quadrupole moment tensor of the retarded energy density $T^{ret}_{00} = \rho^{ret}$. Thus, if the source has a time-dependence

$$\rho(t) \sim \mathrm{e}^{-i\Omega t} \quad , \tag{23.111}$$

[66] See e.g. S. Weinberg, *Gravitation and Cosmology*, B. Schutz, *A first course in general relativity*, or S. Carroll, *Spacetime and Geometry*, and of course C. Misner, K. Thorne, J. Wheeler, *Gravitation* for detailed discussionss.

say (of course, one should in the end take real superpositions of such modes), then

$$\bar{h}_{ik}(t,r) \approx -2G_N\Omega^2 Q^{ret}_{ik}\,\frac{\mathrm{e}^{-i\Omega(t-r)}}{r} \tag{23.112}$$

clearly describes an outgoing spherical wave.

As noted before, the retarded solution is automatically in the harmonic gauge, but it is not yet in the transverse traceless gauge. Transforming the above solution to the transverse traceless gauge, one finds that the (transverse, traceless) components

$$\bar{h}^{TT}_{ab} = h^{TT}_{ab} \tag{23.113}$$

are naturally expressed not in terms of the quadrupole moments Q_{ik} but in terms of the so-called "reduced" (traceless) quadrupole moments

$$\begin{aligned}\mathcal{Q}^{ret}_{ik} &= \int d^3x\;\rho^{ret}(x_ix_k - \tfrac{1}{3}\delta_{ik}r^2)\\ &= Q^{ret}_{ik} - \tfrac{1}{3}\delta_{ik}(Q^{ret})^j_{\ j}\quad .\end{aligned} \tag{23.114}$$

These formulae can now in principle be applied to various specific situations of interest by specifying the source term appropriately.

Finally, one quantity of particular interest is of course the energy radiated away by the source. However, as discussed in section 22.6, the notion of "gravitational energy" or "energy of the gravitational field" is not in general well defined and raises numerous conceptual and technical issues. One might perhaps have hoped that these issues can be completely side-stepped in the linearised theory we are dealing with here, which is after all much more like a standard field theory in Minkowski space. And indeed, several strategies are available, and they all lead to expressions for the energy-density which are of the standard form "quadratic in the derivatives of the fields". For example one can

- proceed by analogy with quadrupole radiation in Maxwell theory,
- use the (Belinfante-improved) Noether energy-momentum tensor of the quadratic Fierz-Pauli action (23.12),
- expand the Einstein equations not only to linear but to quadratic order in the fluctuations $h_{\mu\nu}$ and interpret the quadratic terms as the gravitational contribution to the energy-momentum tensor,
- ...

E.g. in the transverse traceless gauge the Fierz-Pauli action reduces to a "standard" quadratic action

$$\mathcal{L} = -\tfrac{1}{2}(\eta^{\alpha\beta}\partial_\alpha h_{11}\partial_\beta h_{11} + \eta^{\alpha\beta}\partial_\alpha h_{12}\partial_\beta h_{12})\quad . \tag{23.115}$$

For $h_{ab} = h_{ab}(t - x^3)$, say, this gives rise to an energy density and energy flux

$$\Theta_{00} = -\Theta_{03} = (\dot{h}_{11})^2 + (\dot{h}_{12})^2 \sim r^{-2}(\dddot{Q}^{ret})^2 \quad . \tag{23.116}$$

On the basis of such considerations one might expect or anticipate the total radiated energy to be proportional to something like

$$dE/dt \sim -(\dddot{Q}^{ret})^2 \quad . \tag{23.117}$$

However, dealing with quadratic terms in a linearised theory is somewhat dodgy and not strictly speaking internally consistent. As a consequence, equally plausible strategies may not necessarily lead to equivalent results. Nevertheless, there appears to be some consensus that a formula like this is indeed correct, and more specifically that (with certain approximations and averaging) one has the remarkable formula

$$\frac{dE}{dt} = -\frac{G_N}{5}(\dddot{Q}^{ret})_{ik}(\dddot{Q}^{ret})^{ik} \quad . \tag{23.118}$$

While this formula (with its 3rd derivative squared) may look unfamiliar, it is precisely analogous to the corresponding formula for the radiated power of an electric quadrupole in Maxwell theory (also proportional to $\dddot{Q}^2$). The main difference between gravitational and electro-magentic radiation lies in the fact that in the Maxwell case the leading (lowest multipole) contribution arises from dipole radiation, while in the gravitational case the leading contribution is quadrupole radiation.

23.11 Even Briefer Comments on Detection of Gravitational Waves

I will conclude this section with some very general comments on the detection of gravitational waves.

In principle, this ought to be straightforward. In practice, however, because of the extreme weakness of gravitational fields, this is about as far from straightforward as one can possibly imagine. For example, on the basis of the calculations done in section 23.9, one might like to simply try to track the separation of two freely suspended masses (and this is indeed part of the principle of the interferometers I will briefly return to below).

Alternatively, the particles need not be free but could be connected by a solid piece of material. Then gravitational tidal forces will stress the material. If the resonant frequency of this "antenna" equals the frequency of the gravitational wave, this should lead to a detectable oscillation. This is the principle of the so-called *Weber detectors* or *Weber bars* (1966-. . .). While fine in principle, in practice gravitational waves are extremely weak. To the best of my knowledge, such detectors have not produced conclusive results so far, and other detection techniques are favoured in modern generations of detectors.

More sensitive modern experiments are not fine-tuned to a particular resonant frequency but can in principle detect a continuous range of frequencies. These use detectors based on huge laser Michelson interferometers (arms several kilometers long), e.g. LIGO (Laser Interferometer Gravitational Wave Observatory) and VIRGO. These or their upgrades, or the space-based LISA (Laser Interferometer Space Antenna), are widely expected to have reached sufficient sensitivity to finally directly detect gravitational waves in the next couple of years.[67] **See the 2016 Update at the end of this section!**

However, in spite of the absence of direct evidence for gravitational waves, reassuringly there is indirect (and very compelling) evidence for gravitational waves. A binary system of stars rotating around its common center of mass should radiate gravitational waves (much like electro-magnetic synchroton radiation). For two stars of equal mass M at distance $2r$ from each other, the prediction of General Relativity is that the power radiated by the binary system according to the general formula (23.118) is

$$P = dE/dt = -\frac{2}{5}\frac{G_N^4 M^5}{r^5} \quad . \tag{23.119}$$

This energy loss has actually been observed. In 1974, Hulse and Taylor discovered a binary system, affectionately known as PSR1913+16, in which both stars are small neutron stars, both roughly of solar mass, one of them being a *pulsar*, a rapidly spinning neutron star. The period of the orbit is only eight hours, and the fact that one of the stars is a pulsar provides a highly accurate clock with respect to which a change in the period as the binary loses energy can be measured. The observed value is in good agreement with the theoretical prediction for loss of energy by gravitational radiation and Hulse and Taylor were rewarded for these discoveries with the 1993 Nobel Prize. These observations have been confirmed and refined by the discovery and precise measurements and observations of other (even more extreme) binary systems.

Other situations in which gravitational waves might be either detected directly or inferred indirectly are extreme situations like gravitational collapse (supernovae) or matter orbiting black holes.

2016 Update[68]

On February 11, 2016, the LIGO Scientific Collaboration and Virgo Collaboration teams announced that they had made the first observation of gravitational waves, originating from a pair of merging black holes using the Advanced LIGO detectors. On June 15, 2016, a second detection of gravitational waves from coalescing black holes was announced.

[67] If you want (your PC) to contribute to the search for (continous) gravitational wave sources such as pulsars, take a look at the Einstein@Home project at `http://einstein.phys.uwm.edu/`.

[68] Quoted from `https://en.wikipedia.org/wiki/Gravitational_wave`. For more details, see e.g. `https://en.wikipedia.org/wiki/First_observation_of_gravitational_waves`.

D: General Relativity and the Solar System

24 Einstein Equations and Spherical Symmetry

24.1 Introduction

Einstein himself suggested three tests of General Relativity, namely

1. the gravitational redshift
2. the deflection of light by the sun
3. the anomalous precession of the perihelion of the orbits of Mercury and Venus,

and calculated the theoretical predictions for these effects. In the meantime, other tests have also been suggested and performed, for example the time delay of radar echos passing the sun (the Shapiro effect).[69]

All these tests have in common that they are carried out in empty space, with gravitational fields that are to an excellent approximation stationary (time independent) and isotropic (spherically symmetric). Thus our first aim will have to be to solve the vacuum Einstein equations under the simplifying assumptions of isotropy and time-independence. This, as we will see, is

- indeed not too difficult
- and remarkably rewarding.

24.2 Static Spherically Symmetric Metrics

Even though we have decided that we are interested in stationary spherically symmetric metrics, we still have to determine what we actually mean by this statement. After all, a metric which looks time-independent in one coordinate system may not do so in another coordinate system. There are two ways of approaching this issue:

1. One can try to look for a covariant characterisation of such metrics, in terms of Killing vectors etc. In the present context, this would amount to considering metrics which admit four Killing vectors, one of which is timelike, with the remaining three representing the Lie algebra of the rotation group $SO(3)$.
2. Or one works with 'preferred' coordinates from the outset, in which these symmetries are manifest.

[69] For detailed discussions of experimantal tests of general relativity see e.g. (1) the popular account C. Will, *Was Einstein right?*, (2) the detailed monograph C. Will, *Theory and Experiment in Gravitational Physics*, (3) the recent review article S. Turyshev, *Experimental tests of general relativity*, `arXiv:0806.1731 [gr-qc]`, and the useful resource letter C. Will, *Resource Letter PTG-1: Precision Tests of Gravity*, `arXiv:1008.0296 [gr-qc]`.

While the former approach may be conceptually more satisfactory, the latter is much easier to work with and is hence the one we will adopt.

It is important to recall and realise once again that, precisely because the theory is invariant under coordinate transformations, one is allowed to choose whatever coordinate system is most convenient to perform a calculation (much like Lorentz invariance in special relativity allows one to prove Lorentz-invariant statements by proving them in any suitably chosen inertial frame).

We will implement the condition of time-independence by choosing all the components of the metric to be time-independent (this is the condition (16.52) we called stationarity in section 16.4), and we will express the condition of isotropy by the requirement that, in terms of spatial polar coordinates (r, θ, ϕ) the metric can be written as

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + 2C(r)dr\ dt + D(r)r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{24.1}$$

This ansatz, depending on the four functions $A(r), B(r), C(r), D(r)$, can still be simplified a lot by choosing appropriate new time and radial coordinates.

First of all, let us introduce a new time coordinate $T(t, r)$ by

$$T(t, r) = t + \psi(r) \ . \tag{24.2}$$

Then

$$dT^2 = dt^2 + \psi'^2 dr^2 + 2\psi' dr\ dt \ . \tag{24.3}$$

Thus we can eliminate the off-diagonal term in the metric by choosing ψ to satisfy the differential equation

$$\frac{d\psi(r)}{dr} = -\frac{C(r)}{A(r)} \ . \tag{24.4}$$

This is tantamount to making the coordinate choice $C(r) = 0$, so that the metric can be chosen to have the diagonal form

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + D(r)r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{24.5}$$

In the terminology of section 16.4 the metric is then not only stationary but actually static and thus what we have shown is that a stationary spherically symmetric metric is automatically static (as already mentioned in the discussion around (16.87)). Thus in the context of spherical symmetry the two notions coincide and we will not be overly pedantic about the use of the word stationary versus that of the word static in the following.

We can also eliminate $D(r)$ by introducing a new radial coordinate $R(r)$ by $R^2 = D(r)r^2$. This is tantamount to making the coordinate choice $D(r) = 1$. Thus we can assume that the line element of a static isotropic metric is of the form

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ , \tag{24.6}$$

known as the *standard form* of a static spherically symmetric metric.

<u>REMARKS:</u>

1. Of course this choice is only valid if, or in regions where, $D(r) > 0$. More generally, any coordinate choice is a local choice of coordinates, and one needs to be aware of the possibility that such a choice will not provide a global picture of the space-time one wishes to describe. This will be amply illustrated by our discussion in sections 26 and 27.

2. Another useful presentation, related to the above by a coordinate transformation, is provided by the choice $D(r) = B(r)$, so that the metric takes the form

$$ds^2 = -A(r)dt^2 + B(r)(dr^2 + r^2 d\Omega^2) = -A(r)dt^2 + B(r)d\vec{x}^2 \quad , \tag{24.7}$$

with $d\vec{x}^2$ the standard Euclidean line element. This is the static spherically symmetric metric in what is known as *isotropic form*. The advantage of this isotropic form of the metric is that one can, as already indicated in (24.7), replace $dr^2 + r^2 d\Omega^2$ by e.g. the standard metric on $\mathbb{R}^3$ in Cartesian coordinates, or any other metric on $\mathbb{R}^3$. This is useful when (like many astronomers) one likes to think of the solar system as being essentially described by flat space, with some choice of coordinates.

3. As we will see in the course of section 27, other useful coordinate choices that we might have made (at least with the benefit of hindsight), are (24.1) with $D(r) = 1$ and either the condition $B(r) = 1$ (this is what will give rise to the so-called Painlevé-Gullstrand coordinates, section 27.2), or the condition $B(r) = 0$ (which will give rise to Eddington-Finkelstein coordinates, section 27.4).

   In the former case the metric takes the form

$$B(r) = g_{rr} = 1 \quad \Rightarrow \quad ds^2 = -A(r)dt^2 + 2C(r)dr\, dt + (dr^2 + r^2 d\Omega^2) \tag{24.8}$$

   and has the characteristic property that it is non-diagonal while the metric on the slices of constant t (which is not the same coordinate t as that appearing in the standard metric (24.6)) is just the flat Euclidean metric,

$$ds^2|_{t=t_0} = dr^2 + r^2 d\Omega^2 \quad . \tag{24.9}$$

   In the latter case one has

$$B(r) = g_{rr} = 0 \quad \Rightarrow \quad ds^2 = -A(r)dt^2 + 2C(r)dr\, dt + r^2 d\Omega^2 \quad . \tag{24.10}$$

   In this case one has the (at first sight peculiar) feature that the metric induced on the surfaces of constant t (which is again different from any of the other coordinates also called t above), namely

$$ds^2|_{t=t_0} = r^2 d\Omega^2 \quad , \tag{24.11}$$

is degenerate, as it just provides a non-degenerate metric for the two directions along the 2-sphere. In other words, the surfaces of constant t are *lightlike* or *null*, with r a lightlike direction. We will see in section 27.5 that this coordinate choice is particularly convenient for understanding and unravelling some of the more mysterious features of the Schwarzschild metric.

For the time being, however, we will mostly be using the metric in the standard form (24.6), as this coordinate system is well adapted to the description of the exterior of a normal star. Let us note some immediate properties of this metric:

1. By our ansatz, the components of the metric are time-independent. Because we have been able to eliminate the $dtdr$-term, the metric is also invariant under time-reversal $t \to -t$. Thus a stationary spherically symmetric metric is static (cf. the discussion in section 16.4).

2. The surfaces of constant t and r have the metric
$$ds^2|_{r=const.,t=const.} = r^2 d\Omega^2 \quad , \tag{24.12}$$
and hence have the geometry of two-spheres.

3. Because $B(r) \neq 1$, we cannot identify r with the proper radial distance. However, even though r is not a measure of proper radial distance, it has the clear geometrical significance that a 2-sphere of coordinate radius r has the area
$$A(S^2_r) = 4\pi r^2 \quad . \tag{24.13}$$
For this reason, the coordinate r is also known as the *aerea radius* or *aereal radius*.

4. Also, even though the coordinate time t is not directly measurable, up to an affine transformation
$$t \to at + b \tag{24.14}$$
it can be invariantly characterised by the fact that $\partial/\partial t$ is a timelike Killing vector.

5. The functions $A(r)$ and $B(r)$ are now to be found by solving the Einstein field equations.

6. If we want the solution to be asymptotically flat (i.e. that it approaches Minkowski space for $r \to \infty$), we need to impose the boundary conditions
$$\lim_{r\to\infty} A(r) = \lim_{r\to\infty} B(r) = 1 \quad . \tag{24.15}$$

7. We will come back to other aspects of measurements of space and time in such a geometry after we have solved the Einstein equations.

In conclusion to this section I want to stress that in the present discussion we have assumed from the outset that the metric is stationary. However, it can be shown with little effort (see section 24.6) that the vacuum Einstein equations actually imply all by themselves that a spherically symmetric metric is necessarily static!

This result is known as *Birkhoff's theorem*. It is the General Relativity analogue of the Newtonian result that a spherically symmetric body behaves as if all the mass were concentrated in its center.

In the present context it means that the gravitational field not only of a static spherically symmetric body is static and spherically symmetric (as we have assumed), but that the same is true for a radially oscillating/pulsating object. This is a bit surprising because one would expect such a body to emit gravitational radiation. What Birkhoff's theorem shows is that this radiation cannot escape into empty space (because otherwise it would destroy the time-independence of the metric). Translated into the language of waves, this means that there is no s-wave (monopole) gravitational radiation.

24.3 Solving the Einstein Equations: the Schwarzschild Metric

We will now solve the vacuum Einstein equations for the static isotropic metric in standard form, i.e. we look for solutions of $R_{\mu\nu} = 0$ for metrics of the type (24.6). You should have already (as an exercise) calculated all the Christoffel symbols of this metric, using the Euler-Lagrange equations for the geodesic equation, as described in section 3.1.

As a reminder, here is how this method works. To calculate all the Christoffel symbols $\Gamma^r_{\mu\nu}$, say, in one go, you look at the Euler Lagrange equation for $r = r(\tau)$ resulting from the Lagrangian $\mathcal{L} = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu/2$. This is easily seen to be

$$\ddot{r} + \frac{B'}{2B}\dot{r}^2 + \frac{A'}{2B}\dot{t}^2 + \cdots = 0 \tag{24.16}$$

(a prime denotes an r-derivative), from which one reads off that $\Gamma^r_{rr} = B'/2B$ etc. Proceeding in this way, you should find (or have found) that the non-zero Christoffel symbols are given by

$$\begin{aligned}
&\Gamma^r{}_{rr} = \frac{B'}{2B} \qquad \Gamma^r{}_{tt} = \frac{A'}{2B} \\
&\Gamma^r{}_{\theta\theta} = -\frac{r}{B} \qquad \Gamma^r{}_{\phi\phi} = -\frac{r\sin^2\theta}{B} \\
&\Gamma^\theta{}_{\theta r} = \Gamma^\phi{}_{\phi r} = \frac{1}{r} \qquad \Gamma^t{}_{tr} = \frac{A'}{2A} \\
&\Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta \qquad \Gamma^\phi{}_{\phi\theta} = \cot\theta
\end{aligned} \tag{24.17}$$

Now we need to calculate the Ricci tensor of this metric. A silly way of doing this would be to blindly calculate all the components of the Riemann tensor and to then perform all the relevant contractions to obtain the Ricci tensor. A more intelligent and less time-consuming strategy is the following:

1. Instead of using the explicit formula for the Riemann tensor in terms of Christoffel symbols, one should use directly its contracted version
$$\begin{aligned}R_{\mu\nu} &= R^{\lambda}_{\ \mu\lambda\nu} \\ &= \partial_\lambda\Gamma^{\lambda}_{\ \mu\nu} - \partial_\nu\Gamma^{\lambda}_{\ \mu\lambda} + \Gamma^{\lambda}_{\ \lambda\rho}\Gamma^{\rho}_{\ \mu\nu} - \Gamma^{\lambda}_{\ \nu\rho}\Gamma^{\rho}_{\ \mu\lambda}\end{aligned} \tag{24.18}$$
and use the formula (5.49) for $\Gamma^{\lambda}_{\ \mu\lambda}$ derived previously.

2. The high degree of symmetry of a static spherically symmetric metric implies that many components of the Ricci tensor are automatically zero. For example, invariance of the Schwarzschild metric under $t \to -t$ implies that $R_{rt} = 0$. The argument for this is simple:
   - Since the metric is invariant under $t \to -t$, the Ricci tensor should also be invariant.
   - Under the coordinate transformation $t \to -t$, R_{rt} transforms as $R_{rt} \to -R_{rt}$.
   - Hence, invariance requires $R_{rt} = 0$, and no further calculations for this component of the Ricci tensor are required.

3. Analogous arguments, now involving θ or ϕ instead of t, imply that
$$R_{r\theta} = R_{r\phi} = R_{t\theta} = R_{t\phi} = R_{\theta\phi} = 0 \ . \tag{24.19}$$

4. Since the Schwarzschild metric is spherically symmetric, its Ricci tensor is also spherically symmetric. It is easy to prove, by considering the effect of a coordinate transformation that is a rotation of the two-sphere defined by θ and ϕ (leaving the metric invariant), that this implies that
$$R_{\phi\phi} = \sin^2\theta R_{\theta\theta} \ . \tag{24.20}$$
Here is one possible proof (I will give a shorter argument below): Consider a coordinate transformation $(\theta,\phi) \to (\theta',\phi')$. Then
$$d\theta^2 + \sin^2\theta d\phi^2 = \left[\left(\frac{\partial\theta}{\partial\theta'}\right)^2 + \sin^2\theta\left(\frac{\partial\phi}{\partial\theta'}\right)^2\right] d\theta'^2 + \ldots \tag{24.21}$$
Thus, a necessary condition for the metric to be invariant is
$$(\frac{\partial\theta}{\partial\theta'})^2 + \sin^2\theta(\frac{\partial\phi}{\partial\theta'})^2 = 1 \ . \tag{24.22}$$
Now consider the transformation behaviour of $R_{\theta\theta}$ under such a transformation. Using $R_{\theta\phi} = 0$, one has
$$R_{\theta'\theta'} = (\frac{\partial\theta}{\partial\theta'})^2 R_{\theta\theta} + (\frac{\partial\phi}{\partial\theta'})^2 R_{\phi\phi} \ . \tag{24.23}$$

Demanding that this be equal to $R_{\theta\theta}$ (because we are considering a coordinate transformation which does not change the metric) and using the condition derived above, one obtains

$$R_{\theta\theta} = R_{\theta\theta}(1 - \sin^2\theta(\frac{\partial\phi}{\partial\theta'})^2) + (\frac{\partial\phi}{\partial\theta'})^2 R_{\phi\phi} \ , \tag{24.24}$$

which implies (24.20).

5. Alternatively, look at the mixed spherical components $R^a_{\ b}$ of the Ricci tensor, $x^a = (\theta, \phi)$. Rotational invariance implies that $R^a_{\ b} \sim \delta^a_{\ b}$. Since $g_{\theta\phi} = 0$ and $g_{\phi\phi} = \sin^2\theta g_{\theta\theta}$, this implies for the covariant components R_{ab} that $R_{\theta\phi} = 0$ and $R_{\phi\phi} = \sin^2\theta R_{\theta\theta}$.
6. Thus the only components of the Ricci tensor that we need to compute are R_{rr}, R_{tt} and $R_{\theta\theta}$.
7. R_{tt} was already determined in section 13.5 (see (13.37)) using a shortcut procedure based on the Killing vector $\xi = \partial_t$ and some identities relating Killing vectors and the curvature tensor, and $R_{\theta\theta}$ and $R_{\phi\phi}$ can be calculated by the same procedure (or directly).
8. Therefore the only component that remains to be determined is R_{rr}.

Putting everything together, the final result for the Ricci tensor of the general static spherically symmetric metric is

$$\begin{aligned} R_{tt} &= \frac{A''}{2B} - \frac{A'}{4B}(\frac{A'}{A} + \frac{B'}{B}) + \frac{A'}{rB} \\ R_{rr} &= -\frac{A''}{2A} + \frac{A'}{4A}(\frac{A'}{A} + \frac{B'}{B}) + \frac{B'}{rB} \\ R_{\theta\theta} &= 1 - \frac{1}{B} - \frac{r}{2B}(\frac{A'}{A} - \frac{B'}{B}) \ . \end{aligned} \tag{24.25}$$

Inspection of these formulae reveals that there is a linear combination which is particularly simple, namely $BR_{tt} + AR_{rr}$, which can be written as

$$BR_{tt} + AR_{rr} = \tfrac{1}{rB}(A'B + B'A) \ . \tag{24.26}$$

Demanding that this be equal to zero, one obtains

$$A'B + B'A = 0 \ \ \Rightarrow \ \ A(r)B(r) = const. \tag{24.27}$$

Asymptotic flatness fixes this constant to be $= 1$, so that

$$B(r) = \frac{1}{A(r)} \ . \tag{24.28}$$

Plugging this result into the expression for $R_{\theta\theta}$, one obtains

$$R_{\theta\theta} = 0 \ \ \Rightarrow \ \ A - 1 + rA' = 0 \ \ \Leftrightarrow \ \ (Ar)' = 1 \ , \tag{24.29}$$

which has the solution $Ar = r + C$ or

$$A(r) = 1 + \frac{C}{r} \quad . \tag{24.30}$$

Now also $R_{tt} = R_{rr} = 0$.

To fix C, we compare with the Newtonian limit which tells us that asymptotically $A(r) = -g_{00}$ should approach (temporarily reinserting c) $(1 + 2\Phi/c^2)$, where $\Phi = -G_N M/r$ is the Newtonian potential for a static spherically symmetric star of mass M. Thus $C = -2MG/c^2$, and the final form of the metric is

$$\boxed{ds^2 = -(1 - \frac{2MG_N}{c^2 r})c^2 dt^2 + (1 - \frac{2MG_N}{c^2 r})^{-1} dr^2 + r^2 d\Omega^2} \tag{24.31}$$

This is the famous *Schwarzschild metric*, obtained by the astronomer Karl Schwarzschild in 1916, the very same year that Einstein published his field equations, while in hospital as a soldier in World War I.

REMARKS:

1. Due to some of the idiosyncracies of Einstein's earlier versions of his field equations, "Einstein's obsession with coordinate systems in which the determinant of the metric was precisely -1"[70], Schwarzschild originally found this solution in different (and somewhat less convenient) coordinates. The starting point for him was the "unimodular" spherical coordinate system $\{z^k\} = (\rho, \psi, \phi)$ introduced at the end of section 4.4, in terms of which the Euclidean metric takes the form (4.72)

$$dr^2 + r^2 d\Omega^2 = r(\rho)^{-4} d\rho^2 + r(\rho)^2 (d\psi^2/\sin^2\theta(\psi) + \sin^2\theta(\psi) d\phi^2) \tag{24.32}$$

   One can then, as an alternative to the standard ansatz (24.6) for a spherically symmetric static space-time metric, make an ansatz for the space-time metric of the form

$$ds^2 = -a(\rho)dt^2 + b(\rho)d\rho^2 + d(\rho)(d\psi^2/\sin^2\theta(\psi) + \sin^2\theta(\psi)d\phi^2) \quad , \tag{24.33}$$

   subject to the unimodularity condition

$$a(\rho)b(\rho)d(\rho)^2 = 1 \tag{24.34}$$

   and the (compatible) asymptotic flatness conditions (as $\rho \to \infty$)

$$a(\rho) \to 1 \quad , \quad b(\rho) \to r(\rho)^{-4} \quad , \quad d(\rho) \to r(\rho)^2 \quad . \tag{24.35}$$

[70] P. Fromholz, E. Poisson, C.M. Will, *The Schwarzschild metric: It's the coordinates, stupid!*, `arXiv:1308.0394 [gr-qc]`.

If one plugs this into the vacuum Einstein equations (which simplify somewhat for unimodular metrics), one finds the Schwarzschild metric in Schwarzschild's original coordinates. See the reference in footnote 70 for further details.

2. The same solution was apparently (I have not checked this myself, hence the "apparently") discovered independently a few months later by Johannes Droste, a student of Lorentz (cf. the reference in footnote 13 in section 5.9).

3. We will usually not write the constant G_N explicitly (and set $c = 1$), and thus we introduce the abbreviation

$$m = \frac{G_N M}{c^2} \quad , \tag{24.36}$$

in terms of which the Schwarzschild metric takes the form

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} \quad . \tag{24.37}$$

4. The interpretation of m is that of a classical (i.e. no $\hbar$) gravitational length scale or *gravitational mass radius* associated to the mass M. To see that all this is dimensionally correct, note that Newton's constant has dimensions (M mass, L length, T time) $[G_N] = \mathsf{M}^{-1}\mathsf{L}^3\mathsf{T}^{-2}$ so that

$$[G_N] = \mathsf{M}^{-1}\mathsf{L}^3\mathsf{T}^{-2} \quad \Rightarrow \quad [m] = [G_N M/c^2] = \mathsf{L} \quad . \tag{24.38}$$

For examples of the value of m for various objects see section 24.4.

We have seen that, by imposing appropriate symmetry conditions on the metric, and making judicious use of them in the course of the calculation, the complicated Einstein equations become rather simple and manageable, and we will now embark on a detailed investigation of the solution that we have found.

24.4 Schwarzschild Coordinates and Schwarzschild Radius

The metric we have obtained is quite remarkable in several respects. As mentioned before, the vacuum Einstein equations imply that an isotropic metric is static. Furthermore, the metric contains only a single constant of integration, the mass M. This implies that the metric in the exterior of a spherical body is completely independent of the composition of that body. Whatever the energy-momentum tensor for a star may be, the field in the exterior of the star has always got the form (24.31). This considerably simplifies the physical interpretation of General Relativity. In particular, in the subsequent discussion of tests of General Relativity, which only involve the exterior of stars like the sun, we do not have to worry about solutions for the interior of the star and how those could be patched to the exterior solutions.

We begin our investigation of the Schwarzschild metric by taking a look at the coordinates and their range (always a useful first step).

1. The polar coordinates θ and ϕ have their standard interpretation and range.

2. The time coordinate t can be interpreted as the proper time of a static observer infinitely far away from the star, at $r \to \infty$. Thus, given the asymptotic flatness of the solution, we can think of t as measuring Minkowski time. Up to a constant conversion factor, the coordinate time t agrees with the proper time of a static observer, i.e. an observer at fixed values of the spatial coordinates (r, θ, ϕ), and we can therefore say that these Schwarzschild coordinates are adapted to such static observers (but we will see later, in section 26, that t is not always suitable to describe freely falling observers).

   Clearly, the range of t is unrestricted, $-\infty < t < \infty$ (but again we will see later that this may not always be good enough ...).

3. The issue regarding the possible range of r is a priori more interesting.

   (a) We had already discussed above, that the geometrical interpretation of r is that of an area radius, i.e. it is characterised by the fact that, even though r is not proper radial distance, the surface area of a sphere of constant radius r is $4\pi r^2$.

   (b) Moreover, the metric is, by construction, a vacuum metric. Thus, if the star has radius r_0, then the solution is only valid for $r > r_0$, and the range of r is restricted appropriately, $r_0 < r < \infty$.

   (c) However, (24.37) also shows that the metric appears to have a singularity at the *Schwarzschild radius* r_s, given by

   $$r_s = \frac{2G_N M}{c^2} = 2m \quad . \tag{24.39}$$

   Thus, for the time being we will also require $r > r_s$. For most practical purposes, this is not a further constraint on the range of r, since the radius of a physical object is almost always much larger than its Schwarzschild radius. For example, for a proton, for the earth and for the sun one has approximately

   $$\begin{aligned} M_{\text{proton}} &\sim 10^{-24}\ \mathsf{g} \quad \Rightarrow r_s \sim 2,5 \times 10^{-52}\ \mathsf{cm} \ll r_0 \sim 10^{-13}\ \mathsf{cm} \\ M_{\text{earth}} &\sim 6 \times 10^{27}\ \mathsf{g} \quad \Rightarrow r_s \sim 1\ \mathsf{cm} \ll r_0 \sim 6000\ \mathsf{km} \\ M_{\text{sun}} &\sim 2 \times 10^{33}\ \mathsf{g} \quad \Rightarrow r_s \sim 3\ \mathsf{km} \ll r_0 \sim 7 \times 10^5\ \mathsf{km} \quad . \end{aligned} \tag{24.40}$$

   However, for more compact objects, their radius can approach that of their Schwarzschild radius. For example, for neutron stars one can have $r_s \sim 0.1 r_0$, and it is an interesting question (we will take up again later on, in sections 26 and 27) what happens to an object whose size is equal to or smaller than its Schwarzschild radius.

(d) One thing that does *not* occur at r_s, however, in spite of what (24.37) may suggest, is a true physical singularity. The singularity in (24.37) turns out to be a pure coordinate singularity, i.e. an artefact of having chosen a poor coordinate system, and later on we will construct coordinates in which the metric is completely regular at r_s. Nevertheless, it turns out that something interesting *does* happen at $r = r_s$, even though there is no singularity and e.g. geodesics are perfectly well behaved there: r_s is an event-horizon, in a sense a point of no return. Once one has passed the Schwarzschild radius of an object with $r_0 < r_s$, there is no turning back, not on geodesics, but also not with any amount of acceleration.

24.5 Measuring Length and Time in the Schwarzschild Metric

In order to learn how to visualise the Schwarzschild metric (for $r > r_0 > r_s$), we will now discuss some further elementary properties of length and time in the Schwarzschild geometry.

Let us first consider proper time for a static observer, i.e. an observer at rest at fixed values of (r, θ, ϕ). Proper time is related to coordinate time by

$$d\tau = (1 - 2m/r)^{1/2} dt < dt \quad . \tag{24.41}$$

Thus, first of all, as already mentioned above, up to a constant conversion factor for such observers their proper time agrees with their coordinate time, and we can simply and conveniently label events as described by a static observer by the coordinate time t instead of the proper time of the observer.

Secondly, we can interpret the above relationship as the statement that static clocks (measuring the proper time τ) run slower in a gravitational field - something we already saw in the discussion of the gravitational redshift in section 3.5, and also in the discussion of the so-called 'twin-paradox' and the equivalence principle in section 1.1. This formula again suggests that something interesting is happening at the Schwarzschild radius $r = 2m$ - we will come back to this below.

As regards spatial length measurements, thus $dt = 0$, we have already seen above that the slices $r = const.$ have the standard two-sphere geometry. However, as r varies, these two-spheres vary in a way different to the way concentric two-spheres vary in $\mathbb{R}^3$. To see this, note that the proper radius R, obtained from the spatial line element by setting $\theta = const., \phi = const.$, is

$$dR = (1 - 2m/r)^{-1/2} dr > dr \quad . \tag{24.42}$$

In other words, the proper radial distance between concentric spheres of area $4\pi r^2$ and area $4\pi(r + dr)^2$ is $dR > dr$ and hence larger than in flat space. Note that $dR \to dr$

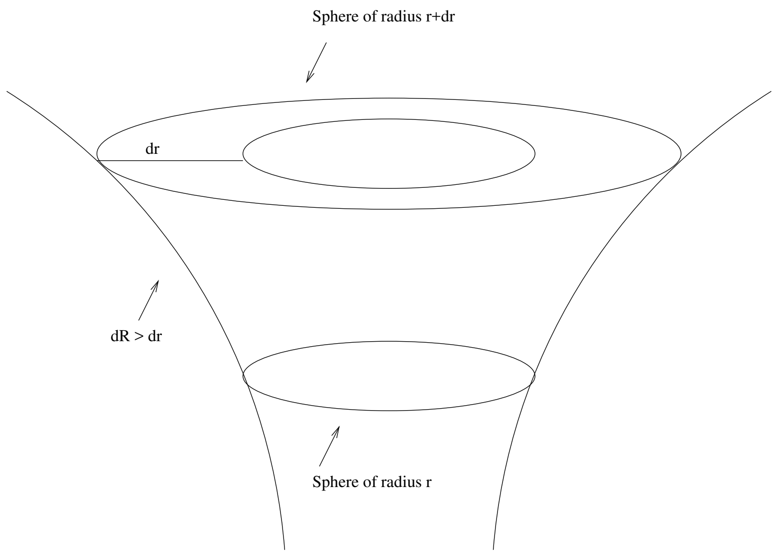


Figure 11: Figure illustrating the geometry of the Schwarzschild metric. In $\mathbb{R}^3$, concentric spheres of radii r and $r+dr$ are a distance dr apart. In the Schwarzschild geometry, such spheres are a distance $dR > dr$ apart. This departure from Euclidean geometry becomes more and more pronounced for smaller values of r, i.e. as one travels down the throat towards the Schwarzschild radius $r = 2m$.

for $r \to \infty$ so that, as expected, far away from the origin the space approximately looks like $\mathbb{R}^3$. One way to visualise this geometry is as a sort of throat or sink, as in Figure 11.

To get some more quantitative feeling for the distortion of the geometry produced by the gravitational field of a star, consider a long stick lying radially in this gravitational field, with its endpoints at the coordinate values $r_1 > r_2$. To compute its length L, we have to evaluate

$$L = \int_{r_2}^{r_1} dr(1 - 2m/r)^{-1/2} \quad . \tag{24.43}$$

It is possible to evaluate this integral in closed form (by changing variables from r to $u = 1/r$), but for the present purposes it will be enough to treat $2m/r$ as a small perturbation and to only retain the term linear in m in the Taylor expansion. Then we find

$$L \approx \int_{r_2}^{r_1} dr(1 + m/r) = (r_1 - r_2) + m \log \frac{r_1}{r_2} > (r_1 - r_2) \quad . \tag{24.44}$$

We see that the corrections to the Euclidean result are suppressed by powers of the Schwarzschild radius $r_s = 2m$ so that for most astronomical purposes one can simply work with coordinate distances.

This is even more evident when one puts the Schwarzschild metric into the isotropic form of the metric (24.7). This is accomplished by the coordinate transformation

$$r = \rho(1 + \frac{m}{2\rho})^2 \quad , \tag{24.45}$$

leading to

$$
\begin{aligned}
ds^2 &= -\frac{(1-\frac{m}{2\rho})^2}{(1+\frac{m}{2\rho})^2}dt^2 + (1+\frac{m}{2\rho})^4(d\rho^2+\rho^2 d\Omega^2) \\
&= -\frac{(1-\frac{m}{2\rho})^2}{(1+\frac{m}{2\rho})^2}dt^2 + (1+\frac{m}{2\rho})^4\, d\vec{x}^2 \qquad (\ \rho^2 = \vec{x}^2\ )\ ,
\end{aligned}
\tag{24.46}
$$

as can easily be verified. To actually find the appropriate coordinate transformation in the first place, one needs to solve the equation

$$
f(r)^{-1}dr^2 + r^2 d\Omega^2 = B(\rho)(d\rho^2 + \rho^2 d\Omega^2)\ , \tag{24.47}
$$

i.e.

$$
\frac{dr}{rf^{1/2}(r)} = \frac{d\rho}{\rho} \quad \Rightarrow \quad \ln(\rho/\rho_0) = \cosh^{-1}(r/m-1)\ . \tag{24.48}
$$

This leads to (24.45) for the choice $\rho_0 = m/2$ (but any other choice would have been just as good). From this one can now read off that the relation between proper and coordinate distance (the latter now referring to the spatial Cartesian coordinates $\vec{x}$ in (24.46)) is

$$
(\Delta x)_{\text{proper}} = (1+\frac{m}{2\rho})^2\,(\Delta x)\ . \tag{24.49}
$$

However, in interpreting this or using this form of the metric for other purposes, one should pay attention to the fact that the region $2m < r < \infty$ of the Schwarzschild metric is covered twice by the isotropic coordinates. In particular, $r \to \infty$ both for $\rho \to 0$ and for $\rho \to \infty$, while $r(\rho)$ reaches its minimal value $r = 2m$ for $\rho = m/2$.

Thus the metric in isotropic coordinates appears to describe not just one but two identical (isometric) asymptotically flat regions, joined together at the 2-sphere $\rho = m/2 \leftrightarrow r = 2m$. This is the first indication that with the Schwarzschild metric we seem to have obtained more than we bargained for, in particular when considering objects whose radius is smaller than the Schwarzschild radius $r_s = 2m$. Later on, we will encounter numerous other coordinate systems for the Schwarzschild metric, providing us with different insights into its physics and geometry. In particular, we will then (re-)discover this second asymptotically flat region in section 27.8 (as the "mirror region III").

24.6 Einstein Equations for Spherical Symmetry and Birkhoff's Theorem

A slightly more general calculation than we have performed in section 24.3 to find the Schwarzschild solution for the exterior of a spherically symmetric static star provides us with

1. a proof of Birkhoff's theorem, mentioned above, that a spherically symmetric vacuum solution of the Einstein equations is necessarily static;

2. some more insight into the interpretation of the parameter M as the mass of the solution;

3. as an added benefit, the basic set of equations governing the solutions of the Einstein equations for the interior of the star (an issue we will briefly consider in section 24.7).

Thus, let us start with a general spherically symmetric (but not necessarily time-independent) metric. Generalising (24.1), we can at first parametrise such a metric as

$$ds^2 = -A(t,r)dt^2 + B(t,r)dr^2 + 2C(t,r)dr\,dt + R(t,r)^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{24.50}$$

This form of the metric is still invariant under transformations

$$(t,r) \to (t'(t,r), r'(t,r)) \ . \tag{24.51}$$

Modulo one *caveat* to be discussed below, arguments analogous to those leading to (24.6) allow one to conclude that by a suitable choice of coordinates the metric can be chosen to be of the form

$$ds^2 = -A(t,r)dt^2 + B(t,r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{24.52}$$

This *caveat* is the following.[71] Clearly, the metric (24.52) has the property that the gradient of the radius of the 2-sphere is spacelike,

$$g^{\alpha\beta}\partial_\alpha r \partial_\beta r = g^{rr} > 0 \ . \tag{24.53}$$

Since $g_{\theta\theta}$ transforms as a scalar under transformations (24.51) of the (t,r)-coordinates among themselves, this statement is invariant under such coordinate transformations. Thus a necessary (and locally sufficient) condition in order to be able to achieve the form (24.52) of the metric is that this gradient was spacelike in the original metric,

$$g^{\alpha\beta}\partial_\alpha R(t,r)\partial_\beta R(t,r) > 0 \ . \tag{24.54}$$

Simple "counterexamples" to the form (24.52) of the metric are thus provided

- e.g. by space-times with a constant $R(t,r)$,

$$R(t,r) = R \ , \tag{24.55}$$

  so that the metric describes a Cartesian product of a 2-dimensional space-time (with coordinates (t,r)) with a 2-sphere of fixed radius R (in which case, no coordinate transformation of the (t,r) will turn this constant into r^2)

[71] Cf. e.g. the discussion in section 14.1 of J. Plebanski, A. Krasinski, *An Introduction to General Relativity and Cosmology.*

- or metrics with $R(t,r)$ with a timelike gradient, such as

$$R(t,r) = t \quad , \tag{24.56}$$

  for which one could choose R as a new time coordinate, but not as a radial coordinate.

These cases (as well as that where the gradient is null) would in principle require a separate analysis, but I will forego this here. The case where the gradient is timelike roughly speaking corresponds to an exchange of the roles of t and r, and since this will be of some interest in our discussion of black holes later on, I will briefly come back to this below.

For now we continue with the understanding that we are considering (regions of) space-times for which the form (24.52) of the metric is valid. However, as we have already seen in the derivation of the Schwarzschild metric, this parametrisation of the metric in terms of the two functions $A(r)$ and $B(r)$ is not ideal. To see what might be more convenient, we first reanalyse the Einstein equations in the time-independent case, but this time with an energy-momentum tensor. Thanks to the relation (24.26) we have

$$R^r_{\ r} - R^t_{\ t} = (rB)^{-1}\frac{(AB)'}{AB} \quad . \tag{24.57}$$

This suggests that it is useful to introduce a new function $h(r)$ through

$$A(r)B(r) = \mathrm{e}^{2h(r)} \quad , \tag{24.58}$$

i.e. through

$$A(r) = \mathrm{e}^{2h(r)}f(r) \quad , \quad B(r) = f(r)^{-1} \tag{24.59}$$

for some arbitrary new function $f(r)$. Using the Einstein equations (19.35) in the form

$$R^\alpha_{\ \beta} = 8\pi G_N (T^\alpha_{\ \beta} - \tfrac{1}{2}\delta^\alpha_{\ \beta}T^\gamma_{\ \gamma}) \quad , \tag{24.60}$$

one sees that one particular linear combination of the Einstein equations now takes the form

$$h'(r) = 4\pi G_N\, r f(r)^{-1}(T^r_{\ r} - T^t_{\ t}) \quad . \tag{24.61}$$

The remaining independent component (in the time-independent case) can be chosen to be

$$R^t_{\ t} - \tfrac{1}{2}R = 8\pi G_N\, T^t_{\ t} \tag{24.62}$$

which, after a bit of algebra with the formulae (24.25), works out to be

$$[r(f(r)-1)]' = 8\pi G_N\, r^2 T^t_{\ t} \quad . \tag{24.63}$$

This suggests that one trades $f(r)$ for another function

$$r(f(r)-1) = -2m(r) \quad \Leftrightarrow \quad f(r) = 1 - 2m(r)/r \quad , \tag{24.64}$$

so that (24.63) becomes

$$m'(r) = 4\pi G_N\, r^2(-T^t_{\ t}) \ . \tag{24.65}$$

Equations (24.61) and (24.65) make it as manifest as possible that the spherically symmetric vacuum solution is the Schwarzschild metric with $f(r) = 1 - 2m/r$, m constant, and $h(r) = 0$ (an arbitrary constant can be absorbed into the definition of the time-coordinate t).

Let us therefore now, in the time-dependent case, and with the benefit of the above hindsight, parametrise the two arbitrary functions $A(t,r)$ and $B(t,r)$ in (24.52) in terms of two other functions $h(t,r)$ and either $f(t,r)$ or $m(t,r)$ by the substitutions

$$A(t,r) = \mathrm{e}^{2h(t,r)} f(t,r) \qquad B(t,r) = f(t,r)^{-1} \tag{24.66}$$

and

$$f(t,r) = 1 - \frac{2m(t,r)}{r} \ . \tag{24.67}$$

Thus, explicitly, the modified ansatz for a general spherically symmetric metric is

$$ds^2 = -\mathrm{e}^{2h(t,r)} f(t,r) dt^2 + f(t,r)^{-1} dr^2 + r^2 d\Omega^2 \ . \tag{24.68}$$

In this gauge, the full (non-vacuum) Einstein equations turn out to take a particularly simple and useful form. The previously obtained equations (24.61) and (24.65) continue to be valid also in the time-dependent case, and there is now one more independent equation, arising from, say, the (rt)-component of the Einstein equation.

Explicitly, the relevant components of the Einstein tensor are

$$\begin{aligned} G^t_{\ t} &= -\frac{2m'(t,r)}{r^2} \\ G^r_{\ t} &= +\frac{2\dot{m}(t,r)}{r^2} \\ G^r_{\ r} &= +\frac{2h'(t,r)f(t,r)}{r} - \frac{2m'(t,r)}{r^2} \ , \end{aligned} \tag{24.69}$$

with a somewhat more complicated and unenlightning expression for the angular components, depending on all of $m', m'', \dot{m}, \ddot{m}, h', h'', \dot{h}$, which we will fortunately not need. In particular, among the 3 above components only $G^r_{\ t}$ contains a time-derivative $\dot{m}(t,r) = \partial_t m(t,r)$. Moreover one can replace $G^r_{\ r}$ by the simpler linear combination

$$G^r_{\ r} \to G^r_{\ r} - G^t_{\ t} = \frac{2h'(t,r)f(t,r)}{r} \ . \tag{24.70}$$

Thus the set of 3 radial/time Einstein equations can be compactly written as

$$\begin{aligned} m'(t,r) &= 4\pi G_N\, r^2(-T^t_{\ t}) \\ \dot{m}(t,r) &= 4\pi G_N\, r^2(+T^r_{\ t}) \\ h'(t,r) &= 4\pi G_N\, r f(t,r)^{-1}(-T^t_{\ t} + T^r_{\ r}) \ . \end{aligned} \tag{24.71}$$

These equations now immediately lead to *Birkhoff's Theorem*:

- For vacuum solutions the equations (24.71) imply that
  - the mass function $m(t,r) = m$ is a constant,

$$T_{\alpha\beta} = 0 \quad \Rightarrow \quad m'(t,r) = \dot{m}(t,r) = 0 \quad \Rightarrow \quad m(t,r) = m \ \text{constant} \ , \tag{24.72}$$

  - and that $h'(t,r) = 0$ so that $h = h(t)$ is only a function of t,

$$T_{\alpha\beta} = 0 \quad \Rightarrow \quad h'(t,r) = 0 \quad \Rightarrow \quad h(t,r) = h(t) \ . \tag{24.73}$$

- Thus $h(t)$, which only appears in the (tt)-component of the metric, can simply be absorbed into a redefinition of t,

$$\begin{aligned} ds^2 &= -\mathrm{e}^{2h(t)} f(r) dt^2 &&+ f(r)^{-1} dr^2 + r^2 d\Omega^2 \\ &= -f(r) \left(\mathrm{e}^{h(t)} dt\right)^2 &&+ f(r)^{-1} dr^2 + r^2 d\Omega^2 \\ &\equiv -f(r)(dt_{new})^2 &&+ f(r)^{-1} dr^2 + r^2 d\Omega^2 \end{aligned} \tag{24.74}$$

  and we can, without loss of generality, assume that $h = 0$.

Thus we uniquely recover the Schwarzschild solution, even without having to assume from the outset that the metric is time-independent. This is Birkhoff's theorem.

Remarks:

1. A *caveat* related to that at the beginning of this section should be added here: if one applies the above reasoning to a region of space-time where $f(r) < 0$ (we will study this region of the Schwarzschild metric in great detail in section 27), so that the roles of t and r are interchanged, then the above argument still shows that an additional Killing vector emerges from the joint requirement of spherical symmetry and the vacuum Einstein equations, but now this Killing vector (misleadingly called ∂_t) is spacelike and not timelike.[72]

2. The above set (24.71) of Einstein equations for spherical symmetry (which should still be supplemented by, say, the conservation law for the energy-momentum tensor), also allows one to read off some fairly simple generalisations of Birkhoff's theorem, such as "spherical symmetry and static sources (i.e. $T^\alpha_{\ \beta} = T^\alpha_{\ \beta}(r)$) $\Rightarrow$ metric is time-independent".

3. Other generalisations of the Birkhoff theorem are reviewed and discussed in an article by H.-J. Schmidt.[73] There it is also stressed that the validity of Birkhoff-like theorems relies crucially on the fact, mentioned in section 11.3 in connection

[72] For a more careful statement and proof of Birkhoff's theorem along these lines see e.g. K. Schleich, D. Witt, *A simple proof of Birkhoff's theorem for cosmological constant*, `arXiv:0908.4110v2 [gr-qc]`.

[73] H.-J. Schmidt, *The tetralogy of Birkhoff theorems*, `arXiv:1208.5237 [gr-qc]`.

with equation (11.31) and (11.32), that the 2-dimensional Ricci tensor (in the (t,r)-directions transverse to the sphere) has only one degenerate eigenvalue, thus strongly constraining higher-dimensional generalisations of such statements.

4. Realistic astrophysical systems are neither exactly spherically symmetric nor exactly vacuum (even outside the star), and typically the sources are not static either. It is therefore of interest to investigate more generally if or to which extent Birkhoff's theorem remains approximately true when the system under consideration is only approximately spherically symmetric or vacuum. This question has been analysed by Goswami and Ellis.[74]

5. Finally, note that the characteristic form

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \tag{24.75}$$

of the Schwarzschild solution is implied not just by the vacuum Einstein equations but, more generally, by the Einstein equations with $T^t_{\ t} = T^r_{\ r}$. This situation is not as uncommon as one may think. For example, solutions of the Einstein-Maxwell equations for spherically symmetric electrically charged stars (the Reissner-Nordstrøm solution, see section 31), even with the inclusion of a cosmological constant, turn out to also be of this form. Some other examples of solutions of this type (in more than 4 dimensions) are presented in section 30.3.[75]

We conclude this section with some remarks about the mass function $m(t,r)$ appearing in the ansatz (24.67) and its interpretation:

1. First of all, Since $T^t_{\ t}$ is (minus) the energy density and $T^r_{\ t}$ represents the radial energy flux, the above equations show that $m(t,r)$ can inded be interpreted as the mass or energy of the solution.

   Indeed, let $\rho(r) = -T^t_{\ t}$ denote the energy density inside a static spherically symmetric star, say, and let $m(r) = G_N M(r)$. Then (24.65) implies

$$M'(r) = 4\pi r^2 \rho(r) \quad \Rightarrow \quad M(r) = 4\pi \int_0^r dr'\ (r')^2 \rho(r')\ , \tag{24.76}$$

   which looks exactly like the ordinary mass inside sphere of radius r (in flat space). In particular, if $\rho(r) = 0$ for $r > r_0$ (with r_0 the radius of the star), then one can interpret

$$M \equiv M(r_0) = 4\pi \int_0^{r_0} dr'\ (r')^2 \rho(r') \tag{24.77}$$

[74] R. Goswami, G. Ellis, *Almost Birkhoff Theorem in General Relativity*, `arXiv:1101.4520 [gr-qc]`, R. Goswami, G. Ellis, *Birkhoff Theorem and Matter*, `arXiv:1202.0240 [gr-qc]`.

[75] For some further reflections on the ubiquity of such solutions, see T. Jacobson, *When is $g_{tt}g_{rr} = -1$?*, `arXiv:0707.3222v3 [gr-qc]`.

as the total mass-energy of the star. One can (try to) attribute the difference between this integral and that of $\rho(r)$ weighted by the proper spatial volume element,

$$\begin{aligned} M_{\text{proper}} &= 4\pi \int_0^{r_0} dr'\ (r')^2 \sqrt{g_{rr}(r')}\rho(r') \\ &= 4\pi \int_0^{r_0} dr'\ (r')^2 \rho(r')(1 - 2m(r'))^{-1/2} > M \end{aligned} \tag{24.78}$$

to the binding energy of the star.

2. It is also worth noting that the *Misner-Sharp mass function* $M_{MS}(t,r) = m(t,r)$ in (24.67) has a coordinate invariant meaning (in spherical symmetry). First of all, for the metric (24.68) one has

$$g^{rr} = f = 1 - \frac{2m}{r} \quad . \tag{24.79}$$

Now consider, as in the argument leading to (12.138), an arbitrary coordinate transformation $(t,r) \to z^a(t,r)$, thus preserving the manifest spherical symmetry, but e.g. abandoning the areal radius r as one of the coordinates. In particular, $r = r(z^a)$ is now a function of the new coordinates. Then the metric will take the general spherically symmetric form

$$ds^2 = g_{ab}(z)dz^a dz^b + r(z)^2 d\Omega^2 \quad , \tag{24.80}$$

and by the usual tensorial transformation rule for the metric one has

$$g^{rr} = g^{ab}\partial_a r(z)\partial_b r(z) \quad . \tag{24.81}$$

Thus in a general spherically symmetric coordinate system the mass function can be expressed in terms of (or defined via) the gradient-squared of the radius function of the transverse sphere,

$$M_{MS}(z) \equiv m(z) = \frac{r(z)}{2}\Big(1 - g^{ab}(z)\partial_a r(z)\partial_b r(z)\Big) \quad , \tag{24.82}$$

and is thus a scalar under these coordinate transformations.

24.7 Interior Solution for a Static Star and the TOV Equation

The Schwarzschild solution is a solution of the vacuum Einstein equations for the exterior of a spherically symmetric (static) star. The Einstein equations also govern and describe the gravitational field = space-time geometry in the interior of the star. In this case one needs to specify the energy-momentum tensor for the matter content in the interior of the star, and in general this is a complicated astrophysics problem which we will not address here. However, a useful idealised model of the energy-momentum tensor, compatible with the symmetry requirements arising from the fact that the star is assumed to be static and spherically symmetric, which we will take to mean that the metric has the form

$$\begin{aligned} ds^2 &= -\mathrm{e}^{2h(r)} f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \\ f(r) &= 1 - \frac{2m(r)}{r} \quad , \end{aligned} \tag{24.83}$$

is provided by the ansatz

$$T^\alpha_\beta = \mathrm{diag}\left(-\rho(r), p(r), p(r), p(r)\right) \quad . \tag{24.84}$$

Here we interpret $\rho = \rho(r)$ as the energy density of the star, and $p = p(r)$ as its pressure density.

<u>REMARKS:</u>

1. This ansatz amounts to neglecting anisotropic stresses in the interior of the star (the spatial off-diagonal components) as well as energy-flow in the form of heat-conduction, say (the off-diagonal time-space components). Depending on the type of star one wishes to describe this may or may not be a justified approximation (but is considered to be an excellent approximation for very compact stars like white dwarves / dwarfs and neutron stars).

2. The covariant components $T_{\alpha\beta}$ of the energy-momentum tensor can be written as

$$T_{\alpha\beta} = (\rho + p)u_\alpha u_\beta + p g_{\alpha\beta} \quad , \tag{24.85}$$

   where

$$u^\alpha = \left((-g_{tt}(r))^{-1/2}, 0, 0, 0\right) \quad \Rightarrow \quad u^\alpha u_\alpha = g_{tt}(r)(-g_{tt}(r))^{-1} = -1 \tag{24.86}$$

   is the velocity-field of static observers, i.e. observers remaining at fixed values of the spatial coordinates (r, θ, ϕ). This is the general form of the energy-momentum tensor of what is known as a *perfect fluid* (cf. sections 7.2 and 7.5), and will be discussed in more detail in the sections on cosmology - see in particular section 35.2.

3. Specifying the energy-momentum content requires specifying not only the energy-momentum tensor (24.84) but also an equation of state, which in this simplified context amounts to postulating a relation $p = p(\rho)$. Again see section 35.2 for a discussion and examples of this. We will sidestep this issue in the discussion below since the only case we will consider explicitly is that of constant energy density $\rho(r) = \rho_0$, in which case the Einstein equations determine $p = p(r)$ via the Tolman-Oppenheimer-Volkoff equation to be derived (in general) below.

With this set-up, the Einstein equations (24.61) and (24.65) for a metric of the form (24.83) and an energy-momentum tensor of the form (24.84) read

$$\begin{aligned} m'(r) &= 4\pi G_N\, r^2 \rho(r) \\ h'(r) &= 4\pi G_N\, r f(r)^{-1}(\rho(r) + p(r)) \ . \end{aligned} \tag{24.87}$$

The first of these can be written in integral form as in (24.76),

$$m(r) = G_N M(r) \quad , \quad M(r) = 4\pi \int_0^r dr'(r')^2 \rho(r') \ , \tag{24.88}$$

where the regularity condition $M(0) = 0$ has been imposed.

The equations (24.87) need to be supplemented either by the conservation-law

$$\nabla_\alpha T^\alpha_\beta = 0 \tag{24.89}$$

for the energy-momentum tensor, or by another component of the Einstein equations, say the ($\theta\theta$)-component, but the former is simpler and more insightful. The only component of (24.89) that is not identically satisfied for an energy-momentum tensor of the form (24.84) and a metric of the form (24.83) is the $\beta = r$ component which reads

$$\begin{aligned} 0 &= \partial_\alpha T^\alpha_r + \Gamma^\alpha_{\ \alpha\beta} T^\beta_r - \Gamma^\beta_{\ \alpha r} T^\alpha_\beta = \partial_r p(r) + \Gamma^\alpha_{\ \alpha r} p(r) - \Gamma^t_{\ tr}(-\rho(r)) - \Gamma^k_{\ kr} p(r) \\ &= p'(r) + \Gamma^t_{\ tr}(\rho(r) + p(r)) = p'(r) + (h'(r) + f'(r)/2f(r))(\rho(r) + p(r)) \ , \end{aligned} \tag{24.90}$$

or

$$p'(r) = -(h'(r) + f'(r)/2f(r))(\rho(r) + p(r)) \ . \tag{24.91}$$

Here both $h'(r)$ and $f'(r)$ can be eliminated in favour of $\rho(r)$ and $p(r)$ by using the Einstein equations (24.87), in particular

$$\tfrac{1}{2} f'(r) = -m'(r)/r + m(r)/r^2 = G_N(-4\pi r \rho(r) + M(r)/r^2) \ , \tag{24.92}$$

and performing these substitutions one obtains the equation

$$\begin{aligned} p'(r) &= -G_N \frac{(\rho(r) + p(r))\left(M(r) + 4\pi r^3 p(r)\right)}{r(r - 2G_N M(r))} \\ &= -\frac{G_N \rho(r) M(r)}{r^2}\left(1 + \frac{p(r)}{\rho(r)}\right)\left(1 + \frac{4\pi r^3 p(r)}{M(r)}\right)\left(1 - \frac{2G_N M(r)}{r}\right)^{-1} \ . \end{aligned} \tag{24.93}$$

This is the famous *Tolman-Oppenheimer-Volkoff (TOV) Equation*, or the *equation of hydrostatic equilibrium*, which determines the pressure in the interior of the star.

<u>Remarks:</u>

1. Note that the right-hand side is manifestly negative so that (reassuringly) the pressure inside the star decreases as one moves to larger values of r.

2. This equation should be supplemented by an equation of state, typically given by an explicit or implicit relation between $\rho(r)$ and $p(r)$. It should also be supplemented by the boundary condition that $p(r_0) = 0$ where r_0 is the radius of the star.

3. Given such an equation of state, in principle one can then integrate the TOV equation and (24.88)
$$\frac{dM(r)}{dr} = 4\pi r^2 \rho(r) \quad , \tag{24.94}$$
for $M(r)$ and $P(r)$. In practice, except for some very special simple equations of state, this needs to be done numerically.

4. The TOV equation can be interpreted as the condition for hydrostatic equilibrium of the star, as it generalises the Newtonian hydrostatic equation
$$p'(r) = -G_N \frac{\rho(r)M(r)}{r^2} \quad . \tag{24.95}$$
The differences between (24.93) and (24.95) can be attributed to (and provide useful insight into) the key differences between Newtonian gravity and general relativity:

   - In general relativity, not only the mass $M(r)$ acts as a source of the gravitational field, but anything that appears in the energy-momentum tensor, in particular in the present situation the pressure $p(r)$. This accounts for the substitution $M(r) \to M(r) + 4\pi r^3 p(r)$.
   - In general relativity, gravity acts not only on $\rho(r)$ but also on $p(r)$. This accounts for the substitution $\rho(r) \to \rho(r) + p(r)$.
   - In general relativity, the gravitational force differs from the Newtonian gravitational force. In the present case this accounts for the additional factor of $f(r) = 1 - 2m(r)/r$ in the denominator.

   All these new terms are suppressed by a factor of c^{-2} relative to the leading Newtonian terms.

We will now consider the solution of this set of equations in the case where the energy density is constant inside the star,
$$\rho(r) = \rho_0 \quad \text{for} \quad r \leq r_0 \quad , \quad \rho(r) = 0 \quad \text{for} \quad r > r_0 \quad . \tag{24.96}$$
This ansatz replaces an assumption about an explicit equation of state relating $\rho(r)$ and $p(r)$, as $p(r)$ can now be determined from the TOV equation.

We begin with $f(r)$ or $m(r)$. Since
$$M(r) = \begin{cases} \frac{4\pi\rho_0}{3} r^3 & \text{for} \quad r \leq r_0 \\ M \equiv \frac{4\pi\rho_0}{3} r_0^3 & \text{for} \quad r \geq r_0 \end{cases} \tag{24.97}$$

$f(r)$ has the form

$$f(r) = 1 - \frac{2m(r)}{r} = \begin{cases} 1 - \frac{8\pi G_N \rho_0}{3} r^2 = 1 - \frac{2G_N M}{r_0}\left(\frac{r}{r_0}\right)^2 & \text{for} \quad r \leq r_0 \\ 1 - \frac{2G_N M}{r} & \text{for} \quad r \geq r_0 \end{cases} \tag{24.98}$$

This already completely determines the spatial part of the interior metric, and matches perfectly with the radial part of the exterior Schwarzschild solution, which has $g_{rr} = f(r)^{-1}$ with $f(r) = 1 - 2G_N M/r$ for all $r > r_0$.

Knowing $m(r)$ (and $\rho(r) = \rho_0$), we can determine $p(r)$ from the TOV equation (24.93), which now reads

$$p'(r) = -\frac{4\pi G_N}{3} \frac{r}{f(r)} (\rho_0 + p(r)) (\rho_0 + 3p(r)) \quad . \tag{24.99}$$

Writing

$$\frac{p'(r)}{(\rho_0 + p(r))(\rho_0 + 3p(r))} = \frac{1}{2\rho_0} \frac{d}{dr} \ln \frac{\rho_0 + 3p}{\rho_0 + p} \tag{24.100}$$

and

$$-\frac{4\pi G_N}{3} \frac{r}{f(r)} = \frac{1}{4\rho_0} \frac{d}{dr} \ln f(r) \quad , \tag{24.101}$$

it follows immediately that the solution satisfying the boundary condition $p(r_0) = 0$ is

$$p(r) = \frac{f(r)^{1/2} - f(r_0)^{1/2}}{3f(r_0)^{1/2} - f(r)^{1/2}} \rho_0 \quad . \tag{24.102}$$

We will come back to one of the consequences of this equation below.

Before turning to that, let us complete the solution of the Einstein equations by determining $h(r)$. This can be found either directly by integration of the second equation in (24.87),

$$h'(r) = 4\pi G_N \, r f(r)^{-1} (\rho_0 + p(r)) \quad , \tag{24.103}$$

or, more efficiently, from (24.91), which we rewrite as

$$\frac{d}{dr}(h(r) + \tfrac{1}{2} \ln f(r)) = -\frac{p'(r)}{\rho_0 + p(r)} \quad \Leftrightarrow \quad \frac{d}{dp}(h + \tfrac{1}{2} \ln f) = -(\rho_0 + p)^{-1} \quad . \tag{24.104}$$

The solution to this equation is evidently

$$\mathrm{e}^{h(r)} f(r)^{1/2} = C(\rho_0 + p(r))^{-1} \tag{24.105}$$

for some integration constant C. Using

$$\rho_0 + p(r) = 2f(r_0)^{1/2} \rho_0 (3f(r_0)^{1/2} - f(r)^{1/2})^{-1} \tag{24.106}$$

and fixing this integration constant by the requirement $h(r_0) = 0$, so that also the (tt)-component of the metric matches onto that of the exterior Schwarzschild solution at $r = r_0$, one finds

$$\mathrm{e}^{h(r)} = \tfrac{1}{2}(3f(r_0)^{1/2}/f(r)^{1/2} - 1) \quad , \tag{24.107}$$

or

$$g_{tt}(r) = -\tfrac{1}{4}(3f(r_0)^{1/2} - f(r)^{1/2})^2 \ . \tag{24.108}$$

This completes the derivation of the solution of the Einstein equations for the interior of a perfect-fluid star with constant energy density.

In summary we have found that the interior metric ($r \leq r_0$) is

$$ds^2 = -\tfrac{1}{4}(3f(r_0)^{1/2} - f(r)^{1/2})^2 dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \tag{24.109}$$

with

$$f(r) = 1 - \frac{r^2}{R^2} = 1 - \frac{8\pi G_N}{3}\rho_0 r^2 \ , \tag{24.110}$$

supported by a perfect fluid matter with $\rho(r) = \rho_0$ and $p(r)$ given in (24.102). This matches continuously (and, in fact, once-differentiably) onto the exterior Schwarzschild solution with $g_{tt}(r) = -f(r)$, $g_{rr} = f(r)^{-1}$, and $f(r) = 1 - 2G_N M/r$, where M is the total mass of the star, $M = 4\pi\rho_0 r_0^3/3$.

Remarks:

1. The spatial part of this interior metric is

$$(ds^2)_{\text{space}} = f(r)^{-1}dr^2 + r^2 d\Omega^2 = \frac{dr^2}{1 - r^2/R^2} + r^2 d\Omega^2 \ , \tag{24.111}$$

with

$$R^2 = \frac{3}{8\pi G_N \rho_0} \ . \tag{24.112}$$

and $r \leq r_0$ (we assume that the energy density ρ_0 is positive, $\rho_0 > 0$). By standard manipulations we can put this metric into the usual form (2.18),

$$r = R\sin\psi \quad \Rightarrow \quad (ds^2)_{\text{space}} = R^2(d\psi^2 + \sin^2\psi d\Omega^2) \quad \text{with} \quad \psi \leq \arcsin(r_0/R) \ , \tag{24.113}$$

of the metric on the 3-sphere. Thus the geometry of the interior of the star is (see section 14) that of a metric of a maximally symmetric space with constant positive curvature, namely the standard metric on the 3-sphere with radius R, restricted to the disc $r \leq r_0$.

Note that in order for the metric to be well-defined, the range of r needs to be such that the maximal value r_0 satisfies $r_0 < R$. Writing R^2 from (24.112) in terms of the total mass M and the radius r_0,

$$R^2 = \frac{3}{8\pi G_N \rho_0} = \frac{r_0^3}{2G_N M} = \frac{r_0^3}{r_s} \ , \tag{24.114}$$

we see that

$$\left(\frac{R}{r_0}\right)^2 = \frac{r_0}{r_s} \ . \tag{24.115}$$

Thus $R > r_0$ iff the radius of the star is larger than its Schwarzschild radius,

$$R > r_0 \quad \Leftrightarrow \quad r_0 > r_s \ . \tag{24.116}$$

This is indeed a necessary condition for a star - an object that is smaller than its Schwarzschild radius will turn out to be not a star but a black hole.

2. A stronger constraint on the relative size of r_0 and r_s arises from an analysis of the solution (24.102) of the TOV equation. Recall that we have $f(r) = 1 - r^2/R^2$. Since $r < r_0$ in the interior, one has $f(r_0) < f(r)$. Thus the pressure can potentially become infinite at values of r where the denominator of (24.102) vanishes.

   The condition for the existence of a stable star, i.e. the requirement that the pressure be non-singular everywhere in the interior of the star, in particular at the origin $r = 0$, is

$$f(r_0)^{1/2} > 1/3 \quad \Leftrightarrow \quad r_0 > \frac{9}{8} r_s \quad \Leftrightarrow \quad M_{max} = \frac{4}{9G_N} r_0 \ . \tag{24.117}$$

   Thus we learn that a star consisting of matter with constant energy density must be larger than 9/8 times its Schwarzschild radius. In other words, the maximal amount of mass that can be contained in a star with radius r_0 is bounded from above by $4r_0/9G_N$, or

$$\frac{2m}{r_0} \leq \frac{8}{9} \ . \tag{24.118}$$

   This result is actually valid for far more general equations of state and is known as the *Buchdahl limit* or *Buchdahl's Theorem* (1959).[76]

3. Here we have considered the simplest model of a static star, with constant energy density. The basic set-up, however, can also be used to analyse in detail the stellar structure and evolution of other compact stars like neutron stars.[77]

4. We will look at some equally idealised models of gravitational collapse of a star, both the exterior solution and a matching interior solution, in section 29. In the latter case, the matter of the star will not be modelled by a constant energy density but rather by pressureless matter, i.e. $p = 0$. A look at (24.93) shows that such an equation of state is evidently incompatible with a static star, i.e. with hydrostatic equilibrium (pressure is required to maintain the star), but it provides a reasonable (free-fall) approximation to a collapsing star.

[76] See e.g. J. Guven, N. O' Murchadha, *Bounds on 2m/R for static spherical objects*, `arXiv:gr-qc/9903067`, H. Andreasson, *Sharp bounds on 2m/r of general spherically symmetric static objects*, `arXiv:gr-qc/0702137`, J. Mark Heinzle, *Bounds on 2m/r for static perfect fluids*, `arXiv:0708.3352 [gr-qc]` for a survey of known results and more recent work along these lines.

[77] See e.g. N. Straumann, *General Relativity: with applications to astrophysics.*

24.8 ADM and Komar Energies of the Schwarzschild Solution

In section 23.4 we had derived two (tentative) expressions for the total energy of an isolated system, namely the ADM energy (23.32)

$$E_{ADM} = \frac{1}{16\pi G_N}\oint_{S^2_\infty} dS_i\,(\partial_k h_{ik} - \partial_i h_{kk}) \quad , \tag{24.119}$$

and the Komar energy (23.41)

$$E_{\text{Komar}}(\Sigma) = -\frac{1}{8\pi G_N}\oint_{S^2_\infty} dS_i\,\partial_i h_{00} = -\frac{1}{8\pi G_N}\oint_{S^2_\infty} dS_{\mu\nu}\,\nabla^\mu K^\nu \quad . \tag{24.120}$$

We can now apply these to the Schwarzschild metric (for which we had already calculated the "canonical" ADM energy in section 21.12).

In standard coordinates the asymptotic behaviour of the spatial part of the metric is

$$(1-2m/r)^{-1}dr^2 + r^2 d\Omega^2 \approx (1+2m/r)dr^2 + r^2 d\Omega^2 \tag{24.121}$$

so that $(2m/r)dr^2$ measures the departure from flat space,

$$(1+2m/r)dr^2 + r^2 d\Omega^2 = dr^2 + r^2 d\Omega^2 + (2m/r)dr^2 = d\vec{x}^2 + (2m/r)dr^2 \tag{24.122}$$

with $r^2 = \vec{x}^2$. Thus one has

$$\frac{2m}{r}dr^2 = \frac{2m}{r^3}x_i x_k dx^i dx^k \quad \Rightarrow \quad h_{ik} = \frac{2m}{r^3}x_i x_k \quad . \tag{24.123}$$

It follows that the ADM integrand is (taking a sphere S^2_r of large but finite radius, and letting $r\to\infty$ at the end)

$$\partial_k h_{ik} - \partial_i h_{kk} = \frac{4m}{r^3}x_i \quad . \tag{24.124}$$

Thus, with $m = G_N M$ and $\oint dS_i x^i = \oint r^2 d\Omega n_i x^i = 4\pi r^3$ one has

$$E_{ADM} = \lim_{r\to\infty}\frac{1}{16\pi G_N}\oint_{S^2_r} dS_i \frac{4m}{r^3}x^i = \lim_{r\to\infty}\frac{1}{16\pi G_N}\frac{4m}{r^3}4\pi r^3 = M \quad . \tag{24.125}$$

Since the ADM expression for the energy appeals to an asymptotically Cartesian coordinate system, this is one instance where it is perhaps more natural (and safer) to use the Schwarzschild metric in isotropic coordinates (24.46). The asymptotic ($\rho\to\infty$, with $\rho^2 = \vec{x}^2$) behaviour of the spatial part of the metric is

$$(1+m/2\rho)^4 d\vec{x}^2 \approx (1+2m/\rho)d\vec{x}^2 \quad \Rightarrow \quad h_{ik} = \frac{2m}{\rho}\delta_{ik} \quad . \tag{24.126}$$

Note that this is not the same as (24.123). Nevertheless, calculating $\partial_k h_{ik} - \partial_i h_{kk}$ in this case, one finds the same expression as in Schwarzschild coordinates (with $r\to\rho$),

$$\partial_k h_{ik} - \partial_i h_{kk} = \frac{4m}{\rho^3}x_i \quad , \tag{24.127}$$

and the remainder of the calculation is then identical, leading again to $E_{ADM} = M$. This agrees with the result of the calculations of the canonical ADM energy of the Schwarzschild metric in section 21.12.

From the alternative Komar expression (23.41) for the energy, this (eminently reasonable and respectable) result arises not from the asymptotic behaviour of the spatial components of the metric but instead form that of the (00)-component of the metric (the relation between the two being provided by the Einstein equations). We have

$$g_{00} = -1 + \frac{2m}{r} \quad \Rightarrow \quad h_{00} = \frac{2m}{r} \quad , \tag{24.128}$$

leading to

$$\oint_{S_r^2} dS_i \, \partial_i h_{00} = 4\pi r^2 \partial_r (2m/r) = -8\pi m \quad . \tag{24.129}$$

Note that for the special case of the Schwarzschild metric, for which (24.128) is exact (and not just true asymptotically), this is independent of r, and thus

$$E_{\text{Komar}} = -\lim_{r\to\infty} \frac{1}{8\pi G_N} \oint_{S_r^2} dS_i \, \partial_i h_{00} = -\frac{1}{8\pi G_N} \oint_{S_r^2} dS_i \, \partial_i h_{00} = M \quad . \tag{24.130}$$

However, more generally, for any metric with the asymptotic behaviour

$$g_{00} = -1 + \frac{a}{r} + \mathcal{O}(r^{-2}) \tag{24.131}$$

this calculation shows that the mass / energy of the solution is determined by the $1/r$ term of the (00)-component of the metric,

$$E_{\text{Komar}} = -\lim_{r\to\infty} \frac{1}{8\pi G_N} \oint_{S_r^2} dS_i \, \partial_i h_{00} = a/2G_N \quad . \tag{24.132}$$

In particular, this applies e.g. to the Reissner-Nordstrøm metric of section 31 and shows that the parameter m appearing in the solution

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2} \tag{24.133}$$

has the interpretation of the total energy of the system, $E = m/G_N$, even in the presence of charge and electrostatic fields.

This also generalises to higher dimensions. Thus in $D = d+1$ dimensions, the mass is determined by the coefficient of r^{2-d} in g_{00}, as in the higher-dimensional Schwarzschild-Tangherlini black hole (30.18) with $f(r) = 1 - \mu/r^{d-2}$. The proportionality factor between μ and the mass is also dimension-dependent, as it involves the area of the transverse $(d-1)$-sphere.

Remarks:

1. The independence of E_{Komar} in (24.130) of r, more generally of the 2-surface over which one integrates in the Schwarzschild case (indeed, one could choose any 2-surface in the region $r > 2m$ enclosing the black hole) can be understood as a

consequence of the relation (13.11), which says that

$$\nabla_\nu(\nabla^\mu K^\nu) = R^\mu_{\ \nu} K^\nu \quad . \tag{24.134}$$

In particular, therefore, in source-free regions of space-time ($T_{\mu\nu} = 0 \Rightarrow R_{\mu\nu} = 0$) one has $\nabla_\nu(\nabla^\mu K^\nu) = 0$ and by the usual arguments the surface integral of $\nabla_\mu K_\nu$ is then independent of the choice of surface.

The Reissner-Nordstrøm metric, on the other hand, is a solution of the combined Einstein-Maxwell equations with a non-trivial energy-momentum tensor throughout space-time. In this case $\nabla_\nu(\nabla^\mu K^\nu) \neq 0$, and the Komar integral will depend on the radius of the sphere, say, as different surface-integrals will enclose different amounts of electrostatic energy.

2. One may be a bit concerned by the fact that we obtained a non-zero result for the ADM or Komar energy for the Schwarzschild metric, allegedly a solution of the vacuum Einstein equations, even though in order to arrive at the expressions for the energy, we started off with a non-trivial energy-momentum tensor, either in the linearised theory, with $T_{00} = \rho \neq 0$, or via the Komar charge associated to the current

$$J^\mu = R^\mu_{\ \nu} K^\nu = 8\pi G_N (T^\mu_{\ \nu} - \tfrac{1}{2}\delta^\mu_{\ \nu} R) K^\nu \tag{24.135}$$

(which is identically zero for a vacuum solution, so it appears that one doesn't have a leg to stand on in that case).

This seeming conflict can be resolved in a number of ways. Let us being with the ADM energy.

- If the Schwarzschild metric describes the gravitational field outside a star, then the constant time 3-surface Σ also includes the interior of the star, and one will get just the right contribution to the mass (namely M) from the interior solution (see the remarks in sections 24.6 and 24.7) in order to match onto the exterior Schwarzschild solution with parameter $m = G_N M$.
- If there is no star but one is dealing with a black hole instead (see section 27), the situation is a bit more subtle. In that case, if Σ is a complete constant time 3-surface (with respect to Kruskal time, say), in particular a surface that avoids the singularity at $r = 0$, then it naturally extends to the mirror region III of the Kruskal-Schwarzschild space-time (section 27.8). In this case the fact that the total integral should be (and is) zero just says that one gets the same contibution with opposite signs from the spatial infinities in regions I and III respectively. Any local observer only has access to one of those regions. Thus the fact that the total integral is zero is irrelevant, and the integral over the surface S^2_∞ of region I gives $E = M$, which is the physically relevant statement.

Now let us turn to the Komar energy. The point to realise (recall) is that it is not the conserved Komar current J^μ (which would indeed be vanishing identically for a vacuum solution) which is the crucial quantity but rather the object $A^{\mu\nu} = \nabla^\mu K^\nu$ satisfying (24.134), and its surface integral. If one chooses the 3-volume V to remain outside the star (or black hole), bounded by 2 spheres at radii $r_1 > r_0$ (radius of the star) or $r_1 > r_s$, and $r_2 > r_1$, then the fact that $J^\mu = 0$ in that region just reproduces the statement that the 2-surface integral is independent of the radius but does not preclude a non-zero value of the integral over either one of those surfaces.

25 Particle and Photon Orbits in the Schwarzschild Geometry

We have now accumulated all the ingredients to address an issue of fundamental interest in general relativity, namely the study of planetary orbits and lightrays in the gravitational field of the sun, i.e. the properties of timelike and null geodesics of the Schwarzschild geometry.

We shall see that, by making good use of the symmetries of the problem, we can reduce the (initially somewhat complicated looking) set of coupled 2nd order geodesic equations to a single 1st order differential equation in one variable, analogous to that for a one-dimensional particle moving in a particular ("effective") potential. Solutions to this equation can then readily be discussed qualitatively and also quantitatively (analytically).

In order to be able to read this section as a concrete ("real-life", so to speak) application and illustration of the general formalism for geodesics introduced in sections 2 and 3 (without knowledge of all the intervening material on tensor analysis and the Einstein equations, say), this section starts with a brief summary of the key features and main properties of the Schwarzschild metric that we will make use of in the following and that were established in detail in the previous section 24.

25.1 Schwarzschild Metric: Summary of its Key Properties

In standard Schwarzschild coordinates (t, r, θ, ϕ), the Schwarzschild metric is

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \tag{25.1}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the line element of the standard (rotationally invariant) metric on the 2-sphere S^2, and $f(r)$ is the function

$$f(r) = 1 - \frac{2m}{r} \quad , \tag{25.2}$$

with m an integration constant (with the dimension of length).

This metric has the following origin / interpretation / properties:

1. As shown in section 24.2, via a suitable choice of coordinates, any metric that is static and spherically symmetric (and thus describes a time-independent spherically symmetric gravitational field) can be put into the form

   $$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2 d\Omega^2 \quad , \tag{25.3}$$

   for some (positive) functions $A(r)$ and $B(r)$.

   Provided that $A(r) \to 1$ and $B(r) \to 1$ (sufficiently rapidly) as $r \to \infty$, this metric is asymptotically Minkowskian (usually referred to as *asymptotically flat*).

An anymptotically flat metric can be considered to describe the gravitational field of an isolated/localised object in space.

In particular, therefore, the Schwarzschild metric is static, spherically symmetric and asymptotically flat.

2. What singles out the Schwarzschild metric among all such metrics is the crucial fact (established in section 24.3) that

The Schwarzschild metric is the unique static, spherically symmetric and asymptotically flat solution of the vacuum Einstein field equations!

[This statement can be sharpened somewhat: the assumption that the metric is static turns out to be unnecessary - see the discussion of Birkhoff's theorem in section 24.6.]

Here "vacuum Einstein field equations" refers to the field equations for the metric in the absence of matter or in matter-free regions of space-time. Thus the physical interpretation of the Schwarzschild metric is that it describes the gravitational field in the exterior region $r > r_0$ of a static spherically symmetric star of radius r_0.

3. Comparison of the Schwarzschild

$$g_{00} = -\left(1 - \frac{2m}{r}\right) \tag{25.4}$$

with the Newtonian limit, where one has (cf. the discussion in section 3.3)

$$g_{00} = -\left(1 + \frac{2\phi}{c^2}\right) \tag{25.5}$$

shows that this physical interpretation is perfectly compatible with the Newtonian limit, where the appropriate Newtonian potential ϕ for a star with mass M has the form

$$\phi(r) = -\frac{G_N M}{r} \quad . \tag{25.6}$$

In particular, the physical interpretation of the integration constant m appearing in the Schwarzschild solution is that it is related to the mass M by

$$m = \frac{G_N M}{c^2} \quad , \tag{25.7}$$

The Schwarzschild metric is thus the extension of the Newtonian $1/r$-potential, the unique exact spherically symmetric solution of the Newtonian "vacuum" field equation

$$\Delta\phi(r) = 0 \quad \text{for} \quad r > r_0 \tag{25.8}$$

vanishing as $r \to \infty$, to an exact solution of the general relativistic vacuum field equations for the metric.

4. Here is what we can say about the coordinates:

   - The Schwarzschild coordinates are adapted to the symmetries of the metric in the sense that invariance under time translations $t \to t + b$ and rotations of the angular coordinates (θ, ϕ) is manifest, or as manifest as possible.
   - The coordinates (θ, ϕ) have the standard interpretation and range as coordinates on the 2-sphere.
   - Even though r is not a measure of proper radial distance, it does have a clean geometric interpretation, namely that r is such that spheres of radius r are those that have area $4\pi r^2$.
   - A priori the range of r is limited by the requirement $r > r_0$ (because the Schwarzschild metric only describes the gravitational field to the exterior of the star). One might also be concerned about the behaviour of the metric as $r \to 2m$ but, as recalled below, for standard astrophysical objects like planets or stars one has $r_0 \gg 2m$, so that the condition $r > r_0$ automatically excludes this potentially dangerous region.
   - Since the metric is independent of t, the range of the coordinate t can be taken to be $(-\infty, +\infty)$. Up to affine transformations $t \to at + b$, t is also characterised by the fact that the time-translation symmetry is realised by translations $t \to t + b$ of the coordinate t (and not some more complicated transformation of the coordinates).
   - The physical interpretation of t is that it is the proper time of a Minkowskian observer "at infinity" (where $f(r) \to 1$) and that, up to a constant conversion factor $\sqrt{f(r)} < 1$ it is the proper time of observers hovering at fixed values of the spatial coordinates (r, θ, ϕ).

   Since symmetries are manifest, and the coordinates have a simple interpretation, these coordinates are ideally suited for studying the geodesics, i.e. the motion of particles and light, in the gravitational field of an object like the sun, with $r_0 \gg 2m$, and we will embark on this below.

   However, for other purposes other coordinates may be more useful or even necessary and can provide additional insight. We will look at this in quite some detail in subsequent sections.

5. As shown in (24.40), the apparently problematic behaviour of the Schwarzschild metric as $r \to 2m$ is irrelevant for applications to the solar system (or the description of the gravitational field of other standard benign astrophysical objects), since for such objects m is much less than the radius of the object, $m \ll r_0$ (while the Schwarzschild metric only applies to the exterior region $r > r_0$).

   E.g. for the earth m is of the order of a centimeter, while for an object with a mass approximately that of the sun it is of the order of a couple of kilometers.

However, we will take a closer look at what happens at (and beyond) $r = 2m$ in section 26 and subsequent sections, when exploring the physics of black holes.

25.2 Symmetries and Effective Potential for the Kepler Problem

To set the stage for the subsequent discussion of the motion of particles (and light) in the Schwarzschild geometry, let us very briefly recall how one goes about this in the corresponding (Kepler) problem in Newtonian mechanics, namely the motion of particles in the potential (25.6).

The crucial point here is to make judicious use of the symmetries of the problem to reduce the problem of the 3-dimensional motion of a particle in the spherically symmetric potential $V(r)$ to that of a 1-dimensional radial motion in a suitable *effective potential*. Then the problem has been reduced to that of solving a single ordinary differential equation (and this can be done in closed form for the Kepler problem), and even without knowledge of the exact solutions the qualitative behaviour of the orbits in the Schwarzschild geometry can essentially be determined "by inspection" from drawing the effective potential.

The steps are the following:

1. Because of time-translation invariance, the Newtonian energy E_N of a particle moving in an arbitary time-independent potential $V(\vec{x})$ is conserved. This energy is the sum of the kinetic energy and the potential energy. Passing straightaway to spherical coordinates (r, θ, ϕ), and just looking at a particle of unit mass (since we already know that in the case of the Kepler problem the mass of the particle will drop out of all equations anyway) this energy is
$$E_N = \frac{1}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\dot{\phi}^2\right) + V(r, \theta, \phi) \quad , \tag{25.9}$$
where (in this Newtonian context) an overdot refers to an ordinary time-derivative, $\dot{r} = dr/dt$ etc.

2. When the potential is purely radial (spherically symmetric), $V = V(r)$, the problem is invariant under spatial rotations, and correspondingly angular momentum $\vec{L}$ is conserved. Thus motion takes place in the plane orthogonal to $\vec{L}$. For any choice of $\vec{L}$ one can choose the coordinate system such that $\vec{L}$ points into the x^3-direction or, equivalently, such that the motion is restricted to the equatorial plane $\theta = \pi/2$,
$$\theta = \pi/2 \quad \Rightarrow \quad \vec{L} = (0, 0, L) \quad , \quad L = r^2\dot{\phi} \ . \tag{25.10}$$

3. Plugging this back into the energy, one finds
$$E_N = \frac{1}{2}\dot{r}^2 + \frac{L^2}{2r^2} + V(r) \equiv \frac{1}{2}\dot{r}^2 + V_{eff}(r) \quad , \tag{25.11}$$

where $V_{eff}(r)$ is known as the *effective potential*. It governs the radial motion of the particle in the potential $V(r)$ and differs from $V(r)$ by the usual centrifugal / angular momentum barrier term (which acts as a potential for the corresponding centrifugal "pseudo-force").

Thus the symmetries have allowed us to reduce the original set of 3 2nd-order differential equations for the coordinates $\vec{x} = \vec{x}(t)$ to a single ordinary 1st-order differential equation for $r = r(t)$,

$$\frac{dr}{dt} = \pm\sqrt{2E_N - 2V_{eff}(r)} \ . \tag{25.12}$$

4. To obtain an equation for the *orbit*, i.e. for $r = r(\phi)$, one can use

$$\frac{dr}{d\phi} = \frac{\dot{r}}{\dot{\phi}} = \pm\frac{r^2}{L}\sqrt{2E_N - 2V_{eff}(r)} \ . \tag{25.13}$$

As is well known, for the Kepler problem, with $V(r) \sim 1/r$, this equation can be solved in closed form in terms of conical sections (this is recalled at the beginning of section 25.8).

25.3 Symmetries and Effective Potential for Schwarzschild Geodesics

Our aim is now to try to develop a similarly efficient strategy to deal with the corresponding general relativistic problem, i.e. for describing the solutions of the geodesic equation in the Schwarzschild geometry.

A convenient starting point in general for discussing geodesics is, as I stressed repeatedly, the Lagrangian $\mathcal{L} = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$. For the Schwarzschild metric this is

$$\mathcal{L} = -(1 - 2m/r)\dot{t}^2 + (1 - 2m/r)^{-1}\dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \ , \tag{25.14}$$

where $2m = 2MG_N/c^2$.

Rather than writing down and solving the (second order) geodesic equations, we will make use of the conserved quantities associated with the continuous symmetries of the Lagrangian (or the metric). After all, conserved quantities correspond to first integrals of the equations of motion and if there are a sufficient number of them (there are) we can directly reduce the second order differential equations to first order equations.

1. Spherical Symmetry and Conserved Angular Momentum

   Since the gravitational field is spherically symmetric, or isotropic, there is conservation of angular momentum. Thus, exactly as in the Newtonian problem, the orbits of the particles or planets are planar. Without loss of generality, we can choose our coordinates in such a way that this plane is the equatorial plane $\theta = \pi/2$, so in particular $\dot{\theta} = 0$ at all times.

In case this is not obvious (even though it should be), here are two explicit ways to establish this:

- One can certainly choose one's coordinates in such a way that at some initial time $\tau = \tau_0$ one has $\theta(\tau_0) = \pi/2$ and that the angular velocity $\dot{\theta}(\tau_0) = 0$. It then follows from the Euler-Lagrange equations of motion for $\theta(\tau)$ that they are solved at all times by $\theta(\tau) = \pi/2$, $\dot{\theta}(\tau) = 0$.
- If you already know about Killing vectors (which made their first appearance in this context in sections 3.2), you may find it somewhat more insightful to use the conserved angular momenta

$$L_{(a)} = g_{\alpha\beta}\dot{x}^{\alpha}V_{(a)}^{\beta} \tag{25.15}$$

associated to the Killing vectors $V_{(a)}$ (9.55),

$$\begin{aligned} L_{(1)} &= r^2\left(-\cos\phi\,\dot{\theta} + \cot\theta\sin\phi(\sin^2\theta\,\dot{\phi})\right) \\ L_{(2)} &= r^2\left(+\sin\phi\,\dot{\theta} + \cot\theta\cos\phi(\sin^2\theta\,\dot{\phi})\right) \\ L_{(3)} &= r^2\sin^2\theta\,\dot{\phi}\ \ , \end{aligned} \tag{25.16}$$

and to use spherical symmetry to rotate the $L_{(a)}$ into the form

$$(L_{(1)}, L_{(2)}, L_{(3)}) = (0, 0, L)\ \ . \tag{25.17}$$

Using the explicit expressions (25.16), it is straightforward to see that $L_{(1)} = L_{(2)} = 0$ implies $\theta = \pi/2$, $\dot{\theta} = 0$.

Either way, we have fixed the direction of $\vec{L}$, the motion is now restricted to the equatorial plane, with

$$\theta = \pi/2 \quad \Rightarrow \quad \vec{L} = (0, 0, L)\ \ , \tag{25.18}$$

and the residual Lagrangian to deal with is

$$\mathcal{L} = -(1-2m/r)\dot{t}^2 + (1-2m/r)^{-1}\dot{r}^2 + r^2\dot{\phi}^2\ \ . \tag{25.19}$$

The magnitude L (or z-component of $\vec{L}$) of the angular momentum (per unit rest mass) is now the conserved quantity

$$L = r^2\dot{\phi} \tag{25.20}$$

associated to the cyclic variable ϕ of the residual Lagrangian,

$$\frac{\partial\mathcal{L}}{\partial\phi} = 0 \Rightarrow \frac{d}{d\tau}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = 0\ \ . \tag{25.21}$$

2. Time Translation Invariance and Conserved Energy

   The Lagrangian (both the original Lagrangian and the reduced Lagrangian) is invariant under translations of t, which is also a cyclic variable,

   $$\frac{\partial \mathcal{L}}{\partial t} = 0 \Rightarrow \frac{d}{d\tau}\frac{\partial \mathcal{L}}{\partial \dot{t}} = 0 \quad . \tag{25.22}$$

   This gives rise to the conserved energy (with a conventional choice of normalisation)

   $$E = (1 - 2m/r)\dot{t} \quad . \tag{25.23}$$

   Calling L the angular momentum (per unit rest mass) required no further justification, but let me pause to explain in what sense E is an energy (per unit rest mass).

   - On the one hand, it is the conserved quantity (3.7) associated to time-translation invariance. As such, it certainly deserves to be called the energy.
   - It is moreover true that for a particle at infinity ($r \to \infty$) E is just the special relativistic energy $E = \gamma(v_\infty)c^2$, with $\gamma(v) = (1 - v^2/c^2)^{-1/2}$ the usual relativistic γ-factor, and v_∞ the coordinate velocity dr/dt at infinity. This can be seen in two ways. First of all, for a particle that reaches $r = \infty$, the constant E can be determined by evaluating it at $r = \infty$. It thus follows from the definition of E that

     $$E = \dot{t}_\infty \quad . \tag{25.24}$$

     In Special Relativity, the relation between proper and coordinate time is given by (setting $c = 1$ again)

     $$d\tau = \sqrt{1 - v^2}dt \quad \Rightarrow \quad \dot{t} = \gamma(v) \quad , \tag{25.25}$$

     suggesting the identification

     $$E = \gamma(v_\infty) \qquad (E = \gamma(v_\infty)c^2 \quad \text{if} \quad c \neq 1) \tag{25.26}$$

     Another argument for this identification will be given below, once we have introduced the effective potential.

3. τ-Translation Invariance and the Conserved Lagrangian $\mathcal{L}$

   As we have seen in section 2.5 (and again in section 5.8), there is also always one more integral of the geodesic equation, namely $\mathcal{L}$ itself,

   $$\frac{d}{d\tau}\mathcal{L} = 2g_{\mu\nu}\dot{x}^\mu D_\tau \dot{x}^\nu = 0 \tag{25.27}$$

   (this can be interpreted as the conserved "Hamiltonian" associated with the invariance of $\mathcal{L}$ under translations of the affine parameter). Thus we set

   $$\mathcal{L} = \epsilon \quad , \tag{25.28}$$

where $\epsilon = -1$ for timelike geodesics and $\epsilon = 0$ for null geodesics. We thus have

$$-(1-2m/r)\dot{t}^2 + (1-2m/r)^{-1}\dot{r}^2 + r^2\dot{\phi}^2 = \epsilon \ , \tag{25.29}$$

Putting everything together, we can now express $\dot{t}$ and $\dot{\phi}$ in terms of the conserved quantities E and L to obtain a first order differential equation for r alone, namely

$$-(1-2m/r)^{-1}E^2 + (1-2m/r)^{-1}\dot{r}^2 + \frac{L^2}{r^2} = \epsilon \ . \tag{25.30}$$

Multiplying by $(1-2m/r)/2$ and rearranging the terms, one obtains

$$\frac{E^2+\epsilon}{2} = \frac{\dot{r}^2}{2} + \epsilon\frac{m}{r} + \frac{L^2}{2r^2} - \frac{mL^2}{r^3} \ . \tag{25.31}$$

Now this equation is of the familiar Newtonian form

$$E_{eff} = \frac{\dot{r}^2}{2} + V_{eff}(r) \ , \tag{25.32}$$

with

$$\begin{aligned} E_{eff} &= \frac{E^2+\epsilon}{2} \\ V_{eff}(r) &= \epsilon\frac{m}{r} + \frac{L^2}{2r^2} - \frac{mL^2}{r^3} \ , \end{aligned} \tag{25.33}$$

describing the energy conservation for the 1-dimensional motion in an effective potential. Except for $t \to \tau$, this is exactly the same as the Newtonian equation of motion in a potential

$$V(r) = \epsilon\frac{m}{r} - \frac{mL^2}{r^3} \ , \tag{25.34}$$

the effective angular momentum term $L^2/r^2 = r^2\dot{\phi}^2$ arising, as usual, from the change to polar coordinates. As in the corresponding Newtonian one-dimensional (radial) problem, the qualitative behaviour of the orbits in the Schwarzschild geometry can thus essentially be determined "by inspection".

Given that in principle we started off with the four coupled non-linear geodesic differential equations, this is an enormous and enormously useful simplification, and the main result of this section.

Remarks:

1. In particular, for $\epsilon = -1$, the general relativistic motion (as a function of τ) is exactly the same as the Newtonian motion (as a function of t) in the potential

$$\epsilon = -1 \ \Rightarrow \ V(r) = -\frac{m}{r} - \frac{mL^2}{r^3} \ . \tag{25.35}$$

The first term is just the ordinary Newtonian potential, so the second term is apparently a general relativistic correction. We will later on treat this as a perturbation but note that the above is an exact result, not an approximation (so, for example, there are no higher order corrections proportional to higher powers of m/r). We expect observable consequences of this general relativistic correction because many properties of the Newtonian orbits (Kepler's laws) depend sensitively on the fact that the Newtonian potential is precisely $\sim 1/r$.

2. In order to see that the new 3rd term really is a (very) small correction to the Newtonian effective potential, note first of all that, restoring a factor of c^2 in the first (Newtonian) potential term, we can write the effective potential as

$$V_{eff}(r) = -\frac{G_N M}{r} + \frac{L^2}{2r^2} - \frac{m}{r}\frac{L^2}{r^2} \quad . \tag{25.36}$$

For the planetary orbits, there is a balance between the attractive 1st and repulsive 2nd term, so these two terms are of the same order of magnitude. The 3rd term is then smaller than these by a factor (m/r). Since r (the radial coordinate of the planet) is much larger than r_0 (the radius of the sun) and, as we had already seen before, for standard astrophysical objects like the sun the radius r_0 is much larger than the Schwarzschild radius $2m$, we have

$$r \gg r_0 \gg m \quad \Rightarrow \quad \frac{m}{r} \ll 1 \quad . \tag{25.37}$$

Thus, the general relativistic corrction term $\sim mL^2/r^3$ indeed provided a very small (but important) correction to the Newtonian potential.

3. Looking at the equation for $\epsilon = -1$,

$$\tfrac{1}{2}\dot{r}^2 + V_{eff}(r) = \tfrac{1}{2}(E^2 - 1) \tag{25.38}$$

and noting that $V_{eff}(r) \to 0$ for $r \to \infty$, we can read off that for a particle that reaches $r = \infty$ we have the relation

$$\dot{r}_\infty^2 = E^2 - 1 \quad . \tag{25.39}$$

This implies, in particular, that for such (scattering) trajectories one necessarily has $E \geq 1$, with $E = 1$ corresponding to a particle initially or finally at rest at infinity. For $E > 1$ the coordinate velocity at infinity can be computed from

$$v_\infty^2 = \frac{\dot{r}_\infty^2}{\dot{t}_\infty^2} \quad . \tag{25.40}$$

Using (25.24) and (25.39), one finds

$$v_\infty^2 = \frac{E^2 - 1}{E^2} \quad \Leftrightarrow \quad E = (1 - v_\infty^2)^{-1/2} \quad , \tag{25.41}$$

thus confirming the result claimed in (25.26).

4. From the effective potential we can deduce the radial geodesic equation of motion, namely

$$\epsilon = -1 \quad \Rightarrow \quad \ddot{r} = -V'_{eff}(r) = -\frac{m}{r^2} + (r - 3m)\frac{L^2}{r^4} \quad . \tag{25.42}$$

5. For null geodesics, $\epsilon = 0$, on the other hand, the Newtonian part of the potential is zero, as one might expect for massless particles, but in General Relativity a photon with $L \neq 0$ feels a non-trivial potential

$$\begin{aligned} \epsilon = 0 \quad \Rightarrow \quad V(r) &= -\frac{mL^2}{r^3} \\ V_{eff}(r) &= f(r)\frac{L^2}{2r^2} \quad . \end{aligned} \tag{25.43}$$

6. It will also turn out to be useful to have the full radial part of the null geodesic equation at our disposal. As in the timelike case above, this can simply be read off from the effective potential and one finds

$$\epsilon = 0 \quad \Rightarrow \quad \ddot{r} = -V'_{eff}(r) = \frac{L^2}{r^4}(r - 3m) \quad . \tag{25.44}$$

   Obviously something potentially interesting is happening at $r = 3m$, provided the star is sufficiently small so that its radius $r_0 < 3m$ of course - see section 25.10). Also note that this differs from its timelike counterpart (25.42) precisely by the absence of the Newtonian term $-m/r^2$.

7. Finally, we note that the success of the above analysis relied only on the symmetries of the metric, not on the fact that the particular metric we were looking at satisfies the vacuum Einstein equations. It is indeed reasonably straightforward to generalise the preceding analysis to arbitrary static spherically symmetric metrics. As mentioned before, among these general static spherically symmeric metrics the class of metrics (24.75)

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \tag{25.45}$$

   is of particular importance and interest. For these metrics one finds that the geodesic equation can still be reduced to the effective potential form (25.32), with E_{eff} precisely as in (25.33), and V_{eff} now given by

$$V_{eff}(r) = -\epsilon\phi(r) + \frac{L^2}{2r^2} + \phi(r)\frac{L^2}{r^2} \quad , \tag{25.46}$$

   where we have written $f(r)$ in terms of the corresponding "Newtonian" potential $\phi(r)$ as

$$f(r) = 1 + 2\phi(r) \quad . \tag{25.47}$$

   For the Schwarzschild metric one has $\phi(r) = -m/r$ and (25.46) reduces to (25.33).

25.4 Caveat: Effective Potential versus Routhian

Because this tends to cause some confusion, not just among students but also occasionally in the literature, before turning to a more detailed analysis of the consequences of the equations we have just derived, I want to insert a cautionary remark, a *caveat*, regarding the above derivation of the effective potential equation:

Namely, starting from the Lagrangian $\mathcal{L}$ (25.14), considerations about spherical symmetry led us to choose $\theta = \pi/2$, and imposing this condition in the Lagrangian we were led to the reduced Lagrangian in (25.19). This is legitimate, since such *holonomic* constraints can always be inserted into the Lagrangian itself (this is one of the main virtues and advantages of the Lagrangian formalism).

Subsequently we used the conserved quantities E and L associated to the cyclic variables t and ϕ in order to eliminate from the Lagrangian the quantities $\dot{t}$ and $\dot{\phi}$. Here one needs to be more careful. In general such non-holonomic constraints like $r^2\dot{\phi} = L$ cannot be inserted into the Lagrangian to obtain a (reduced) Lagrangian from which the equations of motion for the remaining variables (here r) can be obtained as the Euler-Lagrange equations. One correct way to do this is to pass to what is known as the *Routhian*, a partial Legendre transform of the Lagrangian on the cyclic variables.

As a simple example, consider the Lagrangian of a (unit mass) free particle in 2 dimensions, expressed in polar coordinates. Its Lagrangian is

$$\mathcal{L} = \tfrac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) \quad . \tag{25.48}$$

The variable ϕ is cyclic, leading to the conserved angular momentum

$$p_\phi = r^2\dot{\phi} = L \quad , \tag{25.49}$$

and the equation of motion for r is the Euler-Lagrange equation

$$\ddot{r} = r\dot{\phi}^2 \quad . \tag{25.50}$$

It is legitimate to eliminate $\dot{\phi}$ in this equation, leading to the centrifugal force equation

$$\ddot{r} = L^2/r^3 \quad . \tag{25.51}$$

Had one eliminated $\dot{\phi}$ in the Lagrangian instead, one would have found the reduced Lagrangian and resulting Euler-Lagrange equations

$$\mathcal{L} \to \tfrac{1}{2}(\dot{r}^2 + L^2/r^2) \quad \Rightarrow \quad \ddot{r} = -L^2/r^3 \qquad \text{(wrong!)} \quad . \tag{25.52}$$

However, even though you might have gained that impression, this is of course not what we were doing in our derivation of the radial effective potential equation. There we were using this elimination in conjunction with the condition that $\mathcal{L}$ is constant on solutions, to obtain a first-order equation for r (the effective potential equation), and this *is* legitimate. This is a 1st integral of the 2nd order radial equation, the 2nd order equations for r follow from this simply by differentiating with respect to proper time τ, and *not* by treating $\dot{r}^2/2 + V_{eff}(r)$ as a Lagrangian and looking at its Euler-Lagrange equations.

And finally, just to illustrate the claim about the Routhian in the above example, the partial Legendre transform with respect to the cyclic variable ϕ of $\mathcal{L}$ is (expressing $\dot{\phi}$ in terms of p_ϕ)

$$\mathcal{R} = \mathcal{L} - p_\phi\dot{\phi} = \tfrac{1}{2}(\dot{r}^2 - p_\phi^2/r^2) \to \tfrac{1}{2}(\dot{r}^2 - L^2/r^2) \tag{25.53}$$

(note the sign flip compared with the "wrong" Lagrangian above). This is the correct reduced Lagrangian, whose Euler-Lagrange equations give rise to the correct radial equation of motion.

25.5 Equation for the Shape of the Orbit

Typically, one is primarily interested in the shape of an orbit, that is in the radius r as a function of ϕ, $r = r(\phi)$, rather than in the dependence of, say, r on some extraterrestrial's proper time τ. In this case, the above mentioned difference between t (in the Newtonian theory) and τ (here) is irrelevant: In the Newtonian theory one uses $L = r^2 d\phi/dt$ to express t as a function of ϕ, $t = t(\phi)$ to obtain $r(\phi)$ from $r(t)$. In General Relativity, one uses the analogous equation $L = r^2 d\phi/d\tau$ to express τ as a function of ϕ, $\tau = \tau(\phi)$. Hence the shapes of the General Relativity orbits are precisely the shapes of the Newtonian orbits in the potential (25.34). Thus we can use the standard methods of Classical Mechanics to discuss these general relativistic orbits and of course this simplifies matters considerably.

To obtain r as a function of ϕ we proceed as indicated above. Thus we use

$$\left(\frac{dr}{d\phi}\right)^2 = \frac{\dot{r}^2}{\dot{\phi}^2} \tag{25.54}$$

to combine (25.32),

$$\dot{r}^2 = 2E_{eff} - 2V_{eff}(r) \quad , \tag{25.55}$$

and (25.20),

$$\dot{\phi}^2 = \frac{L^2}{r^4} \tag{25.56}$$

into

$$\frac{r'^2}{r^4}L^2 = 2E_{eff} - 2V_{eff}(r) \tag{25.57}$$

where a prime denotes a ϕ-derivative.

In the examples to be discussed below, we will be interested in the angle $\Delta\phi$ swept out by the object in question (a planet or a photon) as it travels along its trajectory between the farthest distance r_2 from the star (sun) ($r_2 = \infty$ for scattering trajectories) and the position of closest approach to the star r_1 (the perihelion or, more generally, if we are not talking about our own solar system, periastron), and back again,

$$\Delta\phi = 2\int_{r_1}^{r_2} \frac{d\phi}{dr} dr \quad . \tag{25.58}$$

In the Newtonian case, these integrals can be evaluated in closed form. With the general relativistic correction term, however, these are elliptic integrals which cannot be expressed in closed form. A perturbative evaluation of these integrals (treating the exact general relativistic correction as a small perturbation) also turns out to be somewhat delicate since e.g. the limits of integration depend on the perturbation.

It is somewhat simpler to deal with this correction term not at the level of the solution (integral) but at the level of the corresponding differential equation. As in the Kepler problem, it is convenient to make the change of variables

$$u = \frac{1}{r} \qquad u' = -\frac{r'}{r^2} \quad . \tag{25.59}$$

Then (25.57) becomes

$$u'^2 = L^{-2}(2E_{eff} - 2V_{eff}(r)) \quad . \tag{25.60}$$

Upon inserting the explicit expression for the effective potential, this becomes

$$u'^2 + u^2 = \frac{E^2 + \epsilon}{L^2} - \frac{2\epsilon m}{L^2} u + 2mu^3 \quad . \tag{25.61}$$

This can be used to obtain an equation for $d\phi(u)/du = u'^{-1}$, leading to

$$\Delta\phi = 2\int_{u_2}^{u_1} \frac{d\phi}{du} du \quad . \tag{25.62}$$

Differentiating (25.61) once more, one finds

$$u'(u'' + u) = u'(-\frac{\epsilon m}{L^2} + 3mu^2) \quad . \tag{25.63}$$

Thus either $u' = 0$, which corresponds to a circular orbit of constant radius (i.e. the solution is $u(\phi) = u_0$ or $r(\phi) = r_0$)), and this is not only a trivial but also an irrelevant solution since neither the planets nor the photons of interest to us travel on circular orbits, or

$$u'' + u = -\frac{\epsilon m}{L^2} + 3mu^2 \quad . \tag{25.64}$$

The unperturbed ("Newtonian") equation is the linear equation obtained by dropping the last term, and its solutions $u_0(\phi)$,

$$u_0'' + u_0 = -\frac{\epsilon m}{L^2} \quad , \tag{25.65}$$

give the familiar conical sections for $\epsilon = -1$ (cf. section 25.8) and straight lines for $\epsilon = 0$ (section 25.11). Treating the last term as a small perturbation, one can then expand the solution u as

$$u = u_0 + u_1 \tag{25.66}$$

with u_1 small, leading to the linear equation

$$u_1'' + u_1 = 3mu_0^2 \tag{25.67}$$

for the perturbation u_1, with u_0 providing the source term (or external force on the harmonic oscillator).

Equations (25.65) and (25.67) are the equations that we will study below to determine the perihelion shift and the bending of light by a star. In the latter case, which is a bit simpler, I will also sketch two other derivations of the result, based on different perturbative evaluations of the elliptic integral.

25.6 Timelike Geodesics

We will first try to gain a qualitative understanding of the behaviour of geodesics in the effective potential

$$V_{eff}(r) = -\frac{m}{r} + \frac{L^2}{2r^2} - \frac{mL^2}{r^3} \quad . \tag{25.68}$$

The standard way to do this is to plot this potential as a function of r for various values of the parameters L and m. The basic properties of $V_{eff}(r)$ are the following:

1. Asymptotically, i.e. for $r \to \infty$, the potential tends to the Newtonian potential,

$$V_{eff}(r) \stackrel{r\to\infty}{\longrightarrow} -\frac{m}{r} \quad . \tag{25.69}$$

2. At the Schwarzschild radius $r_s = 2m$, nothing special happens and the potential is completely regular there,

$$V_{eff}(r = 2m) = -\frac{1}{2} \quad . \tag{25.70}$$

   For the discussion of planetary orbits in the solar system we can safely assume that the radius of the sun is much larger than its Schwarzschild radius, $r_0 \gg r_s$, but the above shows that even for these highly compact objects with $r_0 < r_s$ geodesics are perfectly regular as one approaches r_s. Of course the particular numerical value of $V_{eff}(r = 2m)$ has no special significance because $V(r)$ can always be shifted by a constant.

3. The extrema of the potential, i.e. the points at which $dV_{eff}/dr = 0$, are at

$$mr^2 - L^2 r + 3mL^2 = 0 \quad \Rightarrow \quad r_\pm = (L^2/2m)[1 \pm \sqrt{1 - 12(m/L)^2}] \quad , \tag{25.71}$$

   and the potential has a maximum at r_- and a local minimum at r_+. Thus there are qualitative differences in the shapes of the orbits between $L/m < \sqrt{12}$ and $L/m > \sqrt{12}$.

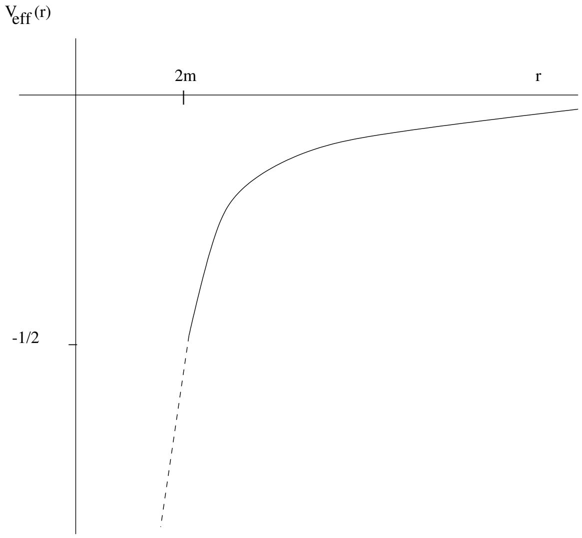


Figure 12: Effective potential for a massive particle with $L/m < \sqrt{12}$. The extrapolation to values of $r < 2m$ has been indicated by a dashed line.

Let us discuss these two cases in turn. When $L/m < \sqrt{12}$, then there are no critical points and the potential looks approximately like that in Figure 12. Note that for the time being we should be careful with extrapolating to values of r with $r < 2m$ because we know that the Schwarzschild metric has a coordinate singularity there. However, the picture (in fact the entire effective potential) turns out to be correct also for $r < 2m$.

From this picture we can read off that there are no bounded orbits for these values of the parameters. Any inward bound particle with $L < \sqrt{12}m$ will continue to fall inwards (provided that it moves on a geodesic). This should be contrasted with the Newtonian situation in which for any $L \neq 0$ there is always the centrifugal barrier reflecting incoming particles since the repulsive term $L^2/2r^2$ will dominate over the attractive $-m/r$ for small values of r. In General Relativity, on the other hand, it is the attractive term $-mL^2/r^3$ that dominates for small r.

Fortunately for the stability of the solar system, the situation is qualitatively quite different for sufficiently large values of the angular momentum, namely $L > \sqrt{12}m$ (see Figure 13).

In that case, there is a minimum and a maximum of the potential. The critical radii correspond to exactly circular orbits, unstable at r_- (on top of the potential) and stable at $r_+ > r_-$ (the minimum of the potential). We will briefly discuss these separately in section 25.7 below.

For given L, for sufficiently large values of E_{eff} a particle will fall all the way down the potential. For $E_{eff} < 0$, there are bound orbits which are not circular and which range between the radii r_1 and r_2, the turning points at which $\dot{r} = 0$ and therefore

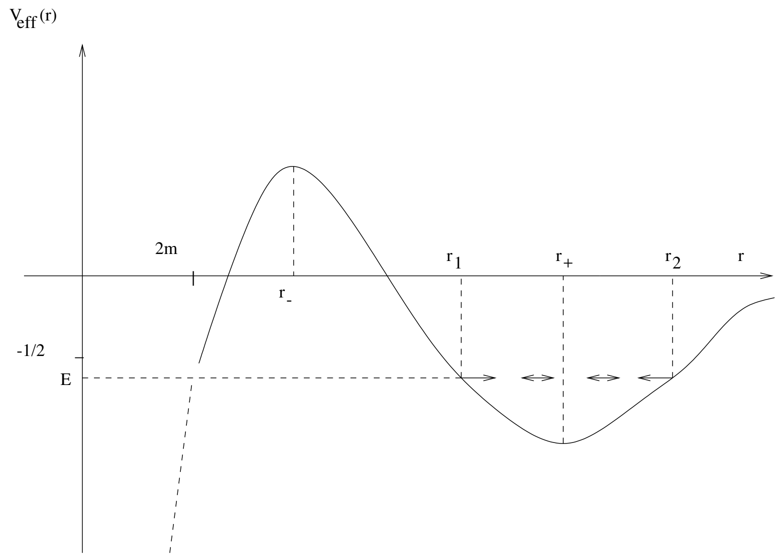


Figure 13: Effective potential for a massive particle with $L/m > \sqrt{12}$. Shown are the maximum of the potential at r_- (an unstable circular orbit), the minimum at r_+ (a stable circular orbit), and the orbit of a particle with $E_{eff} < 0$ with turning points r_1 and r_2.

$E_{eff} = V_{eff}(r_{1,2})$. We will take a closer look at these bound (but not closed) orbits, with their characteristic precession of the perihelion, in section 25.8.

25.7 Some Comments on Circular Timelike Orbits

Recall from above that, provided that $(L/m) > \sqrt{12}$, for a given L there are 2 circular orbits, at the critical points $r_\pm$ of the potential, given by (25.71)

$$r_\pm = (L^2/2m)[1 \pm \sqrt{1 - 12(m/L)^2}] \ , \tag{25.72}$$

For $L \to \sqrt{12}m$ these two circular orbits approach each other, the critical radius tending to $r_\pm \to 6m$. Thus the innermost stable circular orbit (known affectionately as the ISCO in astrophysics) is located at

$$r_{ISCO} = 3r_s = 6m \ . \tag{25.73}$$

This is of course only a relevant quantity when r_{ISCO} lies outside the star, i.e. for $r_0 < 3r_s$ (and thus mainly for black holes or possibly other quite extreme objects like extremely dense neutron stars).

On the other hand, for very large values of L the critical radii are (expand the square root to first order) to be found at

$$(r_+, r_-) \stackrel{L\to\infty}{\longrightarrow} (L^2/m, 3m) \ . \tag{25.74}$$

Thus letting L range over $[\sqrt{12}m, \infty)$, one sees that circular orbits can occur at any radius $r > 3m$ (for a suitable finetuned choice of angular momentum L). However, such an orbit is stable only for $r > 6m$.

Remarks:

1. Another way to understand the appearance and significance of the minimal circular radius at $r = 3m$ is to look directly at the radial geodesic equation (25.42)

$$\ddot{r} = -\frac{m}{r^2} + (r - 3m)\frac{L^2}{r^4} \quad . \tag{25.75}$$

   We see that at $r = 3m$ the "centrifugal" term changes sign, and that we can only compensate the first (Newtonian) term on the right-hand side by this term for $r > 3m$.

   We will see in section 25.10 that $r = 3m$ also has a special (and more prominent) significance for *null* geodesics, namely as the location of (unstable) circular null geodesics (the *photon sphere*).

2. From (25.20) one finds that the angular coordinate velocity on a circular orbit $r_\pm$ is

$$\omega_\pm = d\phi/dt = \dot{\phi}/\dot{t} = (L/E)(1 - 2m/r_\pm)/r_\pm^2 \quad . \tag{25.76}$$

   Since this frequency is constant, this immediately gives $\phi = \phi(t)$,

$$\phi(t) = \omega_\pm t + \phi_0 \quad . \tag{25.77}$$

   However, since the quantities $m, E, L, r_\pm$ are not independent of each other, this is not a particularly useful way of writing $\omega_\pm$. First of all, from the effective potential equation and $\dot{r}_\pm = 0$ one finds that the energy E can be written in terms of the other parameters as

$$E^2 = 1 + 2V_{eff}(r_\pm) = (1 - 2m/r_\pm)(1 + L^2/r_\pm^2) \quad . \tag{25.78}$$

   Similarly, from (25.71) L^2 can be written in terms of m and $r_\pm$ as

$$L^2 = mr_\pm^2/(r_\pm - 3m) \quad . \tag{25.79}$$

   Thus one finds that

$$\frac{L^2}{E^2} = \frac{mr_\pm^3}{(r_\pm - 2m)^2} \quad , \tag{25.80}$$

   and the angular velocity (as a function of m and $r_\pm$) is

$$(\omega_\pm)^2 = \frac{m}{r_\pm^3} \quad \Leftrightarrow \quad \omega_\pm^2 r_\pm^3 = m \quad . \tag{25.81}$$

   In particular, for the ISCO one has the characteristic orbital frequency

$$m\omega_{ISCO} = 6^{-3/2} \quad . \tag{25.82}$$

Since ω is inversely proportional to the period of the motion, (25.81) looks very much like Kepler's law which says that the period-squared is proportional to the radius-cubed of an orbit. However, this is a bit of a fluke because this relation is not coordinate-independent but relies on having expressed time and radius in terms of the Schwarzschild coordinates (rather than in terms of proper time along the orbit, say).

25.8 Anomalous Precession of the Perihelia of the Planetary Orbits

Because of the general relativistic correction $\sim 1/r^3$, the bound orbits will not be closed (elliptical). In particular, the position of the perihelion, the point of closest approach of the planet to the sun where the planet has distance r_1, will not remain constant. However, because r_1 is constant, and the planetary orbit is planar, this point will move on a circle of radius r_1 around the sun.

As described in section 25.5, in order to calculate this perihelion shift one needs to calculate the total angle $\Delta\phi$ swept out by the planet during one revolution by integrating this from r_1 to r_2 and back again to r_1, or

$$\Delta\phi = 2\int_{r_1}^{r_2} \frac{d\phi}{dr} dr \quad . \tag{25.83}$$

Rather than trying to evaluate the above integral via some sorcery, we will determine $\Delta\phi$ by analysing the orbit equation (25.64) for $\epsilon = -1$,

$$u'' + u = \frac{m}{L^2} + 3mu^2 \quad . \tag{25.84}$$

In the Newtonian approximation, this equation reduces to that of a displaced harmonic oscillator,

$$u_0'' + u_0 = \frac{m}{L^2} \quad \Leftrightarrow \quad (u_0 - m/L^2)'' + (u_0 - m/L^2) = 0 \quad , \tag{25.85}$$

and the solution is a Kepler ellipse described parametrically by

$$u_0(\phi) = \frac{m}{L^2}(1 + e\cos\phi) \tag{25.86}$$

where e is the eccentricity ($e = 0$ means constant radius and hence a circular orbit). Plugging this back into the Newtonian non-linear 1st-order equation (cf. (25.61))

$${u_0'}^2 + u_0^2 = \frac{E^2 - 1}{L^2} + \frac{2m}{L^2}u_0 \quad , \tag{25.87}$$

one finds that the integration constant e is related to the energy by

$$e^2 = 1 + \frac{L^2}{m^2}(E^2 - 1) = 1 + \frac{2L^2}{m^2}E_{eff} \quad . \tag{25.88}$$

In particular, $e^2 < 1$ for bound states (bounded orbits) with $E_{eff} < 0$, and we will concentrate on these orbits. The perihelion (aphelion) is then at $\phi = 0$ ($\phi = \pi$), with

$$r_{1,2} = \frac{L^2}{m}\frac{1}{1 \pm e} \quad . \tag{25.89}$$

Thus the semi-major axis a of the ellipse,

$$2a = r_1 + r_2 \quad , \tag{25.90}$$

is

$$a = \frac{L^2}{m}\frac{1}{1-e^2} \quad . \tag{25.91}$$

In particular, in the Newtonian theory, one has

$$(\Delta\phi)_0 = 2\int_0^\pi d\phi = 2\pi \quad . \tag{25.92}$$

The anomalous perihelion shift due to the effects of General Relativity is thus

$$\delta\phi = \Delta\phi - 2\pi \quad . \tag{25.93}$$

In order to determine $\delta\phi$, we now seek a solution to (25.84) of the form

$$u = u_0 + u_1 \tag{25.94}$$

where u_1 is a small deviation. This leads to the equation

$$u_1'' + u_1 = 3mu_0^2 \quad . \tag{25.95}$$

The general solution of this inhomogenous differential equation is the general solution of the homogeneous equation (we are not interested in) plus a special solution of the inhomogeneous equation. Noting that this is a forced harmonic oscillator equation, with the frequency of the force commensurate with the frequency of the oscillator, we expect to encounter a resonance phenomenon leading to a non-periodic (and linearly growing) contribution to the solution (just as a special solution of $\ddot{x}(t) + x(t) = \sin t$ is $x(t) = (t/2)\sin t$).

This expectation is indeed borne out. Writing

$$(1 + e\cos\phi)^2 = (1 + \tfrac{1}{2}e^2) + 2e\cos\phi + \tfrac{1}{2}e^2\cos 2\phi \tag{25.96}$$

and noting that

$$\begin{aligned}(\phi\sin\phi)'' + \phi\sin\phi &= 2\cos\phi \\ (\cos 2\phi)'' + \cos 2\phi &= -3\cos 2\phi\end{aligned} \tag{25.97}$$

one sees that a special solution is

$$u_1(\phi) = \frac{3m^3}{L^4}((1 + \tfrac{1}{2}e^2) - \tfrac{1}{6}e^2\cos 2\phi + e\phi\sin\phi) \quad . \tag{25.98}$$

The term of interest to us is the third (non-periodic, resonance) term which provides a cumulative non-periodic effect over successive orbits. Focussing on this term, we can write the approximate solution to the orbit equation as

$$u(\phi) \approx \frac{m}{L^2}(1 + e\cos\phi + \frac{3m^2 e}{L^2}\phi\sin\phi) \quad . \tag{25.99}$$

If the first perihelion is at $\phi = 0$, the next one will be at a point $\Delta\phi = 2\pi + \delta\phi$ close to 2π which is such that $u'(\Delta\phi) = 0$ or

$$\sin\delta\phi = \frac{3m^2}{L^2}(\sin\delta\phi + (2\pi + \delta\phi)\cos\delta\phi) \quad . \tag{25.100}$$

Using that $\delta\phi$ is small, and keeping only the lowest order terms in this equation, one finds the result

$$\delta\phi = \frac{6\pi m^2}{L^2} = 6\pi(\frac{G_N M}{cL})^2 \quad . \tag{25.101}$$

An alternative way to obtain this result is to observe that (25.99) can be approximately written as

$$u(\phi) \approx \frac{m}{L^2}\left(1 + e\cos[(1 - \frac{3m^2}{L^2})\phi]\right) \tag{25.102}$$

From this equation it is manifest that during each orbit the perihelion advances by

$$\delta\phi = 2\pi\frac{3m^2}{L^2} \tag{25.103}$$

$(2\pi(1 - 3m^2/L^2)(1 + 3m^2/L^2) \approx 2\pi)$ in agreement with the above result. In terms of the eccentricity e and the semi-major axis a (25.91) of the elliptical orbit, this can be written as

$$\delta\phi = \frac{6\pi G_N}{c^2}\frac{M}{a(1-e^2)} \quad . \tag{25.104}$$

Thus general relativity predicts a deviation from the Kepler orbits of the planets, manifesting itself in a precession of the perihelia. As the parameters entering (25.104) are known for the planetary orbits, $\delta\phi$ can be evaluated. Can this also be observed and tested?[78]

At first sight, this seems difficult. Even for Mercury, where this effect is largest (because it has the largest eccentricity) one only finds a $\delta\phi$ of the order of $0,1''$ per revolution. This is of course a tiny effect (1 second, $1''$, is one degree divided by 3600) and not *per se* detectable. However,

1. this effect is cumulative, i.e. after N revolutions one has an anomalous perihelion shift $N\delta\phi$;

2. Mercury has a very short solar year, with about 415 revolutions per century;

3. and accurate observations of the orbit of Mercury go back over more than 200 years.

[78] Since I have no independent way of checking them, the following facts and numbers are taken from S. Weinberg, *Gravitation and Cosmology*, section 8.6. See also B. Berche, E. Medina, *The advance of Mercury's perihelion*, `arXiv:2402.04643 [gr-qc]` for a recent pedagogical discussion and more details.

Thus the cumulative effect is approximately $10^3\ \delta\phi$ and this *is* sufficiently large to be observable in principle. And indeed such an effect is observed (and had for a long time presented a puzzle, an anomaly, for astronomers).

In actual fact, the perihelion of Mercury's orbit, as observed from the Earth, shows a rather significant precession rate of $5601''$ per century, which seems to flatly contradict the Newtonian result that there should be no such precession at all.

However,

- about $5025''$ are due to fact that one is using a non-inertial geocentric coordinate system (precession of the equinoxes)
- $532''$ are due to perturbations of Mercury's orbit caused by the (Newtonian) gravitational attraction of the other planets of the solar system (chiefly Venus, earth and Jupiter).

This much was known prior to General Relativity and left an unexplained anomalous perihelion shift of

$$\delta\phi_{\text{anomalous}} = 43,11'' \pm 0,45''/\text{century} \quad . \tag{25.105}$$

Various ad hoc explanations for this tiny discrepancy between theory and observation (like postulating an additional planet, called *Vulcan*, on an orbit closer to the sun than that of Mercury) were proposed and dismissed.

From (25.104) the prediction of General Relativity for the precession of the perihelion due the the general relativistic correction to the Newtonian theory can be calculated to be

$$\delta\phi_{\text{GR}} = 43,03''/\text{century} \quad . \tag{25.106}$$

Thus general relativity, with its general relativistic correction to the (unperturbed) Kepler orbit of Mercury, appears to account precisely for the observed "anomaly". This is quite a striking, remarkable and impressive confirmation of a prediction of general relativity.

Other observations, involving e.g. the mini-planet Icarus, discovered in 1949, with a huge eccentricity $e \sim 0,827$, binary pulsar systems, and more recently obervations of highly eccentric stars close to the galactic (black hole) center have provided further confirmation of the agreement between General Relativity and observations.

25.9 Null Geodesics

To study the behaviour of massless particles (photons) in the Schwarzschild geometry, we need to study the effective potential

$$V_{eff}(r) = \frac{L^2}{2r^2} - \frac{mL^2}{r^3} = \frac{L^2}{2r^2}(1 - \frac{2m}{r}) \quad . \tag{25.107}$$

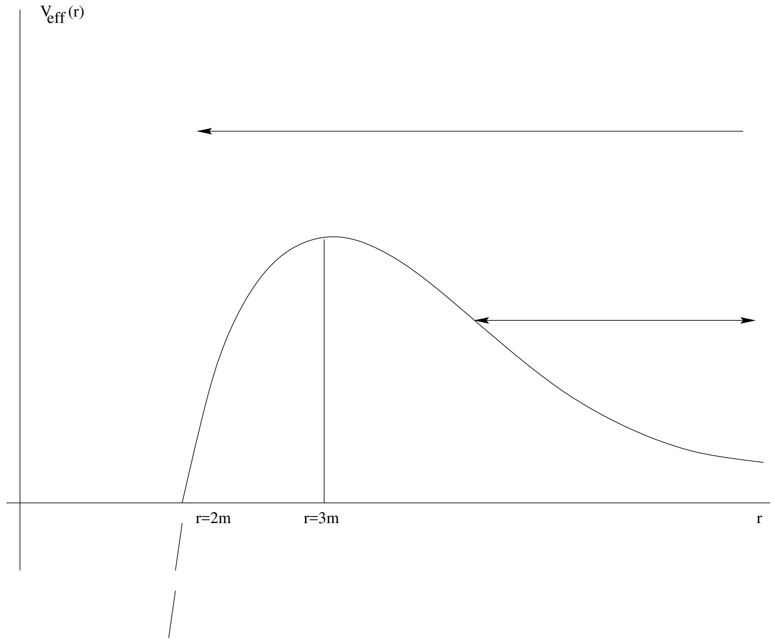


Figure 14: Effective potential for a massless particle. Displayed is the location of the unstable circular orbit at $r = 3m$. A photon with an energy $E^2 < L^2/27m^2$ will be deflected (lower arrow), photons with $E^2 > L^2/27m^2$ will be captured by the star.

The following properties are immediate:

1. For $r > 2m$, the potential is positive, $V(r) > 0$.
2. For $r \to \infty$, one has $V_{eff}(r) \to 0$.
3. $V_{eff}(r = 2m) = 0$.
4. When $L = 0$, the photons feel no potential at all.
5. There is one critical point of the potential, at $r = 3m$, with $V_{eff}(r = 3m) = L^2/54m^2$.

Thus the potential has the form sketched in Figure 14, with the following consequences:

1. For energies $E^2 > L^2/27m^2$, photons are captured by the star and will spiral into it. For energies $E^2 < L^2/27m^2$, on the other hand, there will be a turning point, and lightrays will be deflected by the star.

   As this may sound a bit counterintuitive (shouldn't a photon with higher energy be more likely to zoom by the star without being forced to spiral into it?), think about this in the following way. $L = 0$ corresponds to a photon falling radially towards the star, L small corresponds to a slight deviation from radial motion, while L large (thus $\dot{\phi}$ large) means that the photon is travelling along a trajectory that will not bring it very close to the star at all (see the next subsection for the precise relation between the angular momentum L and the impact parameter b of the photon). It is then not surprising that photons with small L are more likely to be captured by the star (this happens for $L^2 < 27m^2E^2$) than photons with large L which will only be deflected in their path. We will study this in more detail below.

2. Now let us also consider the opposite situation, that of light from or near the star (and we are of course assuming that $r_0 > r_s$). Then for $r_0 < 3m$ and $E^2 < L^2/27m^2$, the light cannot escape to infinity but falls back to the star, whereas for $E^2 > L^2/27m^2$ light will escape. Thus for a path sufficiently close to radial (L small, because $\dot{\phi}$ is then small) light can always escape as long as $r_0 > 2m$.

3. Finally, there is one critical point of the potential, at $r = 3m$. We will (briefly) discuss this case separately below.

25.10 Some Comments on Circular Null Orbits (The Photon Sphere)

As noted above, the effective potential has one critical point, a maximum, at $r = 3m$. This can also be seen from the equation of motion (25.44)

$$\ddot{r} = \frac{L^2}{r^4}(r - 3m) \quad . \tag{25.108}$$

Thus there exists one unstable circular orbit for photons at $r = 3m$ (more precisely one in each "equatorial plane").

While these orbits have some properties in common with the circular timelike orbits discussed in section 25.7, there are also some interesting differences. The first and crucial difference is that this circular orbit can only arise at one particular value of r, namely at $r = 3m$, while (as we have seen) circular timelike orbits can exist for any $r > 3m$. Here are some more brief comments and observations.

Remarks:

1. While not relevant for the applications to the solar system in this section, where we are dealing with objects with a size much larger than $3m$, the existence of this sphere of unstable photon orbits, the so-called *photon sphere*, turns out to be of some interest in black hole astrophysics (as a possibly observable signature of black holes or other very compact objects).

2. Since

$$r = 3m \quad \Rightarrow \quad \dot{t} = E/f(r=3m) = 3E \quad , \quad \dot{\phi} = L/(3m)^2 = L/(9m^2) \ , \tag{25.109}$$

these circular null geodesics are characterised by the null wave vector

$$(k^\alpha) = (\dot{t}, \dot{r}, \dot{\phi}, \dot{\theta}) = (3E, 0, L/9m^2, 0) \ , \tag{25.110}$$

with E and L related either by the null condition or, of course equivalently, by the effective potential equation,

$$\dot{r}|_{r=3m} = 0 \quad \Rightarrow \quad E^2 = 2V_{eff}(r=3m) = L^2/(27m^2) \ . \tag{25.111}$$

Thus the wave vector can e.g. be written as

$$(k^\alpha) = (L/\sqrt{3}m, 0, L/9m^2, 0) \ . \tag{25.112}$$

Of course, the overall normalisation is irrelevant (as it depends on the choice of affine parameter).

3. In this case, the angular velocity (frequency) is given by

$$\omega = d\phi/dt = \dot{\phi}/\dot{t} = (L/E)(1-2m/r)/r^2|_{r=3m} = L/(27Em^2) \ . \tag{25.113}$$

Using the relation (25.111), this can be written as

$$\omega = E/L = 1/(3\sqrt{3}m) \quad \Leftrightarrow \quad m\omega = 3^{-3/2} \tag{25.114}$$

This agrees with the $r_- \to 3m$ ($L \to \infty$) limit of the frequency ω_- of the unstable circular timelike orbit (25.81) at $r = r_-$,

$$\lim_{r_- \to 3m} (m\omega_-) = 3^{-3/2} \tag{25.115}$$

(and is structurally similar to, but different from, the ISCO frequency $m\omega_{ISCO} = 6^{-3/2}$ (25.82)).

25.11 Bending of Light by a Star: 3 Derivations

To study the bending of light by a star, we consider an incoming photon (or lightray) with impact parameter b (see Figure 15) and we need to calculate $\phi(r)$ for a trajectory with turning point at $r = r_1$. At that point we have $\dot{r} = 0$. Here the dot can, as usual, be taken to be the derivative with repect to some affine parameter σ. However, noting that the condition $g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 0$ is reparametrisation-invariant (unlike its massive cousin $g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = -1$), we can equally well choose to parametrise the lightrays by the coordinate time t even though this is not an affine parameter (this matters at the level of the 2nd order geodesic differential equations but not at the level of the 1st integrals and the effective potential).

Either way r_1 is determined by

$$E_{eff} = V_{eff}(r_1) \quad \Leftrightarrow \quad r_1^2 = \frac{L^2}{E^2}(1 - \frac{2m}{r_1}) \ . \tag{25.116}$$

The first thing we need to establish is the relation between b and the other parameters E and L. Consider the ratio

$$\frac{L}{E} = \frac{r^2\dot{\phi}}{(1-2m/r)\dot{t}} \ . \tag{25.117}$$

For large values of r, $r \gg 2m$, this reduces to

$$\frac{L}{E} = r^2\frac{d\phi}{dt} \ . \tag{25.118}$$

On the other hand, for large r we can approximate $b/r = \sin\phi$ by ϕ. Since we also have $dr/dt = -1$ (for an incoming lightray), we deduce

$$\frac{L}{E} = r^2\frac{d}{dt}\frac{b}{r} = b \ . \tag{25.119}$$

In terms of the variable $u = 1/r$ the equation for the shape of the orbit (25.64) is

$$u'' + u = 3mu^2 \tag{25.120}$$

and the elliptic integral (25.62) for $\Delta\phi$ is

$$\Delta\phi = 2\int_{r_1}^{\infty}\frac{d\phi}{dr}dr = 2\int_0^{u_1} du\, [b^{-2} - u^2 + 2mu^3]^{-1/2} \ . \tag{25.121}$$

Moreover, in terms of u we can write the equation (25.116) for $u_1 = 1/r_1$ as

$$b^{-2} = u_1^2 - 2mu_1^3 \ . \tag{25.122}$$

In the absence of the general relativistic correction (calling this 'Newtonian' is perhaps not really appropriate since we are dealing with photons/lightrays) one has $b^{-1} = u_1$ or $b = r_1$ (no deflection). The orbit equation

$$u_0'' + u_0 = 0 \tag{25.123}$$

has the solution

$$u_0(\phi) = \frac{1}{b}\sin\phi \ , \tag{25.124}$$

describing the straight line

$$r_0(\phi)\sin\phi = b \ . \tag{25.125}$$

Obligingly the integral gives

$$(\Delta\phi)_0 = 2\int_0^1 dx(1-x^2)^{-1/2} = 2\arcsin 1 = \pi \ . \tag{25.126}$$

Thus the deflection angle is related to $\Delta\phi$ by

$$\delta\phi = \Delta\phi - \pi \quad . \tag{25.127}$$

We will now determine $\delta\phi$ in three different ways,

- by perturbatively solving the orbit equation (25.120);
- by perturbatively evaluating the elliptic integral (25.121);
- by performing a perturbative expansion (linearisation) of the Schwarzschild metric.

Derivation I: Perturbative Solution of the Orbit Equation

In order to solve the orbit equation (25.120), we proceed as in section 25.8. Thus the equation for the (small) deviation $u_1(\phi)$ is

$$u_1'' + u_1 = 3mu_0^2 = \frac{3m}{b^2}(1 - \cos^2\phi) = \frac{3m}{2b^2}(1 - \cos 2\phi) \tag{25.128}$$

which has the particular solution (cf. (25.97))

$$u_1(\phi) = \frac{3m}{2b^2}(1 + \tfrac{1}{3}\cos 2\phi) \quad . \tag{25.129}$$

Therefore

$$u(\phi) = \frac{1}{b}\sin\phi + \frac{3m}{2b^2}(1 + \tfrac{1}{3}\cos 2\phi) \quad . \tag{25.130}$$

By considering the behaviour of this equation as $r \to \infty$ or $u \to 0$, one finds an equation for (minus) half the deflection angle, namely

$$\frac{1}{b}(-\delta\phi/2) + \frac{3m}{2b^2}\frac{4}{3} = 0 \quad , \tag{25.131}$$

leading to the result

$$\delta\phi = \frac{4m}{b} = \frac{4MG_N}{bc^2} \quad . \tag{25.132}$$

Derivation II: Perturbative Evaluation of the Elliptic Integral

The perturbative evaluation of (25.121) is rather tricky when it is regarded as a function of the independent variables m and b, with r_1 determined by (25.116) (try this!). The trick to evaluate (25.121) is (see R. Wald, *General Relativity*) to regard the integral as a function of the independent variables r_1 and m, with b eliminated via (25.122). Thus (25.121) becomes

$$\Delta\phi = 2\int_0^{u_1} du\, [u_1^2 - u^2 - 2m(u_1^3 - u^3)]^{-1/2} \quad . \tag{25.133}$$

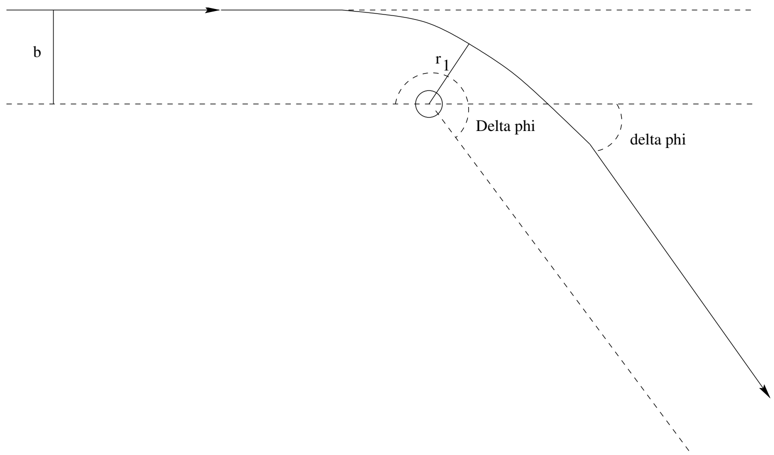


Figure 15: Bending of light by a star. Indicated are the definitions of the impact parameter b, the perihelion r_1, and of the angles $\Delta\phi$ and $\delta\phi$.

The first order correction

$$\Delta\phi = (\Delta\phi)_0 + m(\Delta\phi)_1 + \mathcal{O}(m^2) \tag{25.134}$$

is therefore

$$(\Delta\phi)_1 = \left(\frac{\partial}{\partial m}\Delta\phi\right)_{m=0} = 2\int_0^{b^{-1}} du\, \frac{b^{-3}-u^3}{(b^{-2}-u^2)^{3/2}} \quad . \tag{25.135}$$

This integral is elementary,

$$\int dx\, \frac{1-x^3}{(1-x^2)^{3/2}} = -(x+2)\left(\frac{1-x}{1+x}\right)^{1/2} \quad , \tag{25.136}$$

and thus

$$(\Delta\phi)_1 = 4b^{-1} \quad , \tag{25.137}$$

leading to

$$\delta\phi = \frac{4m}{b} = \frac{4MG_N}{bc^2} \quad , \tag{25.138}$$

in agreement with the result (25.132) obtained above.

Derivation III: Linearising the Schwarzschild Metric

It is instructive to look at the second derivation from another point of view. As we will see, in some sense this derivation 'works' because the bending of light is accurately described by the *linearised solution*, i.e. by the metric that one obtains from the Schwarzschild metric by the approximation

$$\begin{aligned} A(r) &= 1-\frac{2m}{r} \quad \to \quad 1-\frac{2m}{r} \\ B(r) &= (1-\frac{2m}{r})^{-1} \quad \to \quad 1+\frac{2m}{r} \quad . \end{aligned} \tag{25.139}$$

I will only sketch the main steps in this calculation, so you should think of this subsection as an annotated exercise.

First of all, redoing the analysis of sections 25.3 and 25.5 for a general spherically symmetric static metric (24.6),

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad , \tag{25.140}$$

it is easy to see that the orbit equation can be written as

$$B(r)\frac{r'^2}{r^4} + \frac{1}{r^2} = \frac{\epsilon}{L^2} + \frac{E^2}{L^2 A(r)} \tag{25.141}$$

or, in terms of $u = 1/r$, as (abusing notation by writing $A(r = 1/u)$ as $A(u)$ etc., apologies if this causes allergic reactions)

$$B(u)\; u'^2 + u^2 = \frac{\epsilon}{L^2} + \frac{E^2}{L^2 A(u)} \quad . \tag{25.142}$$

We will concentrate on the lightlike case $\epsilon = 0$,

$$B(u)\; u'^2 + u^2 = \frac{E^2}{L^2 A(u)} \quad , \tag{25.143}$$

and express the impact parameter $b = L/E$ in terms of the turning point $r_1 = 1/u_1$ of the trajectory. At this turning point, $u' = 0$, and thus

$$\frac{E^2}{L^2} = A(u_1)u_1^2 \quad , \tag{25.144}$$

leading to

$$B(u)\; u'^2 = \frac{A(u_1)}{A(u)}u_1^2 - u^2 \quad . \tag{25.145}$$

We thus find

$$\frac{d\phi}{du} = \pm B(u)^{1/2}\left[\frac{A(u_1)}{A(u)}u_1^2 - u^2\right]^{-1/2} \quad . \tag{25.146}$$

For the (linearised) Schwarzschild metric the term in square brackets is

$$\begin{aligned} \frac{A(u_1)}{A(u)}u_1^2 - u^2 &= u_1^2(1 + 2m(u - u_1)) - u^2 \\ &= (u_1^2 - u^2)(1 - 2m\frac{u_1^2}{u_1 + u}) \quad . \end{aligned} \tag{25.147}$$

Using this and the approximate (linearised) value for $B(u)$,

$$B(u) \approx 1 + 2mu \tag{25.148}$$

one finds that $d\phi/du$ is given by

$$\begin{aligned} \frac{d\phi}{du} &= \pm B(u)^{1/2}\left[(u_1^2 - u^2)(1 - 2m\frac{u_1^2}{u_1 + u})\right]^{-1/2} \\ &\approx (u_1^2 - u^2)^{-1/2}\left(1 + m(\frac{u_1^2}{u_1 + u} + u)\right) \\ &= (u_1^2 - u^2)^{-1/2} + m\frac{u_1^3 - u^3}{(u_1^2 - u^2)^{3/2}} \quad . \end{aligned} \tag{25.149}$$

The first term now gives us the Newtonian result and, comparing with Derivation II, we see that the second term agrees precisely with the integrand of (25.135) with $b \to r_1$ (which, in a term that is already of order m, makes no difference). We thus conclude that the deflection angle is, as before,

$$\delta\phi = 2\int_0^{u_1} du\, m\frac{u_1^3 - u^3}{(u_1^2 - u^2)^{3/2}} = 4mu_1 \approx \frac{4m}{b} \quad . \tag{25.150}$$

<u>Remarks:</u>

1. This effect is physically measurable and was one of the first true tests of Einstein's new theory of gravity. For light just passing the sun the predicted value is

   $$\delta\phi \sim 1,75'' \quad . \tag{25.151}$$

   Quick sanity check on this estimate: $1'' = \pi/(180 \times 3600) \approx 0.5 \times 10^{-5}$ radians, while using the (rough) values for $r_s = 2m$ and $r_0 \approx b$ for the sun given in (24.40), one has

   $$\frac{4m}{b} \approx (6/7) \times 10^{-5} \approx (12/7)'' \approx 1.7'' \quad . \tag{25.152}$$

   Experimentally this is a bit tricky to observe because one needs to look at light from distant stars passing close to the sun. Under ordinary circumstances this would not be observable, but in 1919 a test of this was performed during a total solar eclipse, by observing the effect of the sun on the apparent position of stars in the direction of the sun. The observed value was rather imprecise, yielding $1,5'' < \delta\phi < 2,2''$ which is, if not a confirmation of, at least consistent with General Relativity.

2. More recently, it has also been possible to measure the deflection of radio waves by the gravitational field of the sun. These measurements rely on the fact that a particular Quasar, known as 3C275, is obscured annually by the sun on October 8th, and the observed result (after correcting for diffraction effects by the corona of the sun) in this case is $\delta = 1,76'' \pm 0,02''$.

3. The value predicted by General Relativity is, interestingly enough, exactly twice the value that would have been predicted by the Newtonian approximation of the geodesic equation alone (but the Newtonian approximation is not valid anyway because it applies to slowly moving objects, and light certainly fails to satisfy this condition). A calculation leading to this wrong value had first been performed by Soldner in 1801 (!) (by cancelling the mass m out of the Newtonian equations of motion before setting $m = 0$) and also Einstein predicted this wrong result in 1908 (his equivalence principle days, long before he came close to discovering the field equations of General Relativity now carrying his name).

This result can be obtained from the above calculation by setting $B(u) = 1$ instead of (25.148), as in the Newtonian approximation only g_{00} is non-trivial.

4. More generally, one can calculate the deflection angle for a metric with the approximate behaviour
$$B(u) \approx 1 + 2\gamma m u \quad , \tag{25.153}$$
for γ a real parameter, with the result
$$\delta\phi \approx \frac{1+\gamma}{2}\frac{4m}{b} \quad . \tag{25.154}$$
This reproduces the previous result for $\gamma = 1$, half its value for $\gamma = 0$, and checking to which extent measured deflection angles agree with the theoretical prediction of general relativity ($\gamma = 1$) constitutes an experimental test of general relativity. In this context γ is known as one of the PPN parameters (PPN for *parametrised post-Newtonian approximation*).

25.12 A Unified Description in terms of the Runge-Lenz Vector

The perhaps slickest way to obtain the orbits of the Kepler problem is to make use of the so-called *Runge-Lenz vector* (even though it was discovered neither by Runge nor by Lenz).[79]

Recall that, due to conservation of angular momentum $\vec{L}$, the orbits in any spherically symmetric potential are planar. The bound orbits of the Kepler problem, however, have the additional property that they are closed, i.e. that the perihelion is constant. This suggests that there is a further hidden symmetry in the Kepler problem, with the position of the perihelion the corresponding conserved charge. This is indeed the case.

Consider, for a spherically symmetric potential $W(r)$, the vector
$$\vec{A} = \dot{\vec{x}} \times \vec{L} + W(r)\vec{x} \tag{25.155}$$
or, in components,
$$A_i = \epsilon_{ijk}\dot{x}_j L_k + W(r)x_i \quad . \tag{25.156}$$
A straightforward calculation, using the Newtonian equations of motion in the potential $W(r)$, shows that
$$\frac{d}{dt}A_i = (r\partial_r W(r) + W(r))\dot{x}_i \quad . \tag{25.157}$$
Thus $\vec{A}$ is conserved if and only if $W(r)$ is homogeneous of degree (-1),
$$\frac{d}{dt}\vec{A} = 0 \quad \Leftrightarrow \quad W(r) = \frac{c}{r} \quad . \tag{25.158}$$

[79] For an account of the history, with further references, see e.g. the Wikipedia entry `https://en.wikipedia.org/wiki/Laplace-Runge-Lenz_vector`.

In our notation, $c = \epsilon m$, and we will henceforth refer to the vector

$$\vec{A} = \dot{\vec{x}} \times \vec{L} + \frac{\epsilon m}{r}\vec{x} \tag{25.159}$$

as the Runge-Lenz vector.

In addition to being preserved for a $(1/r)$-potential, the Runge-Lenz vector has the following properties:

1. $\vec{A}$ is orthogonal to $\vec{L}$ and hence lies in the plane of the orbit,

$$\vec{x}.\vec{L} = 0 \quad \Rightarrow \quad \vec{A}.\vec{L} = 0 \ . \tag{25.160}$$

2. The norm A of $\vec{A}$ can be expressed in terms of the other conserved quantities and parameters (energy E, angular momentum L, mass m) of the problem. In the notation of section 25.3 one has

$$A^2 = E^2L^2 + \epsilon(L^2 + \epsilon m^2) \ . \tag{25.161}$$

3. Thus, even though a priori $\vec{A}$ has 3 components, the only new information is contained in the constant direction of $\vec{A}$, which (since it lies in the orbital plane) is just one angle, a single real number. Thus all in all, in the Kepler problem one has 5 independent constants of motion, $E, \vec{L}$ and the direction of $\vec{A}$.

4. It is well known, and can be shown e.g. by determining the Poisson brackets among the L_i and A_j (the calculation of $\{A_i, A_j\}$ is a bit messy), and a suitable rescaling of the A_i by a function of the (conserved) energy, that $\vec{A}$ extends the manifest symmetry group of rotations $SO(3)$ of the Kepler problem to the (hidden, phase space) symmetry group $SO(4)$ for bound orbits and $SO(3,1)$ for scattering orbits. We will not need to make use of this fact here, though.

Given all this information, it is now straightforward to determine the Keplerian orbits from $\vec{A}$. Let us choose the constant direction of $\vec{A}$ to be in the direction $\phi = 0$. Then $\vec{A}.\vec{x} = Ar\cos\phi$ and from (25.159) one finds

$$Ar\cos\phi = L^2 + \epsilon mr \ . \tag{25.162}$$

Now we consider the two cases $\epsilon = -1$ and $\epsilon = 0$.

For $\epsilon = -1$, (25.162) can be written as

$$\frac{1}{r(\phi)} = \frac{m}{L^2}(1 + \frac{A}{m}\cos\phi) \ . \tag{25.163}$$

Comparing with (25.86), we recognise this as the equation for an ellipse with eccentricity e and semi-major axis a (25.91) given by

$$e = \frac{A}{m} \qquad \frac{m}{L^2} = \frac{1}{a(1-e^2)} \ . \tag{25.164}$$

Moreover, we see that the perihelion is at $\phi = 0$ which establishes that the Runge-Lenz vector points from the center of attraction to the (constant) position of the perihelion. During one revolution the angle ϕ changes from 0 to 2π.

For $\epsilon = 0$ (i.e. no potential), on the other hand, (25.162) reduces to

$$\frac{1}{r(\phi)} = \frac{A}{L^2}\cos\phi \tag{25.165}$$

This describes a straight line (25.124) with impact parameter

$$b = \frac{L^2}{A} = \frac{L}{E} \quad . \tag{25.166}$$

In this case, ϕ runs from $-\pi/2$ to $\pi/2$ and the point of closest approach is again at $\phi = 0$ (distance b).

We see that the Runge-Lenz vector captures precisely the information that in the Newtonian theory bound orbits are closed and lightrays are not deflected. The Runge-Lenz vector will no longer be conserved in the presence of the general relativistic correction to the Newtonian motion, and this non-constancy is a precise measure of the deviation of the general relativistic orbits from their Newtonian counterparts. As shown e.g. in an article by Brill and Goel[80] this provides a very elegant and quick way of (re-)deriving the results about perihelion precession and deflection of light in the solar system.

Calculating the time-derivative of the Newtonian Runge-Lenz $\vec{A}$ (25.159), but now for a particle moving in the general relativistic potential (25.34)

$$V(r) = \epsilon\frac{m}{r} - \frac{mL^2}{r^3} \quad , \tag{25.167}$$

one finds one additional term arising from substituting the equation of motion into $\ddot{x}_j$, leading to (of course we now switch from t to τ)

$$\frac{d}{d\tau}\vec{A} = \frac{3mL^2}{r^2}\frac{d}{d\tau}\vec{n} \quad , \tag{25.168}$$

where $\vec{n} = \vec{x}/r = (\cos\phi, \sin\phi, 0)$ is the unit vector in the equatorial plane $\theta = \pi/2$ of the orbit. Thus $\vec{A}$ rotates in the ϕ-direction in the equatorial plane. If $\vec{A}$ is originally pointing in the x^1-direction $\phi = 0$, then its initial angular velocity in the x^2-direction is

$$\omega = \frac{3mL^2\cos\phi}{Ar^2}\dot{\phi} \quad . \tag{25.169}$$

In principle, here A refers to the norm of the Newtonian Runge-Lenz vector (25.159) calculated for a trajectory $\vec{x}(\tau)$ in the general relativistic potential (25.167). This norm is now no longer constant,

$$A^2 = E^2L^2 + \epsilon(L^2 + \epsilon m^2) + \frac{2mL^4}{r^3} \quad . \tag{25.170}$$

[80] D. Brill, D. Goel, *Light bending and perihelion precession: A unified approach*, Am. J. Phys. 67 (1999) 316, `arXiv:gr-qc/9712082`

However, assuming that the change in $\vec{A}$ is small, we obtain an approximate expression for ω by substituting the unperturbed orbit $r_0(\phi)$ from (25.162),

$$r_0(\phi) = \frac{L^2}{A\cos\phi - \epsilon m} \quad , \tag{25.171}$$

as well as the unperturbed norm (25.161) in (25.169) to find

$$\omega \approx \frac{3m}{AL^2}(A\cos\phi - \epsilon m)^2 \cos\phi\ \dot{\phi} \quad . \tag{25.172}$$

Now the total change in the direction of $\vec{A}$ when the object moves from ϕ_1 to ϕ_2 can be calculated from

$$\begin{aligned} \delta\phi &= \int_{\phi_1}^{\phi_2} \omega d\tau \\ &= \frac{3m}{AL^2}\int_{\phi_1}^{\phi_2} d\phi\ (A\cos\phi - \epsilon m)^2 \cos\phi \quad . \end{aligned} \tag{25.173}$$

For $\epsilon = -1$, and $(\phi_1, \phi_2) = (0, 2\pi)$, this results in (only the $\cos^2\phi$-term gives a non-zero contribution)

$$\delta\phi = 2\pi\frac{3m^2}{L^2} = \frac{6\pi m^2}{L^2} \quad , \tag{25.174}$$

in precise agreement with (25.101,25.103).

For $\epsilon = 0$, on the other hand, one has

$$\delta\phi = \frac{3mA}{L^2}\int_{-\pi/2}^{\pi/2} d\phi\ \cos^3\phi \quad . \tag{25.175}$$

Using

$$\int \cos^3\phi = \sin\phi - \tfrac{1}{3}\sin^3\phi \quad , \tag{25.176}$$

one finds

$$\delta\phi = \frac{4mA}{L^2} = \frac{4m}{b} \quad , \tag{25.177}$$

which agrees precisely with the results of section 25.11.

E: Black Holes

26 Black Holes I: Approaching the Schwarzschild Radius r_s

Recall that the Schwarzschild metric, given in standard coordinates by

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} \quad , \tag{26.1}$$

is the unique spherically symmetric vacuum solution of the Einstein equations. Associated with this solution there is a characteristic length scale, the *Schwarzschild radius* $r = r_s \equiv 2m$. More generally, this is the characteristic lenght scale associated to an object of mass M via the formula

$$r_s = 2m = \frac{2G_N M}{c^2} \quad . \tag{26.2}$$

In our previous discussions of this solution and its properties, in sections 24 and 25, we had considered standard astrophysical objects ("stars") of a size larger (in practice much larger) than their Schwarzschild radius, $r_0 > r_s = 2m$. As a consequence we did not have to address the question what happens to the metric or the geometry as $r \to r_s$ (which was in the deep interior of the star, not desribed by the Schwarzshild metric). We will now contemplate the existence of objects with radius $r_0 < r_s$.

26.1 Preliminary Considerations

Thus we now need to understand the behaviour of the Schwarzschild geometry / gravitational field as one approaches or crosses $r = r_s$. At that radius something special appears to happen to the metric, even though it is not clear what precisely happens there. On the one hand, we noted that the effective potential $V_{eff}(r)$ is perfectly well behaved at r_s. On the other hand, some components of the metric evidently become singular (zero or infinite) there.

Whether this indicates a true singularity of the geometry or the gravitational field (such as infinite tidal forces) or is perhaps simply due to an unfortunate choice of coordinates is usually not something that can just be decided by superficial inspection but requires a more detailed investigation.

Just to illustrate this point, it is a familiar fact that the (apparent) degeneracy of the Euclidean metric on $\mathbb{R}^2$ in polar coordinates,

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2 \quad , \tag{26.3}$$

at $r = 0$ is simply a coordinate singularity. Likewise, the fact that $g_{\phi\phi} \to \infty$ as $r \to \infty$ is not an indication of a singular behaviour of the metric there. By changing variables, $r = 1/u$, one finds that the line element takes the form

$$ds^2 = u^{-4}du^2 + u^{-2}d\phi^2 \quad . \tag{26.4}$$

In this case, both the non-zero components of the metric diverge as $u \to 0$, but again we know that this is just due to an unfortunate choice of coordinates. However, if one had just been given the above metric in the coordinates (u, ϕ) without an explanation how it was obtained, one might not have realised immediately that this is just the Euclidean metric in disguise. In fact, one can compare and contrast this e.g. with the metric

$$ds^2 = du^2 + u^{-2} d\phi^2 \quad , \tag{26.5}$$

which may look somewhat less singular than the previous metric but is actually genuinely singular as $u \to 0$. This can be verified by calculating its curvature following e.g. the examples given in section 11.3: then one finds that its Gauss curvature or Ricci scalar is proportional to u^{-2} (while the Gauss curvature of the Euclidean metric on $\mathbb{R}^2$ is of course identically zero, regardless of which coordinates one uses).

Thus in order to decide what happens or does not happen as $r \to r_s$, we need to take a closer look at the metric. We will do this in several steps, which you can think of as the various stages of an expedition, from the familiar region at $r \gg r_s$, with a map provided by the Schwarzschild coordinates (t, r, θ, ϕ), into potentially dangerous new and uncharted territory:

- First we will consider observers that don't quite dare to cross r_s and who try to remain static at a fixed value of r close to r_s.
- Then we will consider observers that fall freely (and radially) in this geometry, and describe their voyage both from their point of view and from that of a distant static observer (which will turn out to be quite different).
- Then we will study the geometry of the Schwarzschild metric near $r = r_s$, and show that the geometry is completely non-singular (and is in fact closely related to the geometry of the Rindler metric for Minkowski space-time discussed in section 1.3).
- These considerations will indicate that (and explain why) the usual Schwarzschild coordinates (specifically the time-coordinate t) are inadequate for describing the physics across the radius r_s.
- Encouraged by this, in section 27 we then begin to explore the region near and beyond $r = r_s$. To that end we first introduce coordinates that are adapted to infalling observers. This is an illuminating exercise which is interesting in its own right and which will already tell us something about the hidden geometry behind r_s.
- In order to learn more about this region, we next study the behaviour of lightcones and lightrays in this geometry. These considerations will then lead us to the introduction of corresponding adapted coordinates in which the Schwarzschild

metric is non-singular for all $0 < r < \infty$, and then we will be in a position to understand what actually happens (and what characterises) $r = r_s$.

26.2 Static Observers

Some insight into the Schwarzschild geometry, and the difference between Newtonian gravity and general relativistic gravity, is provided by looking at static observers, i.e. observers hovering at fixed values of (r, θ, ϕ). Thus their 4-velocity $u^\alpha = \dot{x}^\alpha$ has the form $u^\alpha = (u^0, 0, 0, 0)$ with $u^0 > 0$. The normalisation $u^\alpha u_\alpha = -1$ then implies

$$u^\alpha = \left(f(r)^{-1/2}, 0, 0, 0\right) = \left(\frac{1}{\sqrt{1 - 2m/r}}, 0, 0, 0\right) \quad . \tag{26.6}$$

Note that, as in (3.125), we could have written this more invariantly as

$$u^\alpha = V^\alpha / \mathbb{V} \tag{26.7}$$

where $V = \partial_t$ is the timelike Killing vector and $\mathbb{V}$ its norm (3.124), but for present (pragmatic and calculational) purposes the explicit coordinate expression is more useful.

The worldline of a static observer is clearly not a geodesic (that would be the worldline of an observer freely falling in the gravitational field), and we can calculate its covariant acceleration (5.99)

$$a^\alpha = D_\tau u^\alpha = u^\beta \nabla_\beta u^\alpha = (d/d\tau) u^\alpha + \Gamma^\alpha_{\beta\gamma} u^\beta u^\gamma \quad . \tag{26.8}$$

Noting that u^α is time-independent, because the observer is at fixed r by hypothesis, one finds

$$a^\alpha = \Gamma^\alpha_{00} (u^0)^2 = \Gamma^\alpha_{00} (1 - 2m/r)^{-1} \quad . \tag{26.9}$$

Thus only a^r is non-zero, and one finds

$$a^\alpha = (0, m/r^2, 0, 0) \quad . \tag{26.10}$$

This looks nicely Newtonian, with a force in the radial direction designed to precisely cancel the gravitational attraction. However, this is a bit misleading since this is still a coordinate dependent statement. A coordinate-invariant quantity (scalar) is the norm of the acceleration,

$$\mathsf{a}(r) \equiv \left(g_{\alpha\beta} a^\alpha a^\beta\right)^{1/2} = \frac{m}{r^2} \left(1 - \frac{2m}{r}\right)^{-1/2} \quad . \tag{26.11}$$

While this approaches the Newtonian value as $r \to \infty$, it diverges as $r \to 2m$, indicating that static observers will find it harder and harder to remain static close to $r = 2m$.

Remarks:

1. At the same time, the discrepancy (gravitational time delay) between the static observer's proper time and the coordinate time (proper time of a static observer at $r \to \infty$) becomes more and more pronounced,
$$\Delta\tau = (1 - \frac{2m}{r})^{1/2}\Delta t \quad , \tag{26.12}$$
and diverges as $r \to 2m$. This leads to an infinite gravitational redshift (see section 26.5 below). It is also the first (but all by itself not conclusive) indication, that the time coordinate t may not be appropriate for understanding the physics at and beyond $r = 2m$.

2. While $\mathsf{a}(r)$ diverges as $r \to 2m$, we will see later (section 27.10) that the finite quantity $\lim_{r\to 2m} f(r)^{1/2}\mathsf{a}(r) = 1/4m$ can be regarded as a measure of the strength of the gravitational field at $r = 2m$ (*surface gravity*).

3. One can also think of $\mathsf{a}(r) = \mathsf{a}^r(r)$ as the radial component of the acceleration with respect to an orthonormal frame (see section 4.8) at that point: a radial unit vector is
$$e_r = (1 - 2m/r)^{1/2}\partial_r : \quad g_{\alpha\beta}e_r^\alpha e_r^\beta = 1 \quad , \tag{26.13}$$
and the acceleration vector can be written as
$$a = a^r\partial_r = a^r(1 - 2m/r)^{-1/2}((1 - 2m/r)^{1/2}\partial_r) \equiv \mathsf{a}^r e_r \quad . \tag{26.14}$$

4. The covariant components of the acceleration can be written as the gradient
$$a_\alpha = \partial_\alpha(\tfrac{1}{2}\ln f(r)) \equiv \partial_\alpha\Phi(r) \quad , \tag{26.15}$$
so that here (and in the static spherically symmetric case in general) $\Phi(r)$ can be regarded as the general relativistic analogue of the Newtonian potential $\phi(r)$, to which it reduces in the Newtonian limit,
$$\Phi(r) = \tfrac{1}{2}\ln f(r) \approx \tfrac{1}{2}\ln(1 + 2\phi(r)) = \phi(r) + \ldots \tag{26.16}$$

5. Using the covariant characterisation (26.7) of the 4-velocity of a static observer, this result can be derived and generalised as follows. First of all the acceleration of u^α is
$$a_\beta = (V^\alpha/\mathbb{V})\nabla_\alpha(V_\beta/\mathbb{V}) = (V^\alpha/\mathbb{V}^2)\nabla_\alpha V_\beta \quad , \tag{26.17}$$
because $\mathbb{V}$ is constant along V (9.61). Using the Killing condition this can be written as
$$a_\beta = -(V^\alpha/\mathbb{V}^2)\nabla_\beta V_\alpha = \tfrac{1}{2}\nabla_\beta \log(-V_\alpha V^\alpha) \quad , \tag{26.18}$$
which provides the appropriate generalisation of (26.15), suggesting that a suitable general relativistic analogue of the scalar Newtonian potential in space-times admitting a timelike Killing vector is provided by the logarithm of the redshift factor $\mathbb{V}$,
$$\Phi = \tfrac{1}{2}\log(-V_\alpha V^\alpha) = \log\mathbb{V} \quad . \tag{26.19}$$

26.3 Vertical Free Fall

We will now consider an object with $r_0 < r_s$ and an observer who is freely falling vertically (radially) towards such an object. "Vertical" means that $\dot{\phi} = 0$, and therefore there is no angular momentum, $L = 0$. Hence the effective potential equation (25.31) becomes

$$E^2 - 1 = \dot{r}^2 - \frac{2m}{r} \quad , \tag{26.20}$$

where the conserved energy E (per unit mass) is given by

$$E = (1 - \frac{2m}{r})\dot{t} \quad . \tag{26.21}$$

In particular, if r_i is the point at which the particle (observer) A was initially at rest,

$$\frac{dr}{d\tau}|_{r=r_i} = 0 \quad , \tag{26.22}$$

we have the relation

$$E^2 = 1 - \frac{2m}{r_i} = f(r_i) \tag{26.23}$$

between the constant of motion E and the initial condition r_i. In particular, $E = 1$ for an observer following a trajectory of an object that would have initially been at rest at infinity. In that case, (26.20) is readily integrated to give

$$r(\tau) \sim (\tau_0 - \tau)^{2/3} \tag{26.24}$$

so that the observer reaches $r = 0$ at $\tau = \tau_0$.

For $r_i < \infty$, we obtain

$$\dot{r}^2 = \frac{2m}{r} - \frac{2m}{r_i} = f(r_i) - f(r) \tag{26.25}$$

and, upon differentiation (for any E)

$$\ddot{r} + \frac{m}{r^2} = 0 \quad . \tag{26.26}$$

This is just like the Newtonian equation (which should not come as a surprise as V_{eff} coincides with the Newtonian potential for zero angular momentum $L = 0$), apart from the fact that r is *not* radial distance and the familiar $\tau \neq t$. Nevertheless, calculation of the time τ along the path proceeds exactly as in the Newtonian theory. For the proper time required to reach the point with coordinate value $r = r_f$ we obtain

$$\tau = -(2m)^{-1/2} \int_{r_i}^{r_f} dr \left(\frac{r_i r}{r_i - r} \right)^{1/2} \quad . \tag{26.27}$$

Since this is just the Newtonian integral, we know, even without calculating it, that it is finite as $r_f \to r_s$ (and even as $r_f \to 0$). This integral can also be calculated in closed form, e.g. via the change of variables

$$r(\eta) = \tfrac{1}{2} r_i \,(1 + \cos\eta) \quad \text{with} \quad \eta_i = 0 \leq \eta \leq \eta_f \leq \pi \quad , \tag{26.28}$$

leading (after some convenient cancellations) to

$$\tau = \left(\frac{r_i}{2m}\right)^{1/2} \int_0^{\eta_f} d\eta\, r(\eta) = \left(\frac{r_i^3}{8m}\right)^{1/2} (\eta_f + \sin\eta_f) \quad . \tag{26.29}$$

In particular, this is finite as $r_f \to 2m$ and our freely falling observer can reach and cross the Schwarzschild radius r_s in finite proper time.

Coordinate time, on the other hand, becomes infinite at $r_f = 2m$. This can be anticipated from the relation (26.21) between t and τ for a freely falling observer, which diverges as $r \to 2m$. We will address this in a more quantitative way in section 26.4 below.

26.4 Vertical Free Fall as seen by a Distant Observer

We will now investigate how the above situation presents itself to a distant observer hovering at a fixed radial distance r_∞. He will observe the trajectory of the freely falling observer as a function of his proper time τ_∞. Up to a constant factor $(1-2m/r_\infty)^{1/2}$, this is the same as coordinate time t, and we will lose nothing by expressing r as a function of t rather than as a function of τ_∞.

From (26.20),

$$\dot{r}^2 + (1 - \frac{2m}{r}) = E^2 \quad , \tag{26.30}$$

which expresses r as a function of the freely falling observer's proper time τ, and the definition of E,

$$E = \dot{t}(1 - \frac{2m}{r}) \quad , \tag{26.31}$$

which relates τ to the coordinate time t, one finds an equation for r as a function of t,

$$\frac{dr}{dt} = -E^{-1}(1 - \frac{2m}{r})(E^2 - (1 - \frac{2m}{r}))^{1/2} \tag{26.32}$$

(the minus sign has been chosen because r decreases as t increases). At large distances, the difference between t and τ, and hence between $r = r(t)$ and $r = r(\tau)$ is small and not particularly interesting in the present context. We want to analyse the behaviour of the solution of this equation as the freely falling observer approaches the Schwarzschild radius, $r \to 2m$. In that region, we can approximate

$$1 - \frac{2m}{r} = \frac{r-2m}{r} \approx \frac{r-2m}{2m} \quad , \tag{26.33}$$

and we get

$$\begin{aligned} \frac{dr}{dt} &= -E^{-1}(\frac{r-2m}{r})(E^2 - \frac{r-2m}{r})^{1/2} \\ &\approx -E^{-1}(\frac{r-2m}{2m})(E^2)^{1/2} = -(\frac{r-2m}{2m}) \quad . \end{aligned} \tag{26.34}$$

We can write this equation as

$$\frac{d}{dt}(r-2m) = -\frac{1}{2m}(r-2m) \quad , \tag{26.35}$$

which obviously has the solution

$$(r-2m)(t) \propto \mathrm{e}^{-t/2m} \quad . \tag{26.36}$$

This shows that, from the point of view of the observer at infinity, the freely falling observer reaches $r = 2m$ only as $t \to \infty$. In particular, the distant observer will never actually see the infalling observer cross the Schwarzschild radius.

This is clearly an indication that there is something wrong with the time coordinate t which runs too fast as one approaches the Schwarzschild radius. We can also see this by looking at the coordinate velocity $v = dr/dt$ as a function of r. Let us choose $r_i = \infty$ for simplicity - other choices will not change our conclusions as we are interested in the behaviour of $v(r)$ as $r \to r_s$. Then $E^2 = 1$ and from (26.32) we find (now dropping the minus sign)

$$v(r) = (2m)^{1/2}\frac{r-2m}{r^{3/2}} \tag{26.37}$$

As a function of r, $v(r)$ reaches a maximum at the critical radius $r_c = 6m = 3r_s$,

$$\frac{d}{dr}v(r) = 0 \quad \Rightarrow \quad r = r_c = 6m \quad , \tag{26.38}$$

where the velocity is (restoring the speed of light c)

$$v(r_c) = \frac{2c}{3\sqrt{3}} \quad . \tag{26.39}$$

The fact that this radius agrees with the ISCO (innermost stable circular orbit) (25.73) mentioned in section 25.6 should (presumably) be regarded as a coincidence.

Beyond that point, i.e. for $r < r_c$, $v(r)$ decreases again and clearly goes to zero as $r \to 2m$. The fact that the coordinate velocity goes to zero is simply another manifestation of the fact that coordinate time goes to infinity, and that the Schwarzschild coordinates are simply not suitable for describing the physics at or beyond the Schwarzschild radius because the time coordinate one has chosen is running too fast. This is the crucial insight that will later on allow us to construct "better" coordinates, which are also valid for $r < r_s$.

Remarks:

1. Sometimes (actually all too frequently) the fact that the coordinate velocity is decreasing as one approaches $r \to 2m$ is claimed to provide evidence for some "repulsive" nature of the gravitational field as one approaches $r = 2m$. While it is understandable that such things caused some confusion in the early days of general relativity, this is of course utter nonsense for which there can be no excuse today.

2. As an aside, if one repeats the above "calculation" for arbitrary E, from (26.32) one finds that the critical radius is

$$r_c = \frac{2m}{1 - 2E^2/3} \quad , \tag{26.40}$$

with the E-dependent maximal velocity

$$v(r_c) = \frac{2c}{3\sqrt{3}} E^2 \quad . \tag{26.41}$$

This reproduces the above result for $E = 1$, shows that $r_c \to 2m$ for $E \to 0$, and moreover curiously shows that there is a maximal E, $E_{max} = \sqrt{3/2}$, for which such a critical point of the coordinate velocity will occur, at $r_c \to \infty$, with maximal velocity

$$E \to E_{max} = \sqrt{3/2} \quad \Rightarrow \quad r_c \to \infty \quad \text{and} \quad v(r_c) \to v_{max} = c/\sqrt{3} \quad . \tag{26.42}$$

Thus for larger values of E, the coordinate velocity will monotonically decrease from its initial value

$$v(r \to \infty) = E^{-1}(E^2 - 1)^{1/2} \tag{26.43}$$

to $v(r) = 0$ as $r \to 2m$.

[And if anybody has an intuitive explanation for these facts, please let me know.]

26.5 Gravitational Redshift in the Schwarzschild Geometry

One dramatic aspect of what is happening at (or, better, near) the Schwarzschild radius for very (very!) compact objects with $r_s > r_0$ is the following. Recall the formula (3.108) for the gravitational redshift, which gave us the ratio between the frequency of light ν_e emitted at the radius r_e and the frequency ν_∞ received at the radius $r_\infty > r_e$ for static observers in a static spherically symmetric gravitational field.

The result, which is in particular also valid for the Schwarzschild metric, was

$$\frac{\nu_\infty}{\nu_e} = \frac{(-g_{00}(r_e))^{1/2}}{(-g_{00}(r_\infty))^{1/2}} \quad . \tag{26.44}$$

In the case of the Schwarzschild metric, this is

$$\frac{\nu_\infty}{\nu_e} = \frac{(1 - 2m/r_e)^{1/2}}{(1 - 2m/r_\infty)^{1/2}} \quad , \tag{26.45}$$

in accordance with the relation (26.12) between the coordinate time and the proper time of a static observer. As $r_e \to r_s$, one clearly finds

$$\frac{\nu_\infty}{\nu_e} \to 0 \quad . \tag{26.46}$$

Expressed in terms of the gravitational redshift factor z,

$$1 + z = \frac{\nu_e}{\nu_\infty} \tag{26.47}$$

this means that there is an *infinite gravitational redshift* as $r_e \to r_s$,

$$r_e \to r_s \quad \Rightarrow \quad z \to \infty \quad . \tag{26.48}$$

This is for static observers. For a freely falling observer whose position is described by $r_e = r(\tau)$ or $r(t)$ and the receiver the static observer at $r_\infty \gg r_s$ we cannot directly apply the above formula (because the emitter is not static but freely falling), and the actual redshift of the emitted lightray will be larger because of an additional Doppler effect due to the observer's proper motion away from the receiver.

If (for the time being) we ignore this additional Doppler contribution, and just use the above formula (this will thus underestimate the actual redshift) in conjunction with the late-time behaviour for $r = r(t)$ derived in (26.36), we can obtain the late-time behaviour of the redshift factor $z = z(t)$ as a function of the distant observer's proper time t, namely

$$1 + z \propto (r - 2m)(t)^{-1/2} \propto \mathrm{e}^{t/4m} \quad . \tag{26.49}$$

Taking into account the proper motion of the freely falling observer one finds (we will derive this below) that, essentially due to the difference between $u^0 = \dot{t} = f^{-1/2}$ for a static observer and $u^0 = \dot{t} = E/f$ for a geodesic observer, as $r \to 2m$ the redshift is enhanced by a further factor $f^{-1/2}$, so that all in all one has $(1 + z) \propto (r - 2m)(t)^{-1}$ instead of (26.49). See (26.67) below.

Either way for the distant observer at late times there is an exponentially growing redshift and the distant observer will never actually see the unfortunate emitter crossing the Schwarzschild radius: he/she will see the freely falling observer's signals becoming dimmer and dimmer and arriving at greater and greater intervals, and the freely falling observer will completely disappear from the distant observer's sight as $r_e \to r_s$. Note that the time-scale t_z for this exponential redshift at late times is set by $t_z \sim m/c$, which is of the order of

$$t_z \sim 10^{-5}\mathsf{s}\left(\frac{M}{M_{\mathsf{sun}}}\right) \quad , \tag{26.50}$$

so that this is pretty much instantaneous for an object the mass of an ordinary star. We will come back to this estimate later on when (briefly) talking about gravitational collapse in section 29.3.

In order to take into account the proper freely falling motion of the emitter, one can use the method described in the second derivation of the gravitational redshift given in section 3.5, based on the equation (3.117),

$$\omega = -u^\alpha k_\alpha \quad , \tag{26.51}$$

for the frequency ω of a lightray with wave vector k^α as seen (or emitted) by an arbitrary (not necessarily static) observer with 4-velocity u^α.

We describe lightrays by null curves $x^\alpha = x^\alpha(\lambda)$ with wave vector

$$k^\alpha = \frac{d}{d\lambda}x^\alpha(\lambda) \equiv x'^\alpha \tag{26.52}$$

(for some affine parameter λ) and the motion of timelike observers (as usual) by $x^\alpha = x^\alpha(\tau)$ with 4-velocity

$$u^\alpha = \frac{d}{d\tau}x^\alpha(\tau) \equiv \dot{x}^\alpha \quad . \tag{26.53}$$

Let us now apply this formula to the case at hand (Schwarzschild geometry and radial motion). The relevant wave and velocity vectors are the following:

- Radial Lightrays

  For lightrays, there is a conserved energy

$$e = f(r)t' \tag{26.54}$$

  and for radial lightrays the effective potential is zero, so that the radial equation of motion for r is simply

$$(r')^2 = e^2 \quad \Rightarrow \quad r' = \pm e \tag{26.55}$$

  (for out/in-going lightrays respectively). Thus out/in-going radial lightrays have wave vector

$$(k^\alpha)_{out/in} = (t', r', 0, 0) = (e/f, \pm e, 0, 0) : \quad g_{\alpha\beta}k^\alpha k^\beta = 0 \quad . \tag{26.56}$$

- Static Observers

  For static observers one has (26.6)

$$(u^\alpha)_s = (f^{-1/2}, 0, 0, 0) : \quad g_{\alpha\beta}u^\alpha u^\beta = -1 \quad . \tag{26.57}$$

- Radial Freely Falling Observers

  For geodesic (freely falling) obervers, there is a conserved energy

$$E = f(r)\dot{t} \tag{26.58}$$

  and for radial free fall the effective potential is just the Newtonian potential, so that the radial equation of motion is simply (26.20)

$$\dot{r}^2 = E^2 - f \quad \Rightarrow \quad \dot{r} = -(E^2 - f)^{1/2} \tag{26.59}$$

  (for infalling observers). Thus their 4-velocity is

$$(u^\alpha)_{ff} = (\dot{t}, \dot{r}, 0, 0) = (E/f, -(E^2 - f)^{1/2}, 0, 0) : \quad g_{\alpha\beta}u^\alpha u^\beta = -1 \quad . \tag{26.60}$$

From these and (26.51) we can now read off the following results:

1. Static Observers and In/Outgoing Lightrays

   The frequency of an in/out-going lightray as seen by a static observer at the radius r is

$$\omega_{s,in/out}(r) = -g_{\alpha\beta}(u^\alpha)_s(k^\beta)_{in/out} = f(r)f(r)^{-1/2}(e/f(r)) = ef(r)^{-1/2} \quad . \tag{26.61}$$

   Note that this only depends on the time component of k^α and is thus the same for in- and outgoing lightrays. In particular,

   - this identifies the energy e as the frequency as seen by an observer "at infinity",

$$\omega_\infty \equiv \lim_{r\to\infty} \omega_{s,in/out} = e \quad , \tag{26.62}$$

   - and reproduces the standard result that the redshift of light emitted by a static observer at radius r relative to a distant static observer scales like $f(r)^{-1/2}$,

$$\frac{\omega_\infty}{\omega_{s,in/out}} = f(r)^{1/2} \quad . \tag{26.63}$$

2. Freely Falling Observers and In- versus Outgoing Lightrays

   For a freely falling observer instead, one has

$$\omega_{ff,in/out} = -g_{\alpha\beta}(u^\alpha)_{ff}k^\beta_{in/out} = f(E/f)(e/f) \mp f^{-1}(E^2-f)^{1/2}e \tag{26.64}$$

   and thus (restoring the r-dependence)

$$\omega_{ff,in/out}(r) = \frac{e}{f(r)}(E \mp (E^2 - f(r))^{1/2}) \quad . \tag{26.65}$$

   Note that, roughly speaking, and as anticipated above, compared to the result for a static emitter or observer, the redshift/blueshift is enhanced by another factor $f(r)^{\pm 1/2}$ due to a Doppler-like effect caused by the motion of the observer or emitter.

   This effect is of course particularly pronounced if either emitter or observer approaches $r = 2m$. More specifically, we can consider the following two situations:

   (a) Outgoing Lightrays

   Taking the limit $f(r) \to 0$, one finds that

$$f(r) \to 0 \quad \Rightarrow \quad \omega_{ff,out}(r) \approx \frac{2eE}{f(r)} \quad \Rightarrow \quad \frac{\omega_\infty}{\omega_{ff,out}(r)} \approx \frac{f(r)}{2E} \to 0 \quad . \tag{26.66}$$

   Thus from the perspective of a static observer an outgoing lightray sent out by an infalling observer will undergo an infinite redshift proportional to

$(r-2m)^{-1}$ as the emitter approaches $r \to 2m$. In particular, here we are now allowed to insert the time dependence (26.36) and instead of (26.49) we will obtain

$$1 + z \propto (r-2m)(t)^{-1} \propto \mathrm{e}^{t/2m} \quad . \tag{26.67}$$

(b) Ingoing Lightrays

We can also consider the opposite situation where an ingoing lightray reaches an infalling observer (from above/behind) as $r \to 2m$. In that case, the lightray is detected by the infalling observer with a finite redshift,

$$f(r) \to 0 \quad \Rightarrow \quad \omega_{ff,in}(r) \approx \frac{e}{2E} \quad \Rightarrow \quad \frac{\omega_\infty}{\omega_{ff,in}} = 2E \quad . \tag{26.68}$$

In particular, for a freely falling observer with $E = 1$, i.e. an oberver who falls as if he/she had started off at rest at infinity, the frequency he/she measures is precisely 1/2 the frequency ω_∞.

This is perhaps one "practical" way for this infalling observer to determine when precisely he/she reaches $r = 2m$: just after the agreed upon almost-UV violet light with emitter frequency 790 THz (at the upper end of the visible spectrum) arrives as the almost-IR red light with frequency 405 THz (at the lower end of the visible spectrum).

26.6 Geometry near r_s and Minkowski Space in Rindler Coordinates

We have now seen in two different ways why the Schwarzschild coordinates are not suitable for exploring the physics in the region $r \leq 2m$: in these coordinates the metric becomes singular at $r = 2m$ and the coordinate time becomes infinite. On the other hand, we have seen no indication that the local physics, expressed in terms of covariant quantities like proper time or the geodesic equation, becomes singular as well. So we have good reasons to suspect that the singular behaviour we have found is really just an artefact of a bad choice of coordinates.

In fact, the situation regarding the Schwarzschild coordinates is quite similar to that of the Rindler coordinates for Minkowski space we discussed (way back) in section 1.3.

Recall from section 1.3 that

- in the coordinates

$$\xi^0(\eta,\rho) = \rho \sinh\eta \qquad \xi^1(\eta,\rho) = \rho\cosh\eta \quad . \tag{26.69}$$

  the $(1+1)$-dimensional Minkowski metric takes the form (1.74)

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 \quad ; \tag{26.70}$$

- that the lines of constant ρ are hyperbolas and that these are the worldlines of observers with constant acceleration $1/\rho$;
- that these coordinates are adapted to observers with constant acceleration in the same way as inertial coordinates are adapted to static observers: they stay at fixed values of their spatial coordinate, and the coordinate time is a direct measure of their proper time;
- the null lines $\xi^0 = \pm\xi^1$ correspond to $\rho = 0$ (infinite acceleration) or $\eta = \pm\infty$, so that the Rindler coordinates cover only the first quadrant $\xi^1 > |\xi^0|$ of Minkowski space;
- inertial (geodesic, freely falling) observers could exit this region in finite proper time.

All this is of course quite reminiscent of the things that we have discovered so far about the Schwarzschild geometry, and this is in fact more than just a loose analogy: as we will see now, remarkably the Rindler metric (26.70) gives an accurate description of the geometry of the Schwarzschild metric close to the Schwarzschild radius.

To confirm this, let us temporarily introduce the variable $\tilde{r} = r - 2m$ measuring the coordinate distance from the critical radius $r = 2m$. In term of $\tilde{r}$ the (t, r)-part of the Schwarzschild metric reads

$$ds^2 = -\left(\frac{\tilde{r}}{\tilde{r} + 2m}\right) dt^2 + \left(\frac{\tilde{r} + 2m}{\tilde{r}}\right) d\tilde{r}^2 \quad . \tag{26.71}$$

Close to $r_s = 2m$, i.e. for small $\tilde{r} << 2m$, we can approximate

$$\frac{\tilde{r}}{\tilde{r} + 2m} \approx \frac{\tilde{r}}{2m} \qquad\qquad \frac{\tilde{r} + 2m}{\tilde{r}} \approx \frac{2m}{\tilde{r}} \quad , \tag{26.72}$$

so that the metric becomes

$$ds^2 = -\frac{\tilde{r}}{2m} dt^2 + \frac{2m}{\tilde{r}} d\tilde{r}^2 \quad . \tag{26.73}$$

Introducing the new radial variable ρ (proper radial distance from $r = r_s$) via

$$d\rho^2 = \frac{2m}{\tilde{r}} d\tilde{r}^2 \quad \Rightarrow \quad \rho = \sqrt{8m\tilde{r}} \quad , \tag{26.74}$$

one finds

$$ds^2 = -\frac{1}{16m^2}\rho^2 dt^2 + d\rho^2 \quad . \tag{26.75}$$

Finally a simple rescaling of t, $\eta = t/4m$, leads to

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 \quad , \tag{26.76}$$

which, remarkably, is identical to the Rindler metric (1.74). Keeping track of the transverse 2-sphere and using $r^2 \approx (2m)^2$ in the near-r_s approximation, the complete metric in this limit and in these coordinates reads

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 + (2m)^2 d\Omega^2 \tag{26.77}$$

If we further restrict to just a small angular region on the sphere, we can approximate

$$d\theta^2 + \sin^2\theta d\phi^2 \approx d\theta^2 + \theta^2 d\phi^2 = (dx^2)^2 + (dx^3)^2 \tag{26.78}$$

which gives us the complete 4-dimensional Rindler metric.

In any case, we see that, up to the harmless redefinition $(r,t) \to (\rho,\eta)$ that we just performed, which is just a reparametrisation

$$\begin{aligned} \rho &= \rho(r) = \sqrt{8m(r-2m)} \\ \eta &= \eta(t) = t/4m \end{aligned} \tag{26.79}$$

Schwarzschild coordinates for the Schwarzschild geometry near $r = 2m$ are just like Rindler coordinates for Minkowski space. This leads to a much improved understanding of the Schwarzschild geometry in general and the observations that we made regarding static and freely falling observers in particular:

1. The first, and most crucial, thing we learn from this is that the singularity of the Schwarzschild metric at $r = 2m$ in the Schwarzschild coordinates (t,r) is, as anticipated, a mere coordinate singularity. Indeed, $r = 2m$ corresponds to $\tilde{r} = 0 \Leftrightarrow \rho = 0$, and we already know that what appears to be a possibly singular point of the Rindler metric at $\rho = 0$ is just a coordinate singularity which can be eliminated e.g. by passing to standard inertial Minkowski coordinates via (1.73).

   Specifically, from (26.79) one could introduce Minkowski-like coordinates (t_M, x_M) (say), via

   $$\begin{aligned} t_M &= \rho \sinh\eta = \sqrt{8m(r-2m)}\sinh t/4m \\ x_M &= \rho\cosh\eta = \sqrt{8m(r-2m)}\cosh t/4m \ , \end{aligned} \tag{26.80}$$

   in terms of which the metric in the region $r \gtrsim 2m$ takes the standard Minkowskian form $-dt_M^2 + dx_M^2$.

   The reason for not pursuing this here is that we will actually be able to do much better than this in various ways in section 27, where we will construct a variety of coordinate systems which exhibit the non-singular nature of the geometry at $r = 2m$ but which are not restricted to only a tiny (strictly speaking infinitesimal) neighbourhood of $r = 2m$. In particular, the above coordinate transformation (26.80) can be understood as the $r \gtrsim 2m$ approximation of the coordinate transformation from Schwarzschild coordinates to *Kruskal coordinates* (to be introduced in section 27.7).

2. The second crucial insight is that, while the surfaces $r =$ constant for $r > 2m$ are all timelike (they contain the timelike worldlines of static observers), the surface $r = 2m$ is actually null (lightlike), just like the line $\rho = 0, \eta = \pm\infty$ in Rindler space. It is then no surprise that the surface $r = 2m$ is special and has some rather unintuitive properties.

   We can now understand physically why the Schwarzschild coordinates break down at $r = 2m$: they are adapted to accelerating massive observers, in the present context the static observers in the Schwarzschild geometry at constant $r > 2m$, with proper time proportional to the Schwarzschild coordinate time t. Referring back to Figure 7 in section 1.3, these are the observers with constant ρ hyperbolic world lines.

   The problem is evidently that the required acceleration of these observers becomes infinite as $\rho \to 0 \Leftrightarrow r \to 2m$ (as we have calculated in section 26.2), because $r = 2m$ is lightlike, not timelike. That these observers appear to see a singular metric is then not the geometry's fault but can be attributed to a bad choice of observers whose perception of the geometry is distorted by their acceleration and their desperate attempt to stay at constant values of r even when they are very close to $r = 2m$.

3. The situation is quite different for the freely falling observers. Their worldlines look like the vertical line labelled "worldline of a static observer" in Figure 7, they cross the "horizon" in finite proper time, experiencing no strong acceleration or gravitational fields. As already noted in section 1.3, they evidently become invisible to observers in the Rindler quadrant (now "Schwarzschild patch") of the geometry, static outside observers noting an infinite gravitational redshift affecting the signals sent out by the freely falling observer.

4. Introducing coordinates adapted to freely falling observers (so that e.g. time is their proper time) would be tantamount to passing from Rindler coordinates to ordinary Minkowski coordinates, and we will consider that option in section 27.2 below (Painlevé-Gullstrand coordinates). These will provide us, as we will see, with a coordinate system that extends in a non-singular way across the Schwarzschild radius.

5. In order to elucidate the significance of the Schwarzschild radius, it will then turn out to be useful to base the construction of new coordinates not on the behaviour of timelike freely falling observers but on ingoing lightrays instead (which evidently also do not suffer from the problems of the static observers). These are the Eddington-Finkelstein coordinates to be discussed in section 27.4.

6. Finally, we can anticipate that upon introduction of suitable analogues of the Minkowski coordinates for the Rindler space-time, we may perhaps uncover not

just one new region (quadrant) of space-time (the one lying to the "future" of $r = 2m$), but also counterparts of the other two quadrants of Minkowski space. This expectation will indeed be borne out.

26.7 Lightcones and Tortoise Coordinates

In section 26.3 we had seen that observers following timelike radial ingoing geodesics reach $r = 2m$ at finite values of the affine parameter (proper time), and in section 26.4 that they do so at infinite values of the coordinate time t. The same is *mutatis mutandis* true for ingoing lightrays travelling along null geodesics, and the argument in this case is even simpler.

Indeed, since for radial lightrays the effective potential (25.43) is zero, radial lightrays are governed by the equation

$$\tfrac{1}{2}\dot{r}^2 = E_{eff} = \tfrac{1}{2}E^2 \quad \Rightarrow \quad \dot{r} = \pm E \ , \tag{26.81}$$

the lower sign corresponding to ingoing lightrays. The solution for ingoing lightrays is thus evidently

$$r(\lambda) = r(0) - E\lambda \tag{26.82}$$

so that from any finite initial position $r(0) > 2m$, the Schwarzschild radius $r = 2m$ is reached for the finite value

$$\lambda = \frac{r(0) - 2m}{E} \tag{26.83}$$

of the affine parameter.

Note, by the way, that (26.82) or $\ddot{r} = 0$ shows that r and λ are related by an affine transformation so that one can equally well use r as the affine parameter along ingoing null geodesics, making it even more manifest that $r = 2m$ arises at a finite value of that affine parameter. To reduce clutter, let us choose $r(0) = 2m$ so that $r = 2m$ is reached at $\lambda = 0$. It then follows from $f(r)\dot{t} = E$, i.e.

$$\dot{t} = E/f(r) = Er/(r - 2m) = E + 2mE/(r - 2m) \ , \tag{26.84}$$

that

$$\dot{t}(\lambda) = E + 2mE/(-E\lambda) = E - 2m/\lambda \quad \Rightarrow \quad t(\lambda) = E\lambda - 2m\log(-\lambda) \ . \tag{26.85}$$

This makes it manifest that

$$r \to 2m \quad \Rightarrow \quad \lambda \to 0_- \quad \Rightarrow \quad t \to +\infty \tag{26.86}$$

also for null geodesics. Thus t is evidently not a good coordinate to describe physics at (or beyond) $r = 2m$.

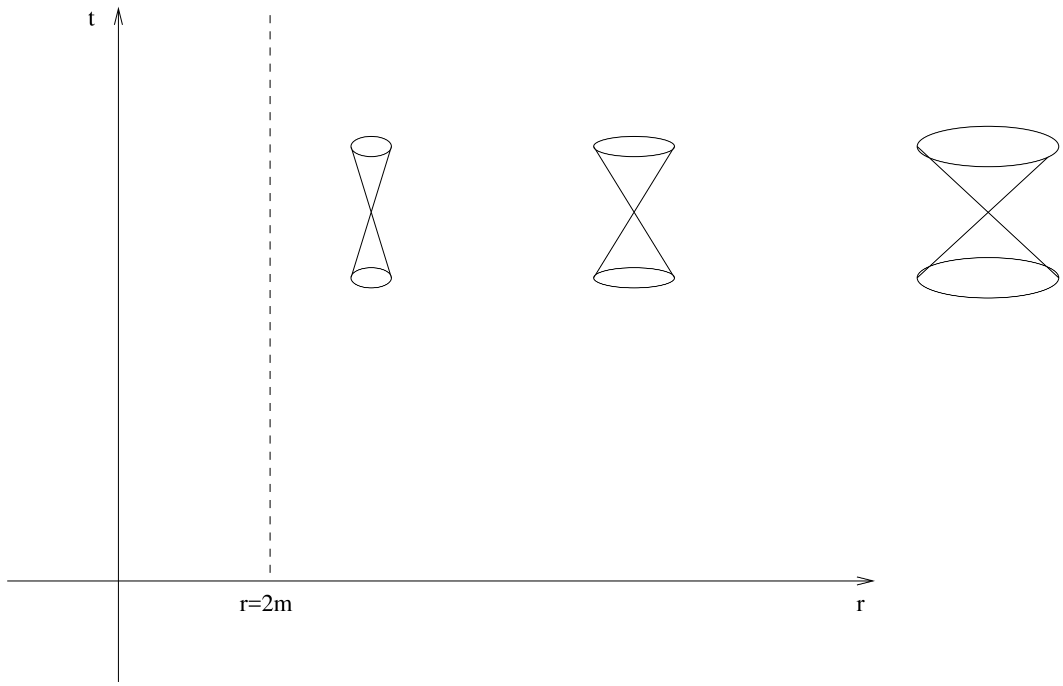


Figure 16: Causal structure of the Schwarzschild geometry in the Schwarzschild coordinates (r,t). As one approaches $r=2m$, the lightcones become narrower and narrower and eventually fold up completely.

To improve our understanding of the Schwarzschild geometry, it is important to study its *causal structure*, i.e. the lightcones. Radial null curves satisfy

$$(1-2m/r)dt^2 = (1-2m/r)^{-1}dr^2 \quad . \tag{26.87}$$

Thus

$$\frac{dt}{dr} = \pm(1-2m/r)^{-1} \quad . \tag{26.88}$$

Recall that, as noted above, this equation for $t=t(r)$ can be equivalently regarded as an equation for $t=t(\lambda)$ as a function of the affine parameter.

In the (t,r)-diagram of Figure 16, dt/dr represents the slope of the lightcones at a given value of r. Now, as $r\to 2m$, one has

$$\frac{dt}{dr} \stackrel{r\to 2m}{\longrightarrow} \pm\infty \quad , \tag{26.89}$$

so the light cones 'close up' as one approaches the Schwarzschild radius. This is the same statement as before regarding the fact that the coordinate velocity goes to zero at $r=2m$, but this time for null rather than timelike geodesics.

As our first step towards introducing coordinates that are more suitable for describing the region around r_s, let us write the Schwarzschild metric in the form

$$ds^2 = (1-2m/r)(-dt^2 + (1-2m/r)^{-2}dr^2) + r^2 d\Omega^2 \quad . \tag{26.90}$$

We see that it is convenient to introduce a new radial coordinate r^* via

$$dr^* = (1 - 2m/r)^{-1}dr = \frac{r}{r-2m}dr = (1 + \frac{2m}{r-2m})dr \quad . \tag{26.91}$$

The solution to this equation is (up to an arbitrary finite constant)

$$r^* = r + 2m\log(r/2m - 1) \quad . \tag{26.92}$$

This new radial coordinate r^*, known as the *Regge - Wheeler radial coordinate* or *tortoise coordinate*, also provides us with the solution

$$t = \pm r^* + C_\pm \tag{26.93}$$

(with $C_\pm$ constants of integration) to the equation (26.88) describing the radial light-cones:

- lines of constant $C_+ = t - r^*$ describe *outgoing* lightrays;
- lines of constant $C_- = t + r^*$ describe *ingoing* lightrays.

In terms of r^* the metric simply reads

$$ds^2 = (1 - 2m/r)(-dt^2 + dr^{*2}) + r^2 d\Omega^2 \quad , \tag{26.94}$$

where r is to be thought of as a function of r^*.

Now the lightcones, defined by

$$dt^2 = dr^{*2} \quad , \tag{26.95}$$

do not seem to fold up as the lightcones have the constant slope $dt/dr^* = \pm 1$ (see Figure 17), and there is no singularity at $r = 2m$. However, r^* is still only defined for $r > 2m$ and the surface $r = 2m$ has been pushed infinitely far away ($r = 2m$ is now at $r^* = -\infty$). Moreover, even though non-singular, the metric components g_{tt} and $g_{r^*r^*}$ (as well as $\sqrt{g}$) vanish at $r = 2m$.

Thus the tortoise coordinate has so far not really allowed us to make dramatic progress in our exploration of the region near or behind r_s, but we will substantially improve this situation in section 27.

26.8 Massless Klein-Gordon Scalar Field in the Schwarzschild Geometry

As an aside, and to conclude this section, I just want to point out that the tortoise coordinate r^* and the corresponding retarded and advanced time-coordinates $u, v = t \mp r^*$ (which we will reencounter in section 27.4 as part of the Eddington Finkelstein coordinates) are not only useful for clarifying the causal structure of the Schwarzschild

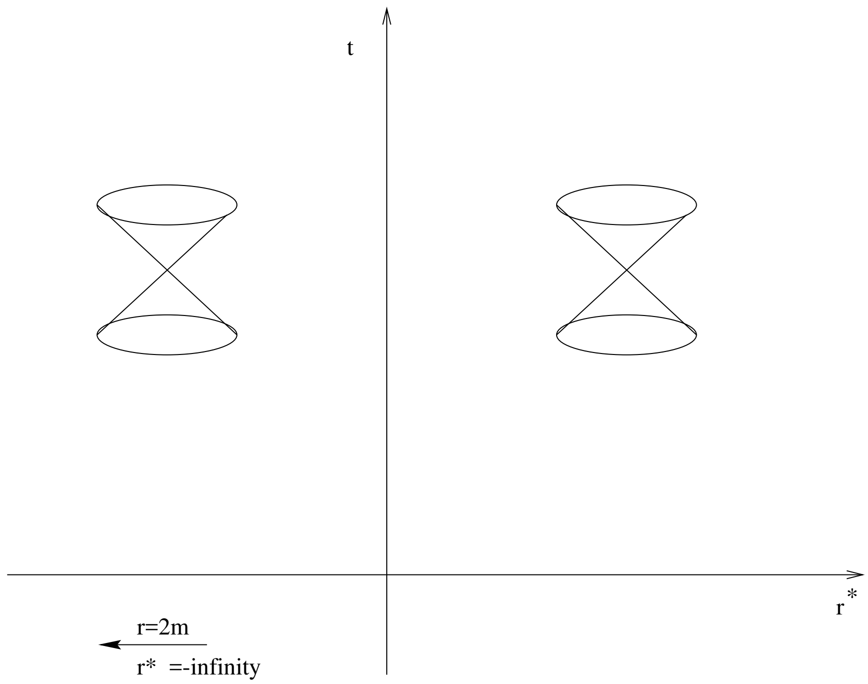


Figure 17: Causal structure of the Schwarzschild geometry in the tortoise coordinates (r^*, t). The lightcones look like the lightcones in Minkowski space and no longer fold up as $r \to 2m$ (which now sits at $r^* = -\infty$).

geometry, but also for the analysis of the propagation of scalar (and other) fields. This is principally due to the fact that in these coordinates the (t, r)-part of the metric is conformally flat - see (26.94) or (27.141) below. Combined with the observation of section 7.7 that the massless Klein-Gordon action is conformally invariant in $(1 + 1)$-dimensions, this leads to a canonical form for the (3+1) wave operator in the (t, r)-sector (and the (θ, ϕ)-sector is standard anyway).

Specifically, we start with the action for a massless scalar field ϕ in the metric (26.94),

$$
\begin{aligned}
S[\phi] &\sim \int \sqrt{g} d^4x\, g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi \\
&= \int dt\, dr^*\, d\Omega\, r^2 f(r) \left[f(r)^{-1}(-(\partial_t\phi)^2 + (\partial_{r^*}\phi)^2) - r^{-2}\phi\, \Delta_{S^2}\phi\right] \qquad (26.96) \\
&= \int dt\, dr^*\, d\Omega\, \left[-(r\partial_t\phi)^2 + (r\partial_{r^*}\phi)^2) - f(r)\phi\, \Delta_{S^2}\phi\right] \quad ,
\end{aligned}
$$

where $f(r) = 1 - 2m/r$, $d\Omega = \sin\theta d\theta\, d\phi$ denotes the solid angle on the 2-sphere, and Δ_{S^2} is the Laplace operator on the 2-sphere. Separating variables according to

$$\phi(x) = r^{-1}\sum_{\ell,m}\psi_{\ell m}(t, r^*)Y_{\ell m}(\theta, \phi) \quad , \qquad (26.97)$$

using

$$\Delta_{S^2}Y_{\ell m} = -\ell(\ell+1)Y_{\ell m} \quad , \qquad (26.98)$$

and using

$$\partial_{r^*} r = f(r) \quad \Rightarrow \quad r\partial_{r^*}(\psi_{\ell m}/r) = \partial_{r^*}\psi_{\ell m} - r^{-1} f(r)\psi_{\ell m} \ , \tag{26.99}$$

one finds for $\psi = \psi_{\ell m}$ the equation of motion

$$(\partial_t^2 - \partial_{r^*}^2)\psi + V_\ell(r^*)\psi = 0 \ , \tag{26.100}$$

where

$$V_\ell(r^*) = f(r)\left(\frac{\ell(\ell+1)}{r^2} + \frac{2m}{r^3}\right) \ . \tag{26.101}$$

<u>Remarks:</u>

1. Note that this potential is quite similar to the effective potential (25.43)

$$V_{eff}(r) = f(r)\frac{L^2}{2r^2} \tag{26.102}$$

for massless particles in the Schwarzschild geometry derived in section 25.3. Indeed, in the large ℓ limit the last term in (26.101) can be neglected and then V_ℓ reduces to V_{eff} with the identification $\ell(\ell+1) \to L^2$. This is as it should be and as one expects (and can show) on general grounds: in a suitable geometric optics (large ℓ, high frequency) limit the massless Klein-Gordon equation reduces to the Hamilton-Jacobi equation for null geodesics.

2. This potential is non-negative for all $2m < r < \infty$ and goes to zero quite rapidly as $r \to \infty$ or $r \to 2m$, i.e. as $r^* \to \pm\infty$,

$$\begin{aligned} r^* \to +\infty \Leftrightarrow r \to \infty : &\quad V_\ell(r) \sim (r^*)^{-2} \\ r^* \to -\infty \Leftrightarrow r \to 2m : &\quad V_\ell(r) \sim \mathrm{e}^{r^*/2m} \ . \end{aligned} \tag{26.103}$$

This means that at infinity (in r) and near the horizon, the solutions of this equation can be chosen to have the standard right-moving (outgoing) / left-moving (ingoing) form

$$\psi(t, r^*) \sim \mathrm{e}^{\pm i\omega(t - r^*)} = \mathrm{e}^{\pm i\omega u} \quad \text{or} \quad \psi(t, r^*) \sim \mathrm{e}^{\pm i\omega(t + r^*)} = \mathrm{e}^{\pm i\omega v} \ . \tag{26.104}$$

3. However, this does not mean that a mode having the above form near infinity, say, evolved from a mode that had also had such a form near the horizon. Rather, the almost infinite exponential gravitational redshift between the near-horizon and asymptotic regions discussed in section 26.5 leads to an exponential relation between the parameter u, say, labelling an outgoing wave at infinity, and the corresponding parameter near the horizon. This exponential relation is analogous to that encountered for a scalar field in Rindler versus inertial Minkowski coordinates (section 7.8). For the Schwarzschild metric it is encoded in the precisely

analogous relation (27.156) between the coordinates u, v and the Kruskal coordinates u_K, v_K to be introduced below, the latter being the analogues of Minkowski inertial coordinates, and the former the analogues of Rindler coordinates. These observations are at the heart of the so-called *Hawking Effect*, i.e. the quantum radiation of black holes. See the references given in section 27.7 for introductions to these topics.

4. By separating out the time-dependence,

$$\psi(t, r^*) = \mathrm{e}^{-i\omega t}\psi(r^*) \quad , \tag{26.105}$$

the exact equation to be solved takes the form of a standard time-independent Schrödinger equation,

$$-\partial_{r^*}^2\psi + V_\ell(r^*)\psi = \omega^2\psi \quad . \tag{26.106}$$

It plays an important role in numerous aspects of Black Hole physics, e.g. in the analysis of the stability of the Schwarzschild solution. In this context, the above equation and its counterparts for vectors and symmetric tensors are known as the Regge-Wheeler(-Zerilli) equations.

27 Black Holes II: the Schwarzschild Black Hole

27.1 Preliminary Remarks

The primary purpose of this section is to understand the significance and physics of the Schwarzschild radius and the region $r < 2m$. We will acccomplish this principally via a construction of appropriate, physically motivated coordinate systems and, indeed, a secondary purpose of this section is to illustrate how to go about constructing such coordinate systems in a systematic way (instead of just introducing them without further explanations).

Even though the details will differ, the general principles of how to construct coordinates to explore and understand a given space-time (by constructing and using coordinates adapted to preferred classes of observers or geodesics) can be applied to other metrics, e.g. some of those listed in section 30.

Let me also make clear from the outset what the issue is *not*. Namely, the issue is not one of constructing appropriate coordinates solely in the region $0 < r < 2m$. Indeed, we have known such coordinates all along, simply the original Schwarzschild coordinates (t, r). The Schwarzschild metric is a vacuum solution of the Einstein equations also in that region, and the coordinates (t, r) give a valid non-singular description of the metric there. Since $f(r) = 1 - 2m/r < 0$, their interpretation differs, i.e. r is a timelike coordinate, and t plays the role of a radial coordinate, but this notational issue can easily be rectified by renaming $r = T, t = R$, and writing the Schwarzschild metric in the region $0 < r = T < 2m$ as

$$\begin{aligned} & f(r) = 1 - \frac{2m}{r} \equiv -(\frac{2m}{T} - 1) \\ \Rightarrow \quad & ds^2 = -(\frac{2m}{T} - 1)^{-1} dT^2 + (\frac{2m}{T} - 1) dR^2 + T^2 d\Omega_2^2 \ . \end{aligned} \tag{27.1}$$

While this provides some minimal insight (e.g. that for $r < 2m$ surfaces of constant r are spacelike and that the metric looks time-dependent), what we are looking for is a way of describing the physics of the Schwarzschild solution that encompasses both the region $r > 2m$ and the region $r < 2m$ (and that therefore e.g. provides a valid continuous map for the freely falling observer as he crosses $r = 2m$).

The way we will go about this is to use either the worldlines of ingoing timelike geodesic observers (along with their proper time) or those of ingoing lightrays (along with an affine parameter along them), to provide us with coordinates in the region $r < 2m$ (recall that in both cases $r = 2m$ lies at a finite value of that affine parameter so that this affine parameter is also a good coordinate beyond $r = 2m$).

We will explore the 1st option in sections 27.2 and 27.3 and the 2nd option in section 27.4 and subsequent sections. The latter, based on null geodesics, is technically somewhat simpler than the former (and is therefore also the approach commonly adopted in the literature), and will lead us rather painlessly and quickly to the maximal analytic extension of the Schwarzschild geometry in section 27.8.

I have also included the former option here, based on timelike goedesics, precisely because the resulting coordinate systems, which are quite interesting in their own right and are useful for certain more advanced applications, are usually not dealt with in any detail in the standard textbooks (and I therefore had to work out many of these details myself at some point). However, it is possible to skip sections 27.2 and 27.3 and go directly to section 27.4 and continue from there (and this is what I usually do in a first course on General Relativity).

27.2 Crossing r_s with Painlevé-Gullstrand Coordinates

We had seen that the Schwarzschild time coordinate t is adapted to static observers (whose proper time is proportional to t), and therefore not useful for describing the region $r < 2m$. We had also seen that freely falling observers cross $r = 2m$ in finite proper time τ. This suggests to choose some family of freely falling observers (geodesics) and to use their proper time $\tau = T(t, r, \theta, \phi)$ as the new time coordinate, i.e. to perform the coordinate transformation

$$(t, r, \theta, \phi) \to (T, r, \theta, \phi) \tag{27.2}$$

(retaining, for the time being, the coordinates (r, θ, ϕ)).

A natural (and the simplest) choice is to consider the family of freely falling observers which fall radially (angular momentum $L = 0$) and start off at rest from infinity (energy $E = 1$). In this case, the geodesic equations (26.20) and (26.21) take the form

$$\dot{r} = -(2m/r)^{1/2} \quad , \quad \dot{t} = (1 - 2m/r)^{-1} \ , \tag{27.3}$$

which also imply that

$$dt/dr = -(r/2m)^{1/2}(1 - 2m/r)^{-1} \ . \tag{27.4}$$

The solutions for $r = r(\tau)$ and $t = t(r)$ are

$$\begin{aligned} r(\tau) &= [3\sqrt{2m}(\tau_0 - \tau)/2]^{2/3} \quad \Leftrightarrow \quad \tau - \tau_0 = -(2/3)(2m)^{-1/2} r^{3/2} \\ t(r) &= t_0 - (2/3)(2m)^{-1/2} r^{3/2} - \Theta(r) \ , \end{aligned} \tag{27.5}$$

where

$$\Theta(r) = 2m \left(2(r/2m)^{1/2} + \ln \left| \frac{(r/2m)^{1/2} - 1}{(r/2m)^{1/2} + 1} \right| \right) \tag{27.6}$$

is the solution of

$$d\Theta(r)/dr = (2m/r)^{1/2}(1 - 2m/r)^{-1} \ . \tag{27.7}$$

In particular, we see that, up to the irrelevant constant $\tau_0 - t_0$, $\tau(r)$ and $t(r)$ are related by

$$\tau(r) = t(r) + \Theta(r) \quad . \tag{27.8}$$

We are thus led to introduce the new time coordinate $T(t,r)$ by

$$T(t,r) = t + \Theta(r) = t + 2(2mr)^{1/2} + 2m \ln \left| \frac{(r/2m)^{1/2} - 1}{(r/2m)^{1/2} + 1} \right| \quad . \tag{27.9}$$

Since the metric does not depend explicitly on t, to determine the metric in the coordinates (T, r, θ, ϕ) we only need to substitute dt by

$$dt = dT - d\Theta = dT - (r/2m)^{-1/2}(1 - 2m/r)^{-1} dr \quad , \tag{27.10}$$

or

$$dt = dT - \frac{\sqrt{r/2m}}{(r/2m) - 1} dr \quad . \tag{27.11}$$

Then one immediately finds the simple result

$$\begin{aligned} ds^2 &= -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \\ &= -f(r)dT^2 + 2\sqrt{2m/r}\, dT dr + (dr^2 + r^2 d\Omega^2) \\ &= -dT^2 + \left(dr + \sqrt{2m/r}\, dT\right)^2 + r^2 d\Omega^2 \quad . \end{aligned} \tag{27.12}$$

This is the Schwarzschild metric in *Painlevé-Gullstrand Coordinates* (Painlevé (1921), Gullstrand (1922)), abbreviated to PG coordinates in the following, and this form of the metric already reveals a number of important properties of the Schwarzschild geometry.

Most importantly, we see that due to the non-singular off-diagonal term the metric in these coordinates is well-defined and non-degenerate for all $0 < r < \infty$, in particular at $r = 2m$. This is the definitive proof that the singularity at $r = 2m$ in Schwarzschild coordinates is really just a coordinate singularity.

REMARKS:

1. The metric has the characteristic property that the metric induced on the slices of constant T is just the flat Euclidean metric for any T,

$$ds^2|_{T=T_1} = dr^2 + r^2 d\Omega^2 = d\vec{x}^2 \quad . \tag{27.13}$$

   In particular, even though the radial coordinate distance of the coordinate r is most certainly not the proper radial distance at constant Schwarzschild time t, $ds = f(r)^{-1/2} dr \neq dr$, the above form of the metric shows that the coordinate r has the property that it does measure the proper radial distance on surfaces of constant T. In this sense, r appears to be (perhaps somewhat surprisingly) adapted to freely falling observers with $E = 1$.

In any case, this makes this form of the metric (and choice of time coordinate) particularly convenient e.g. for the canoncial quantisation of fields in the Schwarzschild space-time, and it is also for this reason that this coordinate system has become increasingly popular in recent years.[81]

2. It is easy to verify / confirm directly in the PG coordinates that the new time-coordinate T really has the interpretation as measuring the proper time τ of radially freely falling observers starting off at rest at infinity. To that end note that in PG coordinates a radial timelike geodesic satisfies

$$-\dot{T}^2 + \left(\dot{r} + \sqrt{2m/r}\dot{T}\right)^2 = -1 \ , \tag{27.14}$$

while the conserved energy associated with the T-translation invariance of the metric has the form

$$E = f\dot{T} - \sqrt{2m/r}\dot{r} \ . \tag{27.15}$$

It immediately follows that

$$T = \tau \quad \Rightarrow \quad \dot{T} = +1 \quad \Rightarrow \quad \dot{r} = -\sqrt{2m/r} \quad \Rightarrow \quad E = 1 \ . \tag{27.16}$$

These are precisely the geodesic paths with which we began the construction.

The coordinate transformation (27.9) is of the general form $T(t,r) = t + \psi(r)$ (24.2) discussed previously, and preserves the t-independence and manifest spherical symmetry of the metric. It leads to the metric in the form anticipated in (24.8).

Turning this around, once one has found the Schwarzschild metric in Schwarzschild coordinates, say, one can of course "discover" an infinite number of new coordinate systems for the Schwarzschild metric by performing such a (or a more general) coordinate transformation. Only in special cases, however, will one find a coordinate system that is actually useful. Let us see how to recover the PG coordinate system in this way, and how to "precover" another coordinate system that we will discuss in more detail in section 27.4.

Starting with the Schwarzschild metric in Schwarzschild coordinates, and performing the coordinate transformation $t \to T(t,r) = t + \psi(r)$, one finds the metric

$$\begin{aligned} ds^2 &= -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \\ &= -f(r)(dT - \psi'(r)dr)^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \\ &= -f(r)dT^2 + 2f(r)\psi'(r)dTdr + f(r)^{-1}(1 - f(r)^2(\psi'(r))^2)dr^2 + r^2 d\Omega^2 \\ &\equiv -f(r)dT^2 + 2C(r)dTdr + f(r)^{-1}(1 - C(r)^2)dr^2 + r^2 d\Omega^2 \ , \end{aligned} \tag{27.17}$$

[81] See e.g. P. Kraus, F. Wilczek, *A Simple Stationary Line Element for the Schwarzschild Geometry, and Some Applications*, `arXiv:gr-qc/9406042` and citations thereto; A. Nielsen, M. Visser, *Production and decay of evolving horizons*, `arXiv:gr-qc/0510083` for PG-like metrics for general time-dependent spherically symmetric metrics; C. Barcelo, S. Liberati, M. Visser, *Analogue Gravity*, `arXiv:gr-qc/0505065` for a detailed discussion of PG-like metrics in the context of so-called analogue models of gravity.

where

$$C(r) = f(r)\psi'(r) \tag{27.18}$$

is an essentially arbitrary function of r. This metric represents the Schwarzschild metric for any choice of $C(r)$. In particular, the vacuum Einstein equations do not impose any constraints on $C(r)$ - this can be checked explicitly, but it would be silly to do so since we have just seen explicitly that $C(r)$ is not determined and just corresponds to the freedom of performing a particular class of coordinate transformations. One can therefore now choose $C(r)$ at will.

1. One natural choice is

   $$g_{Tr} = 0 \quad \Leftrightarrow \quad C(r) = 0 \ . \tag{27.19}$$

   This makes the metric diagonal, corresponds to $\psi(r)$ constant, and evidently returns one to Schwarzschild coordinates. Any other choice of $C(r)$ will lead to a non-diagonal metric in the coordinates (T, r).

2. Another attractive choice is

   $$\begin{aligned} g_{rr} = 1 \quad \Leftrightarrow \quad ds^2|_{T=T_0} = dr^2 + r^2 d\Omega^2 \quad &\Leftrightarrow \quad 1 - C(r)^2 = f(r) \\ &\Leftrightarrow \quad C(r) = \pm\sqrt{2m/r} \end{aligned} \tag{27.20}$$

   We see that with the upper sign this is precisely the choice giving rise to Painlevé-Gullstrand coordinates introduced at the beginning of this section, confirmed by the fact that for $\psi(r)$ this implies the differential equation

   $$C(r) = +\sqrt{2m/r} \quad \Rightarrow \quad \psi'(r) = \sqrt{2m/r}f(r)^{-1} \quad \Rightarrow \quad \psi(r) = \Theta(r) \ . \tag{27.21}$$

   The same argument shows that the lower sign gives a corresponding set of PG coordinates based on outgoing rather than on ingoing geodesics.

3. Yet another appealing choice is

   $$g_{rr} = 0 \quad \Leftrightarrow \quad C(r) = \pm 1 \quad \Leftrightarrow \quad \psi'(r) = \pm f(r)^{-1} \ , \tag{27.22}$$

   so that the metric has the particularly simple (and non-singular at $r = 2m$) form

   $$ds^2 = -f(r)dT^2 \pm 2dTdr + r^2 d\Omega^2 \ . \tag{27.23}$$

   Referring back to the discussion of section 26.7, in particular (26.91), we see that the solution to this equation is

   $$\psi(r) = \pm r^* \quad \Rightarrow \quad T(t, r) = t \pm r^* \ . \tag{27.24}$$

   These are just the advanced and retarded time coordinates (u, v) of section 26.8, and we will encounter them in section 27.4 as Eddington-Finkelstein coordinates.

4. Finally, the Kerr-Schild coordinates to be discussed in section 27.6 can be obtained by the simple choice

$$C(r) = 1 - f(r) \quad \Rightarrow \quad T(t,r) = t + r^* - r = t + 2m\log(r/2m - 1) \ \ . \tag{27.25}$$

This procedure can in principle be applied to any static, spherically symmetric metric, and we will also make use of it later, e.g. in section 39.2 when constructing coordinates for de Sitter space.

A geometrically more satisfactory construction of such adapated coordinates can be based on the attempt to find a function T such that the 4-velocity of a family (congruence) of observers u^α is orthogonal to the slices of constant T, in the sense that $u_\alpha = -\partial_\alpha T$. The two conditions $u_\alpha u^\alpha = -1$ (for a timelike congruence, say) and $u_\alpha = -\partial_\alpha T$ imply that u^α is geodesic,

$$u^\alpha \nabla_\alpha u_\beta = u^\alpha \nabla_\beta u_\alpha = \tfrac{1}{2}\nabla_\beta (u^\alpha u_\alpha) = 0 \ \ , \tag{27.26}$$

(this is a special case of the general result established in (9.65) that a gradient vector field is geodesic iff it is of constant length), and that the metric component g^{TT} with respect to this new coordinate function T is

$$g^{TT} = g^{\alpha\beta}\partial_\alpha T \partial_\beta T = g^{\alpha\beta} u_\alpha u_\beta = -1 \ \ , \tag{27.27}$$

and the construction can proceed from there.[82]

Conversely, this kind of reasoning may allow one to discover the geometric interpretation of a coordinate system that one has selected through other criteria, e.g. via a convenient choice of the function $C(r)$ in (27.17). In particular, applied to PG coordinates, one can argue as follows:

1. Given the metric in PG coordinates, one seeks the interpretation of the gradient covector

$$u_\alpha = -\partial_\alpha T \ \ , \tag{27.28}$$

which is correctly normalised for a 4-velocity field, $u_\alpha u^\alpha = -1$, because $g^{TT} = -1$ in PG coordinates;

2. in Schwarzschild coordinates, u_α has the components

$$(-\partial_t T - \partial_r T, -\partial_\theta T, -\partial_\phi T) = (-1, -(r/2m)^{-1/2} f(r)^{-1}, 0, 0) \ \ ; \tag{27.29}$$

[82] See K. Martel, E. Poisson, *Regular coordinate systems for Schwarzschild and other spherical spacetimes*, `arXiv:gr-qc/0001069`, which also introduces generalised PG coordinates adapted to geodesic observers with $E > 1$, i.e. with non-zero velocity at infinity. Corresponding coordinates for $E < 1$, i.e. adapted to freely falling observers with maximal radius $r_i < \infty$, have been constructed by Gautreau and Hoffmann (see the references in this article).

3. thus its contravariant components give the 4-velocity field

$$u^\alpha = (f(r)^{-1}, -(r/2m)^{-1/2}, 0, 0) \quad , \tag{27.30}$$

and comparison with (27.3) shows that this is precisely the family of tangent vectors

$$u^\alpha = \dot{x}^\alpha = (\dot{t}, \dot{r}, \dot{\theta}, \dot{\phi}) \tag{27.31}$$

characterising the radial ingoing geodesics with $E = 1$.

Even though we now have a coordinate system that extends in a non-singular way across $r = 2m$, so that $r = 2m$ is not a true singularity, this does not mean that nothing interesting at all happens at that locus. Indeed, the (legitimate) static observers at large radii $r \gg 2m$ are still waiting for an explanation for their observations, described in sections 26.4 and 26.5. Simply telling them that their coordinates are no good near $r = 2m$ will hardly be considered by them to be a satisfactory explanation of what they observe.

In order to address this issue, we now look at the behaviour of (radial) lightrays in PG coordinates, characterised by

$$-f(r)dT^2 + 2\sqrt{2m/r}dTdr + dr^2 = 0 \quad . \tag{27.32}$$

Parametrising the lightrays by $r = r(T)$, with a prime denoting a T-derivative, $r' = dr/dT$, we can write this equation as

$$(r')^2 + 2\sqrt{2m/r}r' = f \quad \Leftrightarrow \quad (r' + \sqrt{2m/r})^2 = 1 \quad . \tag{27.33}$$

Thus this has the two solutions

$$r' + \sqrt{2m/r} = \pm 1 \quad \Leftrightarrow \quad r'_\pm = \pm 1 - \sqrt{2m/r} \quad , \tag{27.34}$$

and we can now understand the role that $r = 2m$ plays.

For $r > 2m$ one has

$$r > 2m : \quad r'_+ > 0 \quad \text{and} \quad r'_- < 0 \quad . \tag{27.35}$$

This is the usual situation. There is one outgoing direction along which $r = r(T)$ grows, and one ingoing direction along which $r(T)$ decreases. In particular, while remaining within the future lightcone one can choose to go to either larger or smaller values of r.

For $r < 2m$, on the other hand, one has

$$r < 2m : \quad r'_+ < 0 \quad \text{and} \quad r'_- < 0 \quad . \tag{27.36}$$

Thus along both directions lightrays (and therefore massive particles as well) must move to smaller values of r. In particular, *no observer, no lightray and no information can escape from the region $r < 2m$.* No wonder that the asymptotic static observers never "see" the infalling observer cross $r = 2m$. This new region is a future extension of the Schwarzschild "patch" $r > 2m$ of the space-time, uncovered by ingoing radial geodesics.

With the opposite choice of sign in (27.20), corresponding to outgoing rather than ingoing radial geodesics, one would instead have discovered a region $r < 2m$ in which $r'_{\pm} > 0$. This can therefore not possibly be the "same" region $r < 2m$, and indeed is a past extension of the Schwarzschild patch, uncovered by the back-tracking of outgoing radial geodesics.

While it is possible to further explore the consequences of all this in terms of the present PG coordinates, since we have been led to consider the structure of the lightcones, i.e. null geodesics, it turns out to be more convenient, also for the following, to discuss this in terms of coordinates that are adapted to lightrays rather than to the timelike geodesics. These are the Eddington-Finkelstein coordinates to be discussed in section 27.4 below.

27.3 Lemaître and Novikov Coordinates

Before turning to Eddington-Finkelstein coordinates, I want to briefly mention two more coordinate systems that are instructive and occasionally useful, and that combine one of the attractive feature of PG coordinates (namely that the time coordinate is physical proper time for suitable geodesic observers) with one of the useful features of Schwarzschild coordinates (namely that the metric is diagonal), albeit at the expense of a time-dependent coordinate transformation rendering also the metric time-dependent.

This will be accomplished by constructing *comoving coordinates* for the geodesic observers underlying the (generalised) PG coordinate system, where comoving means that these observers remain at fixed values of all the spatial coordinates and only evolve in time = proper time. Since radial geodesics already remain at fixed values of the angular coordinates (θ, ϕ), what this amounts to is to trade the coordinate r for another coordinate that simply labels the individual geodesics (and that is therefore, tautologically, guaranteed to also remain constant along such a geodesic). Such a label is provided by an integration constant appearing in the solution to the radial geodesic equation since (tauto-)logically such an integration constant is constant along the geodesic.

Let us first consider the case $E = 1$. In this case, the effective potential for radial motion gives us the radial equation

$$\dot{r}^2 = \frac{2m}{r} \quad , \tag{27.37}$$

with the ingoing solution (27.5),

$$\dot{r} = -\sqrt{2m/r} \quad \Rightarrow \quad r(\tau) = [3\sqrt{2m}(\tau_0 - \tau)/2]^{2/3} \quad . \tag{27.38}$$

We can thus immediately identify τ_0 as a candidate for a comoving radial coordinate. This label τ_0 can be interpreted as the proper time at which the geodesic ends up at

$r = 0$, $r(\tau_0) = 0$. Alternatively and preferably, if one is (understandably) reluctant to label geodesics by their behaviour at $r = 0$, τ_0 can be thought of as being related to the value r_0 of $r(\tau)$ at $\tau = 0$,

$$r_0 = r(\tau = 0) = [3\sqrt{2m}\tau_0/2]^{2/3} \quad . \tag{27.39}$$

We therefore introduce a new radial variable $\rho = \rho(r, \tau)$ by

$$r(\tau, \rho) = [3\sqrt{2m}(\rho - \tau)/2]^{2/3} \quad . \tag{27.40}$$

This satisfies

$$dr = \dot{r}d\tau + (\partial r/\partial\rho)d\rho = \dot{r}(d\tau - d\rho) = -\sqrt{2m/r}(d\tau - d\rho) \quad , \tag{27.41}$$

so that (denoting the PG proper time coordinate T now by τ) the Schwarzschild line element in PG form (27.12) turns into

$$\begin{aligned} ds^2 &= -d\tau^2 + \left(dr + \sqrt{2m/r}d\tau\right)^2 + r^2 d\Omega^2 \\ &= -d\tau^2 + \left(\sqrt{2m/r}d\rho\right)^2 + r^2 d\Omega^2 \\ &= -d\tau^2 + \frac{2m}{r(\tau,\rho)}d\rho^2 + r(\tau,\rho)^2 d\Omega^2 \quad , \end{aligned} \tag{27.42}$$

with $r(\tau, \rho)$ given explicitly by (27.40). This is the Schwarzschild metric in *Lemâitre Coordinates* (Lemaître (1938)).

<u>Remarks:</u>

1. These coordinates have a clear physical interpretation and also manifestly extend in a non-singular way across $r = 2m$, the Schwarzschild radius being located at the innocuous value
$$r = 2m \quad \Rightarrow \quad \rho - \tau = 4m/3 \tag{27.43}$$
of the Lemaître coordinates, to all $r > 0$.

2. This is the first (but will not be the last) time that we see that it can be useful to work with a radial coordinate, here $\rho = \rho(r, \tau)$, that is not equal to the standard (aereal radius) radial coordinate r.

3. For comparison purposes with the metric in Novikov coordinates to be discussed below, note that from the explicit change of variables (27.40) one has
$$\frac{2m}{r(\tau,\rho)} = (\partial_\tau r)^2 = (\partial_\rho r)^2 \quad , \tag{27.44}$$
so that the metric (27.42) can also be written as
$$\begin{aligned} ds^2 &= -d\tau^2 + \frac{2m}{r(\tau,\rho)}d\rho^2 + r(\tau,\rho)^2 d\Omega^2 \\ &= -d\tau^2 + (\partial_\rho r)^2 d\rho^2 + r(\tau,\rho)^2 d\Omega^2 \quad . \end{aligned} \tag{27.45}$$

This also shows that the form of the metric does not depend on the precise choice of integration constant used to label the geodesics, since it is manifestly invariant under transformations

$$\rho \to \sigma = F(\rho) \quad \Rightarrow \quad (\partial_\rho r)^2 d\rho^2 = (\partial_\sigma r)^2 d\sigma^2 \ . \tag{27.46}$$

4. Since

$$dr = 0 \quad \Rightarrow \quad d\tau = d\rho \ , \tag{27.47}$$

the metric induced on surfaces of constant $r = r_1$ is given by

$$ds^2|_{r=r_1} = -\left(1 - \frac{2m}{r_1(\tau,\rho)}\right) d\tau^2 + r_1(\tau,\rho)^2 d\Omega^2 \ . \tag{27.48}$$

Thus, as could have (partially) been anticipated from the Schwarzschild form of the metric, a hypersurface of constant $r = r_1$ is timelike for $r_1 > 2m$, null for $r_1 = 2m$, and spacelike for $r_1 < 2m$.

5. The surfaces of constant time $\tau = \tau_1$, on the other hand, are manifestly spacelike everywhere,

$$ds^2|_{\tau=\tau_1} = \frac{2m}{r(\tau_1,\rho)} d\rho^2 + r(\tau_1,\rho)^2 d\Omega^2 \ . \tag{27.49}$$

They run into the spacelike singularity at $r = 0$ at $\rho = \tau_1$.

6. Since from (27.41) one sees that at constant $\tau = \tau_1$ one has

$$\sqrt{\frac{2m}{r}} d\rho|_{\tau=\tau_1} = dr \ , \tag{27.50}$$

this spatial metric induced on surfaces of constant τ is again, as in the case of PG coordinates (27.13), just the Euclidean metric on $\mathbb{R}^3$,

$$ds^2|_{\tau=\tau_1} = dr^2 + r^2 d\Omega^2 \ . \tag{27.51}$$

This should not come as a surprise since the PG coordinate T is equal to the Lemaître coordinate τ.

7. In these coordinates, the volume element $\sqrt{g}$ has the simple form

$$\sqrt{g} = \sqrt{2m} r^{3/2} \sin\theta = 3m(\rho - \tau) \sin\theta \ . \tag{27.52}$$

8. Even though the metric is explicitly time-dependent in Lemaître coordinates, the time-translation invariance of the Schwarzschild metric is still manifest in these coordinates. Indeed, since $r(\tau,\rho)$ only depends on the difference $\rho - \tau$, the metric is invariant under simultanous constant translations

$$(\tau,\rho) \to (\tau + c, \rho + c) \tag{27.53}$$

of τ and ρ. Equivalently, in Lemaître coordinates the Killing vector ∂_t takes the form

$$\partial_t = \partial_\tau + \partial_\rho \quad . \tag{27.54}$$

The norm of this Killing vector is

$$g_{\alpha\beta}(\partial_t)^\alpha(\partial_t)^\beta = -(1 - \frac{2m}{r}) \quad , \tag{27.55}$$

and thus ∂_t is timelike for $r(\tau, \rho) > 2m$, null for $r = 2m$, and spacelike for $0 < r < 2m$.

9. Note that $g_{\tau\tau} = -1$ and $g_{\tau\rho} = 0$ are automatic consequences of this construction: by definition, in comoving coordinates the velocity field has the form $u^\alpha = (1, 0, 0, 0)$, and $u_\alpha = -\partial_\alpha \tau$ implies

$$\begin{aligned} u^\alpha = -g^{\alpha\beta}\partial_\beta\tau = -g^{\alpha\tau} \stackrel{!}{=} (1,0,0,0) \quad &\Rightarrow \quad g^{\tau\tau} = -1 \quad , \quad g^{\tau\rho} = 0 \\ &\Rightarrow \quad g_{\tau\tau} = -1 \quad , \quad g_{\tau\rho} = 0 \quad . \end{aligned} \tag{27.56}$$

We will discuss comoving coordinates, and metrics employing them, in much more detail in the context of cosmology in sections 33-38, cf. in particular section 34.2 for geodesics of comoving observers in comoving coordinates,

10. Somewhat surprisingly, Lemaître coordinates appear not to have attracted a lot of attention or found widespread use. However, the fact that in Lemaître coordinates the metric extends all the way to $r = 0$, is diagonal, and is explicit (unlike the Novikov coordinates, briefly discussed below, or the Kruskal-Szekeres coordinates to be described in detail later on) again makes them an attractive coordinate system to use e.g. when studying quantum field theory in a Schwarzschild background.[83]

Closely related to Lemaître coordinates are the so-called *Novikov coordinates*, comoving coordinates based on radial geodesics with a finite maximal radius, $r_i < \infty$ (i.e. $E < 1$) in the notation of section 26.3. In this case, the equation to solve is

$$\dot{r}^2 = \frac{2m}{r} + 2E_{eff} = \frac{2m}{r} + E^2 - 1 \quad , \tag{27.57}$$

where E_{eff} or E depends on the choice of geodesic via a choice of integration constant (one could even choose E itself, say, as a comoving coordinate). Using r_i instead, related to the energy by

$$E(r_i)^2 = 1 - \frac{2m}{r_i} = f(r_i) = 1 + 2E_{eff}(r_i) \quad , \tag{27.58}$$

[83] See e.g. K. Melnikov, M. Weinstein, *On Unitary Evolution of a Massless Scalar Field In A Schwarzschild Background: Hawking Radiation and the Information Paradox*, `arXiv:hep-th/0205223`.

the trajectories are implicitly defined by (26.28) and (26.29), i.e.

$$\begin{aligned} r(\eta) &= \tfrac{1}{2} r_i \,(1+\cos\eta) \\ \tau(\eta) &= \left(\frac{r_i^3}{8m}\right)^{1/2} (\eta + \sin\eta) \quad . \end{aligned} \tag{27.59}$$

We now think of these relations, together with the usual relation $E = f\dot{t}$ (26.21) as defining a change of variables

$$r = r(\tau, r_i) \quad , \quad t = t(\tau, r_i) \quad . \tag{27.60}$$

Note that r_i is clearly a comoving coordinate, as it can be used to label the geodesic. Note also that, if required/desired, from (27.59) one can solve for $\tau = \tau(r, r_i)$,

$$\begin{aligned} &\eta = \arccos(2r/r_i - 1) \\ \Rightarrow\quad &\tau(r, r_i) = \left(\frac{r_i^3}{8m}\right)^{1/2} \left(\arccos(2r/r_i - 1) + 2\sqrt{(r/r_i) - (r/r_i)^2}\right) \quad . \end{aligned} \tag{27.61}$$

However, even without having to solve or invert these implicit equation, it is easy to see that the resulting metric will have the form

$$ds^2 = -d\tau^2 + f(r_i)^{-1} r'(\tau, r_i)^2 dr_i^2 + r(\tau, r_i)^2 d\Omega^2 \quad , \tag{27.62}$$

where $r'(\tau, r_i) = \partial r(\tau, r_i)/\partial r_i$. This is the metric in *Novikov Coordinates* (Novikov, 1964).

In order to establish (27.62), we proceed as follows:

- Writing
$$dt = \dot{t}d\tau + t'dr_i \quad , \quad dr = \dot{r}d\tau + r'dr_i \quad , \tag{27.63}$$
and plugging this into the Schwarzschild metric, one finds
$$ds^2 = -(f\dot{t}^2 - f^{-1}\dot{r}^2)d\tau^2 + 2(f^{-1}\dot{r}r' - f\dot{t}t')d\tau dr_i + (f^{-1}(r')^2 - f(t')^2)dr_i^2 \quad . \tag{27.64}$$

- The first term in brackets is equal to 1 because $(t(\tau), r(\tau))$ parametrise a timelike radial geodesic,
$$\text{timelike radial geodesic} \quad \Rightarrow \quad -f(r)\dot{t}^2 + f(r)^{-1}\dot{r}^2 = -1 \quad . \tag{27.65}$$
Thus $g_{\tau\tau} = -1$, as expected for comoving coordinates.

- Moreover, for comoving coordinates the second term in brackets, i.e. the off-diagonal term, is zero, $g_{\tau r_i} = 0$ (27.56). Thus we can deduce
$$g_{\tau r_i} = 0 \quad \Rightarrow \quad f^{-1}\dot{r}r' - f\dot{t}t' = 0 \quad . \tag{27.66}$$

Using

$$\dot{r} = -(f_i - f)^{1/2} \quad , \quad f\dot{t} = E = f_i^{1/2} \qquad (f_i \equiv f(r_i)) \ , \tag{27.67}$$

this allows us to eliminate t' in favour of r',

$$(t')^2 = f^{-2}(1 - f/f_i)(r')^2 \ . \tag{27.68}$$

- This implies for the third term in brackets that

$$g_{r_i r_i} = f^{-1}(r')^2 - f(t')^2 = f_i^{-1}(r')^2 \ , \tag{27.69}$$

leading to the result given in (27.62).

REMARKS:

1. Note that, just as the Lemaître metric in (27.45), the form of the Novikov metric does not depend on the precise choice of integration constants used to label the geodesics, as the term $(\partial_{r_i} r)^2 (dr_i)^2$ is invariant under transformations

$$r_i \to \rho = F(r_i) \ . \tag{27.70}$$

Both metrics can then uniformly be presented as

$$ds^2 = -d\tau^2 + \frac{(\partial_\rho r)^2}{1 + 2E_{eff}(\rho)} d\rho^2 + r(\tau, \rho)^2 d\Omega^2 \ , \tag{27.71}$$

where $r(\tau, \rho)$ is the solution to the radial equation of motion

$$\left(\frac{\partial r}{\partial \tau}\right)^2 = \frac{2m}{r} + 2E_{eff}(\rho) \ . \tag{27.72}$$

For $E_{eff}(\rho) = 0$ this reduces to the metric in Lemaître coordinates, for $E_{eff} < 0$ and $\rho = r_i$ this reduces to the above metric in Novikov coordinates, and one can likewise consider Novikov coordinates based on radial geodesics with $E_{eff} > 0$, i.e. with a non-zero velocity at infinity.

2. Usually, the metric is expressed not in terms of the variable r_i but in terms of

$$R = (r_i/2m - 1)^{1/2} \quad \Leftrightarrow \quad r_i = 2m(R^2 + 1) \quad \Leftrightarrow \quad E^2 = f_i = \frac{R^2}{R^2 + 1} \tag{27.73}$$

so that the metric takes the form

$$ds^2 = -d\tau^2 + \frac{R^2 + 1}{R^2} \left(\frac{\partial r}{\partial R}\right)^2 dR^2 + r(\tau, R)^2 d\Omega^2 \ . \tag{27.74}$$

In terms of R, the relation (27.61) can also, using

$$\arccos(2r/r_i - 1) = 2\arccos(r/2m(R^2 + 1))^{1/2} \ , \tag{27.75}$$

which in turn can be established by using the trigonometric identity

$$\cos 2\varphi = 2\cos^2\varphi - 1 \ , \tag{27.76}$$

be written as

$$\frac{\tau(r,R)}{2m} = (R^2+1)\left[\frac{r}{2m} - \frac{(r/2m)^2}{R^2+1}\right]^{1/2} + (R^2+1)^{3/2}\arccos\left[\left(\frac{r/2m}{R^2+1}\right)^{1/2}\right] \ . \tag{27.77}$$

This expression is occasionally found and used in the literature.

3. The fact that $r(\tau, r_i)$ or $r(\tau, R)$ is only determined implicitly makes Novikov coordinates somewhat more awakward to use in practice than Lemaître coordinates.[84] Nevertheless, this is compensated by their clear physical interpretation. In particular, they are useful e.g. in numerical simulations and other investigations of gravitational collapse, and the Novikov metric will indeed naturally arise in this context in our brief discussion of gravitational collapse in section 29, in particular first in section 29.3.

4. Since the timelike geodesics that Novikov coordinates are based on oscillate in finite proper time from $r = 0$ in the past through the maximal radius r_i to $r = 0$ in the future, Novikov coordinates cover both the past and future extensions of the metric in the Schwarzschild patch discovered in terms of PG coordinates in section 27.2, and to be discussed again below in section 27.5. The reflection symmetry $R \to -R$ of the metric also exchanges the Schwarzschild patch and its "mirror region" (first encountered in the context of isotropic coordinates at the end of section 24.5, and to be discussed in detail in section 27.8), and thus Novikov coordinates actually turn out to provide a complete covering of the fully extended Kruskal-Schwarzschild space-time.[85]

27.4 Eddington-Finkelstein Coordinates I: Regularity of the Metric

In previous sections we have used the worldlines of freely falling observers and in particular the proper time along such worldlines, to construct (in various ways) coordinates in which the Schwarzschild metric extends in a non-singular way across $r = 2m$.

Instead of singling out some class of timelike observers in order to construct coordinates, it is at least equally (if not more) natural to introduce coordinates that are adapted to null geodesics.

[84] See e.g. the appendix of J. Makela, A. Peltola, *Thermodynamical Properties of Horizons*, `arXiv:gr-qc/0205128` for somewhat more explicit expressions, including a Taylor expansion of r'.

[85] See e.g. Figure 31.2 of C. Misner, K. Thorne, J. Wheeler, *Gravitation*.

We can easily accomplish this by promoting the integration constants $C_\pm$ in (26.93) labelling the lightray to new coordinates, namely the retarded or advanced time coordinate

$$C_+ \to u = t - r^* \quad , \quad C_- \to v = t + r^* \quad , \tag{27.78}$$

where

$$r^* = r + 2m \log |r/2m - 1| \tag{27.79}$$

is the solution of

$$\frac{dr^*}{dr} = f(r)^{-1} \quad . \tag{27.80}$$

Then ingoing radial null geodesics ($dr^*/dt = -1$) are characterised by $v = const.$ and outgoing radial null geodesics by $u = const.$ (and u and v can be thought of as "comoving" coordinates for outgoing resp. ingoing lightrays).

Then we can label space-time points (in addition to by their angular coordinates) e.g. in terms of ingoing lightrays by specifying the lightray (i.e. v) and the affine parameter λ indicating how far one has to travel along that lightray to reach the point. As we had seen explicitly e.g. in (26.82), this affine parameter can conveniently be chosen to be the radial coordinate r itself, and evidently extends across $r = 2m$.

Thus we now pass to the ingoing or outgoing *Eddington-Finkelstein* coordinates (v, r, θ, ϕ) or (u, r, θ, ϕ) (i.e. we keep r but eliminate t). Note that this coordinate transformation $u(t, r) = t - r^*$ or $v(t, r) = t + r^*$ is also of the general form $T(t, r) = t + \psi(r)$ whose consequences were already explored in section 27.2. Since the metric does not depend explicitly on t, we only need to substitute dt using

$$dt = dv - dr^* = dv - dr/f(r) \tag{27.81}$$

(and likewise for u). It follows that in terms of these coordinates the Schwarzschild metric reads

$$ds^2 = -(1 - 2m/r)dv^2 + 2dv\, dr + r^2 d\Omega^2 \tag{27.82}$$

in ingoing (advanced) Eddington-Finkelstein coordinates, and

$$ds^2 = -(1 - 2m/r)du^2 - 2du\, dr + r^2 d\Omega^2 \tag{27.83}$$

in outgoing (retarded) Eddington-Finkelstein coordinates, as already anticipated in (27.23). This way of writing the Schwarzschild metric turns out to be extremely informative and useful. We will focus on the metric and its regularity in this section, and discuss the causal structure (lightcones) and its implications in section 27.5 below.

Remarks:

1. Even though the metric coefficent g_{uu} or g_{vv} vanishes at $r = 2m$, as for the PG coordinates of section 27.2 there is no real degeneracy, the two-dimensional

metric in the (v,r)- or (u,r)-directions having the completely non-singular and non-degenerate form

$$(g_{(v,r),(v,r)}) = \begin{pmatrix} -f(r) & \pm 1 \\ \pm 1 & 0 \end{pmatrix} \tag{27.84}$$

Indeed, the determinant and inverse of this (2×2)-block of the metric are

$$|\det(g_{(v,r),(v,r)})| = 1 \tag{27.85}$$

and

$$(g^{(v,r),(v,r)}) = \begin{pmatrix} 0 & \pm 1 \\ \pm 1 & f(r) \end{pmatrix} \quad , \tag{27.86}$$

both of which are completely regular for all $r > 0$, in particular at $r = 2m$.

Therefore we can now extend the range of r to the region $0 < r < 2m$ with impunity. Thus this provides another explicit proof that the singularity of the Schwarzschild metric in the Schwarzschild coordinates at $r = 2m$ is a removable (pure coordinate) singularity.

2. Given the Eddington-Finkelstein coordinates, in which the metric is regular at and across $r = 2m$, it is now instructive to revisit some of the preliminary calculations that we performed in section 26 while exploring the Schwarzschild metric in the region $r \gtrsim 2m$.

   In particular, let us reconsider the calculation of the acceleration of a static observer (section 26.2). In Schwarzschild (SS) coordinates (which for the purposes of this remark we will denote by y^μ, reserving x^α for the Eddington-Finkelstein (EF) coordinates), the acceleration vector of a static oberver with 4-velocity (26.6)

$$(u^\mu)_{SS} = (u^t, u^r, u^\theta, u^\phi) = (f(r)^{-1/2}, 0, 0, 0) \tag{27.87}$$

   was found to be (26.10)

$$(a^\mu)_{SS} = (a^t, a^r, a^\theta, a^\phi) = (0, m/r^2, 0, 0) \quad , \tag{27.88}$$

   which looks nicely regular and unspectacularly "Newtonian". This was misleading, however, as the norm of the acceleration vector involves the component $g_{rr} = f(r)^{-1}$ of the metric, leading to

$$g_{\mu\nu}(a^\mu)_{SS}(a^\nu)_{SS} = f(r)^{-1}(m/r^2)^2 \quad , \tag{27.89}$$

   which diverges as $r \to 2m$.

   Let us now look at how this calculation presents itself in EF coordinates in which, as we have seen, the metric is regular at $r = 2m$. In EF coordinates, a static observer has 4-velocity

$$(u^\alpha)_{EF} = (u^v, u^r, u^\theta, u^\phi) = (u^v, 0, 0, 0) \quad , \tag{27.90}$$

with

$$g_{\alpha\beta}(u^\alpha)_{EF}(u^\beta)_{EF} = -f(r)(u^v)^2 = -1 \quad \Rightarrow \quad u^v = f(r)^{-1/2} \ . \tag{27.91}$$

Thus

$$(u^\alpha)_{EF} = (f(r)^{-1/2}, 0, 0, 0) \ . \tag{27.92}$$

This agrees with the result in SS coordinates, as could have also been deduced from the vectorial transformation behaviour

$$(u^v)_{EF} = \frac{\partial v}{\partial y^\mu}(u^\mu)_{SS} = \frac{\partial v}{\partial t}(u^t)_{SS} = (u^t)_{SS} \ . \tag{27.93}$$

under the coordinate transformation

$$dv = dt + dr_* = dt + dr/f(r) \quad \Rightarrow \quad \frac{\partial v}{\partial t} = 1 \quad , \quad \frac{\partial v}{\partial r} = f(r)^{-1} \ . \tag{27.94}$$

The acceleration vector can now be determined

- either by calculating

$$a^\alpha = \Gamma^\alpha_{\beta\gamma} u^\beta u^\gamma = \Gamma^\alpha_{vv}(u^v)^2 = f(r)^{-1}\Gamma^\alpha_{vv} \ . \tag{27.95}$$

  in EF coordinates (Exercise!),

- or by transforming $(a^\mu)_{SS}$ to EF coordinates,

$$(a^\alpha)_{EF} = \frac{\partial x^\alpha}{\partial y^\mu}(a^\mu)_{SS} = \frac{\partial x^\alpha}{\partial r}(a^r)_{SS} \quad \Rightarrow \quad \begin{cases} (a^v)_{EF} = f(r)^{-1} m/r^2 \\ (a^r)_{EF} = m/r^2 \end{cases} \tag{27.96}$$

Either way, the result is

$$(a^\alpha)_{EF} = (a^v = f(r)^{-1} m/r^2, a^r = m/r^2, 0, 0) \ . \tag{27.97}$$

Thus the (non-singular, "Newtonian") r-component agrees with that of the acceleration in SS coordinates, but in addition in EF coordinates there is now a v-component which is singular as $r \to 2m$.

Thus the norm of the acceleration in EF coordinates is

$$g_{\alpha\beta}(a^\alpha)_{EF}(a^\beta)_{EF} = -f(r)(a^v)^2 + 2a^v a^r = f(r)^{-1}(m/r^2)^2 \ , \tag{27.98}$$

in complete agreement with the result in SS coordinates (as it should, since this is a scalar).

We see that, in these coordinates in which the metric is regular, the divergence of the acceleration shows up in the components of that quantity itself (whereas in the "singular" SS coordinates that divergence was misleadingly hidden in the metric).

3. One can of course check directly that the Eddington-Finkelstein metric (27.82) is a solution of the vacuum Einstein equations. One can also obtain the metric directly from integrating the Einstein vacuum equations if one starts not with the standard form of a static isotropic metric (24.6) but makes an ansatz of the form

$$ds^2 = -f(v,x)dv^2 + 2dv\,dx + r(v,x)^2 d\Omega^2 \tag{27.99}$$

in terms of two unknown functions $f(v,x)$ and $r(v,x)$ of two variables. The same arguments as in the discussion of ingoing null geodesics in section 27.5 below show that this general form of the metric is characterised by the fact that that the integral curves of the null vector ∂_x (i.e. the curves with constant (v,θ,ϕ)) are null geodesics, with affine parameter x,

$$g_{xx} = 0 \quad , \quad \nabla_{\partial_x}\partial_x = 0 \ . \tag{27.100}$$

As one can always choose to build a coordinate system in such a way, this is a valid a priori ansatz for the metric, and from this perspective one never encounters the question if the singularity at $r = 2m$ is real or not, since it is not even a coordinate singularity in these coordinates.

Alternatively, one can start with the metric in the Bondi gauge (40.30),

$$ds^2 = -\mathrm{e}^{2h(v,r)} f(v,r)dv^2 + 2\mathrm{e}^{h(v,r)} dv\,dr + r^2 d\Omega^2 \ . \tag{27.101}$$

This is tantamount to introducing the aereal radius $r(v,x)$ as a new coordinate instead of x, and this ansatz is the Eddington-Finkelstein-like counterpart of the Schwarzschild-Birkhoff ansatz (24.68). It is shown in section 40.3 that solving the vacuum Einstein equations in this gauge, one also obtains directly the Schwarzschild solution in Eddington-Finkelstein coordinates.

4. In the Eddington-Finkelstein coordinates, the Killing vector $\xi = \partial_t$ generating the time translation invariance of the metric in Schwarzschild coordinates can be written as

$$\xi = \partial_t = (\partial_t r)\partial_r + (\partial_t v)\partial_v = \partial_v \tag{27.102}$$

or

$$\xi = \partial_t = (\partial_t r)\partial_r + (\partial_t u)\partial_u = \partial_u \ . \tag{27.103}$$

That this is indeed a Killing vector is obvious from the fact that in Eddington-Finkelstein coordinates the components of the metric do not depend on v or u.

In particular, ξ now extends smoothly across $r = 2m$, with norm

$$g_{\alpha\beta}\xi^\alpha\xi^\beta = -(1 - 2m/r) \ . \tag{27.104}$$

Thus ξ is timelike for $r > 2m$, null on $r = 2m$ and spacelike for $0 < r < 2m$. This is a crucial and characteristic feature we will come back to on various occasions in subsequent sections.

27.5 Eddington-Finkelstein Coordinates II: Event Horizons and Black Holes

We now take a closer look at the causal structure of the extended Schwarzschild geometry in Eddington-Finkelstein coordinates. To determine the lightcones in ingoing Eddington-Finkelstein coordinates we again look at radial null geodesics which this time are solutions to

$$(1-2m/r)dv^2 = 2dv\,dr \quad . \tag{27.105}$$

There are thus two possibilities:

1. Ingoing null geodesics are described by $dv/dr = 0$ or $v = const.$. These null geodesics can also be understood in a different way. Note that from the Eddington-Finkelstein form (27.82) of the metric it is evident that ∂_r is a null vector, since $g_{rr} = 0$. Moreover, its integral curves, i.e. the curves

$$x^\alpha(\lambda) = (r = r_0 - \lambda, v = v_0, \theta = \theta_0, \phi = \phi_0) \tag{27.106}$$

   are geodesics. The sign $-\lambda$ has been chosen so that the tangent vector $\dot{x}^\alpha = dx^\alpha/d\lambda$ is future-oriented, i.e. such that its scalar product with $\partial_t = \partial_v$ is negative,

$$g_{\alpha\beta}\dot{x}^\alpha(\partial_t)^\beta = g_{\alpha v}\dot{x}^\alpha = \dot{r} \stackrel{!}{<} 0 \quad . \tag{27.107}$$

   Thus the radius indeed decreases along future-oriented ingoing null geodesics (as the name was meant to suggest) and the radial coordinate is an affine parameter along these geodesics (as we also already saw in (26.82)).

   That these curves are geodesics can be seen explicitly by calculating the acceleration of $u^\alpha = (0, -1, 0, 0)$ (in the coordinates (v, r, θ, ϕ)),

$$u^\alpha \nabla_\alpha u^\beta = -\nabla_r u^\beta = -\Gamma^\beta_{r\gamma}u^\gamma = \Gamma^\beta_{rr} \quad , \tag{27.108}$$

   or more concisely (see (5.21)) by

$$\nabla_{\partial_r}\partial_r = \Gamma^\beta_{rr}\partial_\beta \quad . \tag{27.109}$$

   Either way this shows that ∂_r is geodesic, as $\Gamma^\beta_{rr} = 0$ (since $g_{rr} = 0$, the only possible contribution to $\Gamma_{\beta rr}$ could have arisen from $\partial_r g_{vr}$, but $g_{vr} = 1$ is constant).

2. Outgoing null geodesics are described by the solutions to the equation

$$\frac{dv}{dr} = 2(1-2m/r)^{-1} \quad . \tag{27.110}$$

   The solution to this equation is evidently

$$v(r) = 2r^* + C \quad \Leftrightarrow \quad u = t - r^* = C \quad . \tag{27.111}$$

   We thus reassuringly recover the fact that outgoing lightrays are described by lines of constant u.

Thus the metric and the lightcones remain well-behaved (do not fold up) at $r = 2m$, the surface $r = 2m$ is at a finite coordinate distance, namely (to reiterate the obvious) at $r = 2m$, and there is no problem with following geodesics beyond $r = 2m$.

In particular, this means that we now encounter no difficulties when entering the region $r < 2m$, e.g. along lines of constant v and this region should be included as part of the physical space-time. Note that because $v = t + r^*$ and $r^* \to -\infty$ for $r \to 2m$, we see that decreasing r along lines of constant v amounts to $t \to \infty$. Thus the new region at $r \leq 2m$ we have discovered is in some sense a future extension of the original Schwarzschild space-time.

To understand the nature of this new region, we now take a more detailed look at the behaviour of the lightcones. Even though the lightcones do not fold up at $r = 2m$, something interesting is certainly happening there. Whereas, in a (v, r)-diagram (see Figure 18), one side of the lightcone always remains horizontal (at $v = const.$), the other side becomes vertical at $r = 2m$ ($dv/dr = \infty$) and then tilts over to the other side. In particular, beyond $r = 2m$ all future-directed paths, those within the forward lightcone, now have to move in the direction of decreasing r: clearly the ingoing null geodesics move towards smaller values of r, but so do those that for $r > 2m$ were outgoing,

$$(27.110) \quad \Rightarrow \quad dr/dv < 0 \quad \text{for} \quad r < 2m \ . \tag{27.112}$$

There is thus no way to turn back to larger values of r, not on a geodesic but also not on any other path (i.e. not even with a powerful rocket) once one has gone past $r = 2m$.

Thus, even though locally the physics at $r = 2m$ is well behaved, globally the surface $r = 2m$ is very significant as it is a point of no return:

- Since $r = 2m$ is a null surface, once one has reached the event horizon one has to travel at the speed of light to stay there and not be forced further towards smaller values of r.
- Once one has passed the Schwarzschild radius, there is no turning back to larger values of r.

Therefore nothing, absolutely nothing, no information, no lightray, no particle, can escape from the region $r < 2m$. Thus we have a *Black Hole*, an object that is (classically) completely invisible from the outside. In particular, no information about any events occurring in the black hole region $r < 2m$ can reach the asymptotic ($r >> 2m$) region. The boundary of this region is the surface $r = 2m$, and a (necessarily null) surface with this property is known as an *Event Horizon*.

Even though the Eddington-Finkelstein coordinates (more precisely the closely related Eddington time coordinate - cf. section 27.6 below) were already introduced by Eddington back in 1924, this full significance of the Schwarzschild radius and its interpretation as a "one-way membrane" were only understood much later (Finkelstein, 1958).

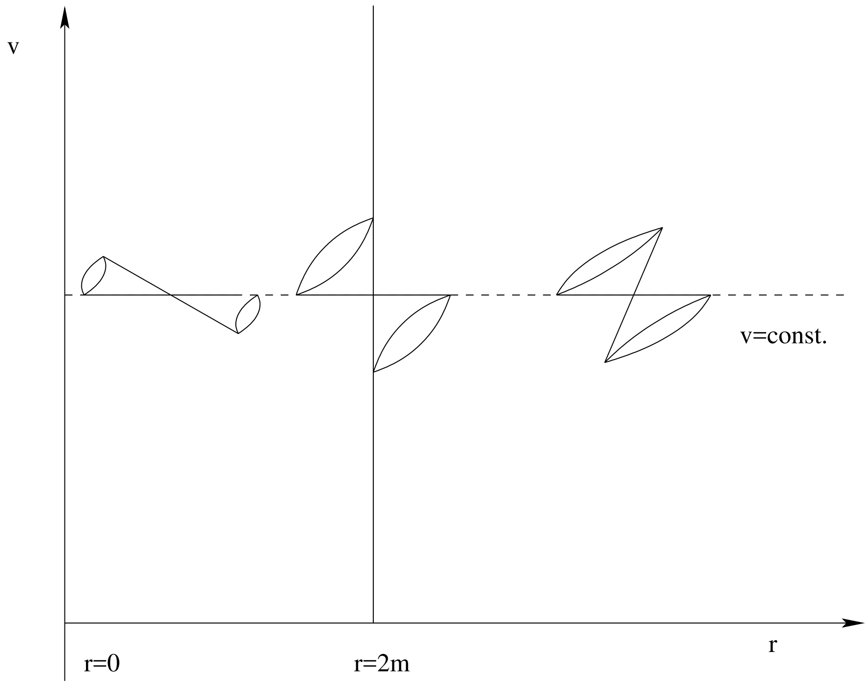


Figure 18: Behaviour of lightcones in ingoing Eddington-Finkelstein coordinates. Lightcones do not fold up at $r = 2m$ but tilt over so that for $r < 2m$ only movement in the direction of decreasing r towards the singularity at $r = 0$ is allowed.

<u>REMARKS:</u>

1. In the above (v, r) coordinate system we can cross the event horizon only on future directed paths, not on past directed ones, and only in the direction of decreasing r. However, clearly this cannot be the whole story: the Schwarzschild metric in the Schwarzschild coordinates is invariant under time-reflections $t \to -t$. Hence there must also exist a time-reversed version of the future extension and its event horizon. Noting that $t \to -t$ implies

$$t \to -t \quad \Rightarrow \quad v \to -u \quad , \quad dr/dt \to -dr/dt \ , \tag{27.113}$$

   it is clear that one will have access to this new region when working with the outgoing Eddington-Finkelstein coordinates (u, r) instead, and backtracking outgoing lightrays beyond $t = -\infty$.

   Indeed, when one uses the coordinates (u, r) instead of (v, r), the lightcones in Figure 18 are flipped (either up-down or left-right), and one can now pass through the horizon along future directed paths only in the outgoing direction of increasing r.

   The new region of space-time covered by the coordinates (u, r) is thus definitely different from the new region we uncovered using (v, r) even though both of them lie 'behind' $r = 2m$. In fact, this one is a past extension (beyond $t = -\infty$) of the original Schwarzschild 'patch' of space-time. In this patch, the region behind $r = 2m$ acts like the opposite (time-reversal) of a black hole (a white hole) which cannot be entered on any future-directed path.

As we will discuss later, in section 29.3, this white hole region is unphysical, i.e. an artefact of the idealisation of an eternal black hole metric. The future black hole region, however, is definitely of relevance as it can be created by gravitational collapse.

2. The above statement about the behaviour of lightrays for $r = r_s = 2m$ and $r < r_s$ can be phrased somewhat more invariantly and geometrically in terms of the *expansion* (12.101), (12.102)

$$\theta_\ell = \tfrac{1}{2} s^{\alpha\beta} L_\ell s_{\alpha\beta} = \frac{1}{\sqrt{s}} L_\ell \sqrt{s} \quad . \tag{27.114}$$

of a null geodesic congruence introduced in connection with the Raychaudhuri equation in section 12.4, and measuring the change in the cross-sectional area element $\sqrt{s}$ of the congruence along the congruence.

To that end let us introduce the null vector fields

$$n = -\partial_r \quad , \quad \ell = \partial_v + \tfrac{1}{2} f(r) \partial_r \quad . \tag{27.115}$$

It is easy to check that these are 2 linearly independent future-pointing null vector-fields, n corresponding to ingoing lightrays and ℓ to (would-be) outgoing lightrays, cross-normalised to $n_\alpha \ell^\alpha = -1$,

$$n_\alpha n^\alpha = \ell_\alpha \ell^\alpha = 0 \quad , \quad n_\alpha \ell^\alpha = -1 \quad . \tag{27.116}$$

With $\sqrt{s} = r^2 \sin\theta$, one finds from (27.114) that

$$\theta_n = -\frac{2}{r} \quad , \quad \theta_\ell = \frac{r f(r)}{r^2} = \frac{r - 2m}{r^2} \quad . \tag{27.117}$$

In particular, in Minkowski space one has the standard behaviour that the ingoing radial null congruence always has negative expansion (it is contracting) while the outgoing radial null congruence has positive expansion (it is expanding),

$$f(r) = 1 \quad \Rightarrow \quad \theta_n = -\frac{2}{r} < 0 \quad , \quad \theta_\ell = +\frac{1}{r} > 0 \quad . \tag{27.118}$$

While one still has $\theta_n < 0$ in the Schwarzschild geometry (ingoing lightrays are contracting),

$$f(r) = 1 - \frac{2m}{r} \quad \Rightarrow \quad \theta_n < 0 \quad \forall\, r \quad , \tag{27.119}$$

for the congruence ℓ one has

$$f(r) = 1 - \frac{2m}{r} \quad \Rightarrow \quad \begin{cases} \theta_\ell > 0 & r > r_s \\ \theta_\ell = 0 & r = r_s \\ \theta_\ell < 0 & r < r_s \end{cases} \tag{27.120}$$

Thus $r = r_s$ is characterised by the fact that the "outgoing" lightrays have zero expansion. We will come back to this, and the related notion of *trapped surfaces*, briefly in the context of the discussion of horizons of black holes in section 32.

Since we now have at our disposal

- the defining feature of a black hole, namely the existence of an event horizon which causally seals off the interior (black hole) region from the outside (we will discuss this in more general terms in section 32.4),
- and the prime example of a black hole gravitational field, namely that described by the future extension of the Schwarzschild metric,

let us briefly, to conclude this section, discuss (and equally briefly dispose of) several popular misconceptions about black holes that are unfortunately quite common in the pop-sci and sci-fi literature:

1. black holes as cosmic vacuum cleaners

   Often a black hole is pictured as an object travelling through space and in the process violently sucking in everything around it. This is quite misleading: at a coordinate distance r in the Schwarzschild metric, the gravitational field (i.e. the metric) is the same regardless of whether in the interior there is a star of mass $M = m/G_N$ or a black hole of the same mass (there is no extra "black hole force" that enters the scene, it is all just gravity).

   In particular, if our sun, say, suddenly collapsed to form a black hole (as we will discuss in section 29, this is the way black holes are believed to form in astrophysical processes, even though this will not be the fate of our sun), then quite a number of things will change for humanity but what will not change appreciably (if the mass is appropriately conserved in this process) is the outside gravitational field and thus e.g. the orbit of the earth.

2. "locking in" of lightrays by black holes as a consequence of extremely strong gravitational fields

   It is natural to think that the properties of black holes are due to extremely strong gravitational effects at or near the horizon which prevent even lightrays from escaping the gravitational field. This point of view is of course not completely incorrect, and is occasionally reinforced by a classical "derivation" of the Schwarzschild radius as the radius at which the escape velocity from an object of mass M with that radius equals the speed of light ($v^2/2 = G_N M/r$ and $v = c$ imply $r = 2G_N M/c^2 = 2m$; I hope I do not have to tell you that you should not take this "calculation" seriously). More to the point, perhaps, we had seen that static observers close to $r = 2m$ require very large accelerations to balance what these observers perceive to be the gravitational attraction.

Nevertheless, this is somewhat misleading since this gravitational attraction is observer-dependent (e.g. there is none for a freely falling observer) and thus does not provide us with an objective measure of the strength of the gravitational field. Objective information about the gravitational field is contained in the tidal forces, as encoded in the Riemann curvature tensor via the geodesic deviation equation discussed e.g. in section 8.4. For instance, the Kretschmann scalar (8.61) of the Schwarzschild metric is (27.163)

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} \sim \frac{m^2}{r^6} \quad . \tag{27.121}$$

This shows that the strength of the gravitational field in the Schwarzschild geometry, as measured by tidal forces, is $\sim m/r^3$ (as in the Newtonian theory). Near the horizon at $r = 2m$ this is

$$m/r^3 \sim 1/m^2 \tag{27.122}$$

and can thus be arbitrarily weak for sufficiently massive (and large) black holes. Related to this, we will estimate in section 29.3 the average density of an object with the size of its Schwarzschild radius and will see that for sufficiently massive objects (e.g. galaxies) this density can be as small as one likes.

The crucial point, as we will discuss in more detail in section 32, is that by definition a black hole is characterised in terms of global properties of space-time (in the sense of a region of space-time that is invisible to an asymptotic observer), and these are not necessarily unambiguously detectable by local observers. In particular, as we will see, event horizons can exist in (and emerge from) flat Minkowskian regions of space-time, in which there is definitely no trace of a strong gravitational field (at that time!).

27.6 Eddington Time Coordinate and Kerr-Schild Form of the Metric

Ordinary space-time diagrams are more familiar (and therefore more intuitive) than space-null diagrams such as the above (r, v)-diagram (in which, for example, in the asymptotically flat regime $r \to \infty$ the lightcone has slopes 0 and 2 rather than the usual $\pm 45°$ slopes ± 1). In the present case this can easily be rectified by introducing a new time-coordinate $\tilde{t}$ instead of v by the relation

$$v = t + r^* = \tilde{t} + r \quad , \tag{27.123}$$

i.e.

$$\tilde{t} = t + 2m\log(r/2m - 1) \quad . \tag{27.124}$$

This time coordinate is also known as the *Eddington time coordinate* (it was discovered and used first by Eddington, and then rediscovered by Finkelstein, but neither of them wrote down explicitly the null form of the metric that is now known as the Eddington-Finkelstein form of the metric and which we discussed above).

Initially the metric looks perhaps somewhat less illuminating in the $(\tilde{t}, r)$-coordinates,

$$ds^2 = -(1 - 2m/r)d\tilde{t}^2 + (4m/r)d\tilde{t}\,dr + (1 + 2m/r)dr^2 + r^2 d\Omega^2 \quad , \tag{27.125}$$

but the lightcones and the horizon now have the following simple and easy to visualise description, which follows from factorising the line element as

$$ds^2 = (d\tilde{t} + dr)(-(1 - 2m/r)d\tilde{t} + (1 + 2m/r)dr) + r^2 d\Omega^2 \quad , \tag{27.126}$$

namely:

- the ingoing side of the lightcone is described by

$$d\tilde{t} + dr = 0 \quad \Leftrightarrow \quad d\tilde{t}/dr = -1 \quad \Leftrightarrow \quad v = \text{const.} \tag{27.127}$$

  (as we already knew) and is thus at $-45°$ everywhere in a $(\tilde{t}, r)$-diagram.

- the "outgoing" side of the lightcone is described by

$$ds^2 = -(1-2m/r)d\tilde{t}+(1+2m/r)dr = 0 \quad \Leftrightarrow \quad d\tilde{t}/dr = (1+2m/r)/(1-2m/r) \quad . \tag{27.128}$$

  Therefore $d\tilde{t}/dr = +1$ for $r \to \infty$ (slope $+1$), so lightcones have the standard Minkowskian form there, $d\tilde{t}/dr \to \infty$ for $r \to 2m$, and $d\tilde{t}/dr \to -1$ for $r \to 0$.

- in particular, the horizon is (again) vertical in such a diagram, with the (would-be) outgoing side of the lightcone vertical and tangent to the horizon, while the lightcone degenerates as $r \to 0$.

Nice diagrams that you can find in many places depicting the collapse of a spherically symmetric star to a black hole and the formation of the horizon typically (either implicitly or explicitly) use these $(\tilde{t}, r)$-coordinates.

Remarks:

1. This coordinate transformation is once again of the general form $T(t, r) = t+\psi(r)$, whose effect on the Schwarzschild metric was studied in some detail in section 27.2. In particlar, the time coordinate $\tilde{t}$ (27.124) already appeared in (27.25).

2. Note that in these coordinates the metric can be split as

$$\begin{aligned} ds^2 &= -(1 - 2m/r)d\tilde{t}^2 + (4m/r)d\tilde{t}\,dr + (1 + 2m/r)dr^2 + r^2 d\Omega^2 \\ &= \left(-d\tilde{t}^2 + dr^2 + r^2 d\Omega^2\right) + \frac{2m}{r} dv(\tilde{t}, r)^2 \quad . \end{aligned} \tag{27.129}$$

Thus it has the curious property that in components it has the form

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{2m}{r}\partial_\alpha v \partial_\beta v \qquad (27.130)$$

where $\eta_{\alpha\beta}$ is the Minkowski metric (here written in spatial spherical coordinates) and $\partial_\alpha v$ is null with respect to the inverse Minkowski metric $\eta^{\alpha\beta}$,

$$\eta^{\alpha\beta}\partial_\alpha v \partial_\beta v = \eta^{vv} = 0 \qquad (27.131)$$

because $v = \tilde{t} + r$ is a null coordinate for the Minkowski metric $-d\tilde{t}^2 + dr^2 + (\ldots)$. This is the Schwarzschild metric in what is known as *Kerr-Schild form.* I will therefore also occasionally refer to $\tilde{t}$ as the Kerr-Schild time coordinate (but, as mentioned above, $\tilde{t}$ is commonly also known as the Eddington time coordinate).

3. By introducing standard inertial (Cartesian) coordinates $\tilde{x}^\alpha = (\tilde{t}, \vec{x})$ for Minkowski space, the metric can now also be written in the form

$$ds^2 = -d\tilde{t}^2 + d\vec{x}^2 + \frac{2m}{r}(\ell_\alpha d\tilde{x}^\alpha)^2 \qquad (27.132)$$

where $r^2 = \vec{x}^2$, and ℓ_α has the components

$$\ell_\alpha = \partial_\alpha(\tilde{t} + r) = (1, x^k/r) \quad . \qquad (27.133)$$

4. There is evidently also a corresponding Kerr-Schild form of the metric in outgoing Eddington-Finkelstein coordinates, now with

$$u = t - r^* = \tilde{t} - r \quad \Rightarrow \quad \tilde{t} = t - 2m\log(r/2m - 1) \quad , \qquad (27.134)$$

so this is not the same $\tilde{t}$ as before, namely

$$ds^2 = \left(-d\tilde{t}^2 + dr^2 + r^2 d\Omega^2\right) + \frac{2m}{r}du(\tilde{t}, r)^2 \quad . \qquad (27.135)$$

More generally, metrics of the form

$$g_{\alpha\beta} = \eta_{\alpha\beta} + f(x)N_\alpha N_\beta \quad \text{with} \quad \eta^{\alpha\beta}N_\alpha N_\beta = \eta_{\alpha\beta}N^\alpha N^\beta = 0 \qquad (27.136)$$

are known as *Kerr-Schild metrics.* This ansatz for the metric played an important role in the search for other exact solutions of the Einstein equations.

The form (27.136) implies that

1. $N^\alpha \equiv \eta^{\alpha\beta}N_\beta$ is also null with respect to $g_{\alpha\beta}$:

$$g_{\alpha\beta}N^\alpha N^\beta = \eta_{\alpha\beta}N^\alpha N^\beta + f(x)(N_\alpha N^\alpha)(N_\beta N^\beta) = 0 \qquad (27.137)$$

2. the inverse metric is given by

$$g^{\alpha\beta} = \eta^{\alpha\beta} - f(x)N^\alpha N^\beta \tag{27.138}$$

Indeed, for $g^{\alpha\beta}$ of this form one has

$$\begin{aligned} g^{\alpha\beta}g_{\beta\gamma} &= (\eta^{\alpha\beta} - f(x)N^\alpha N^\beta)(\eta_{\beta\gamma} + f(x)N_\beta N_\gamma) \\ &= \delta^\alpha_\gamma - f(x)N^\alpha N_\gamma + f(x)N^\alpha N_\gamma = \delta^\alpha_\gamma \quad . \end{aligned} \tag{27.139}$$

3. the contravariant components of N_α with respect to $g_{\alpha\beta}$ are the same as those with respect to $\eta_{\alpha\beta}$:

$$g^{\alpha\beta}N_\beta = \eta^{\alpha\beta}N_\beta - f(x)N^\alpha N^\beta N_\beta = \eta^{\alpha\beta}N_\beta \quad , \tag{27.140}$$

so that the notation N^α is unambiguous.

27.7 Kruskal-Szekeres Coordinates

Using Eddington-Finkelstein coordinates, we have discovered two "new" regions of the space-time. Are there still other regions of space-time to be discovered? The answer is yes. Not only is this suggested by the analogy with Rindler and Minkowski space (but this is after all just an analogy at the moment). We had actually already seen a "doubling" of the Schwarzschild patch in the discussion of the Schwarzschild metric in isotropic coordinates (see the comment at the end of section 24.5). One way to rediscover this region from the present perspective would be to study spacelike rather than null geodesics. Alternatively, let us try to short-cut this somewhat and let us guess how one might be able to describe the maximal extension of space-time.

The first guess might be to use the coordinates u and v simultaneously, instead of r and t. In these coordinates, the metric takes the form

$$ds^2 = -(1 - 2m/r)du\, dv + r^2 d\Omega^2 \quad , \tag{27.141}$$

with $r = r(u, v)$. While this is a good idea, the problem is that in these coordinates the horizon is once again infinitely far away, at $u = +\infty$ or $v = -\infty$ (i.e. at $2r^* = v - u = -\infty$). We can rectify this by introducing coordinates U and V with

$$U = -\mathrm{e}^{-u/4m} \quad , \quad V = \mathrm{e}^{v/4m} \tag{27.142}$$

say, so that the horizon is now at either $U = 0$ or $V = 0$. To better understand why we choose these coordinates and why exactly this factor in the exponent (and not any other positive number, which so far would have had the same effect of moving the horizon to $U = 0$ or $V = 0$), note that

$$\frac{v - u}{4m} = \frac{r}{2m} + \log\left(\frac{r}{2m} - 1\right) \quad , \tag{27.143}$$

so that the prefactor $f(r) = 1 - 2m/r$ can be written as

$$1 - \frac{2m}{r} = \frac{2m}{r}\left(\frac{r}{2m} - 1\right) = \frac{2m}{r}\mathrm{e}^{-r/2m}\mathrm{e}^{(v-u)/4m} \quad . \tag{27.144}$$

Thus the metric can then be written as

$$ds^2 = \frac{2m}{r}\mathrm{e}^{-r/2m}\left(\mathrm{e}^{v/4m}dv\right)\left(-\mathrm{e}^{-u/4m}du\right) + r(u,v)^2 d\Omega^2 \quad . \tag{27.145}$$

It is now clear why we made the choice (27.142). The metric now takes the simple form

$$ds^2 = -\frac{32m^3}{r}\mathrm{e}^{-r/2m}dUdV + r(U,V)^2 d\Omega^2 \quad , \tag{27.146}$$

with $r = r(U,V)$ given implicitly by

$$UV = -\mathrm{e}^{(v-u)/4m} = -\mathrm{e}^{r^*/2m} = -(r/2m - 1)\mathrm{e}^{r/2m} \tag{27.147}$$

and $t = t(U,V)$ explicitly by

$$U/V = -\mathrm{e}^{-(u+v)/4m} = -\mathrm{e}^{-t/2m} \quad . \tag{27.148}$$

Finally, we pass from the null coordinates (U,V) (meaning that ∂_U and ∂_V are null vectors) to more familiar timelike and spacelike coordinates (T,X) defined, in analogy with $(u,v) = t \mp r^*$, by

$$U = T - X \quad , \quad V = T + X \quad , \tag{27.149}$$

in terms of which the metric is

$$ds^2 = \frac{32m^3}{r}\mathrm{e}^{-r/2m}(-dT^2 + dX^2) + r^2 d\Omega^2 \quad . \tag{27.150}$$

Here $r = r(T,X)$ is now implicitly given by

$$X^2 - T^2 = (r/2m - 1)\mathrm{e}^{r/2m} \quad . \tag{27.151}$$

Chasing through the above sequence of coordinate transformations

$$(t,r) \to (t,r^*) \to (u,v) \to (U,V) \to (T,X) \quad , \tag{27.152}$$

one finds that the coordinate transformation $(t,r) \to (T,X)$ is explicitly, and in its full glory, given by

$$\begin{aligned} X(t,r) &= \tfrac{1}{2}(V-U) = (r/2m-1)^{1/2}\mathrm{e}^{r/4m}\cosh t/4m \\ T(t,r) &= \tfrac{1}{2}(U+V) = (r/2m-1)^{1/2}\mathrm{e}^{r/4m}\sinh t/4m \quad , \end{aligned} \tag{27.153}$$

Remarks:

1. In these coordinates, the original "Schwarzschild patch" $r > 2m$, the region of validity of the Schwarzschild coordinates, corresponds to the region $-\infty < U < 0$, $0 < V < \infty$, or, in terms of X and T, to the region $X > 0$ and $X^2 - T^2 > 0$, or $|T| < X$. As Figure 19 shows, this 'Schwarzschild patch' is mapped to the first quadrant of the Kruskal-Szekeres metric, bounded by the lines $X = \pm T$.

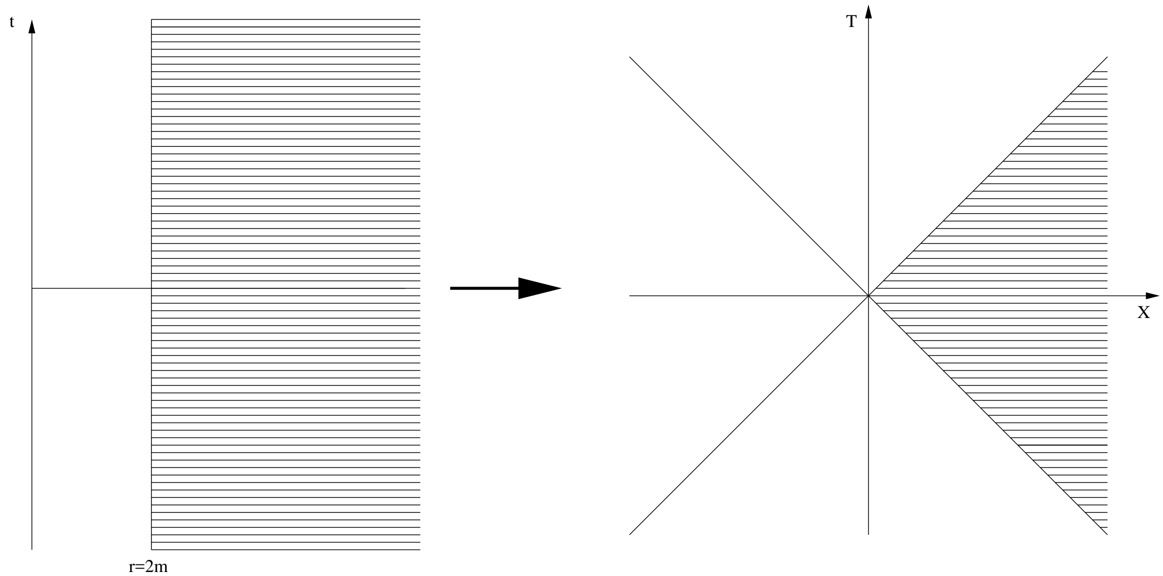


Figure 19: Schwarzschild patch in the Kruskal-Szekeres metric: the half-plane $r > 2m$ is mapped to the quadrant between the lines $X = \pm T$ in the Kruskal-Szekeres metric.

2. As in Minkowski space, null lines are given by $X = \pm T + const.$. In particular, therefore, lightcones have the standard Minkowskian form (slope ± 1) everywhere.

3. Surfaces of constant r are given by the lines (hyperboloids) $X^2 - T^2 = const.$

4. In particular, the boundary of the Schwarzschild patch is the horizon $r = 2m$. In terms of (U, V) this is mapped to $UV = 0$ which is the union of the two lines (null surfaces) $U = 0$ and $V = 0$, and in terms of (T, X) one has

$$r = 2m \quad \Rightarrow \quad X = \pm T \ , \tag{27.154}$$

5. In the region $r \gtrsim 2m$, the above coordinate transformation (27.153) reduces approximately (and up to constant factors) to the transformation (26.80) between the Rindler-like coordinates in that region and Minkowski coordinates,

$$T \sim t_M \sim (r - 2m)^{1/2} \sinh t/4m \quad , \quad X \sim x_M \sim (r - 2m)^{1/2} \cosh t/4m \tag{27.155}$$

6. In these coordinates the metric is now manifestly completely non-singular and regular not only at the horizons but everywhere except possibly at $r = 0$. We will discuss this extension of the Schwarzschild space-time in more detail in the next section.

7. Minor cosmetic improvements can be obtained by introducing, instead of U and V (27.142), the rescaled coordinates (u_K, v_K) through

$$\begin{aligned} du_K = \mathrm{e}^{-u/4m} du \quad &\Rightarrow \quad u_K = -4m\mathrm{e}^{-u/4m} = 4m\, U \\ dv_K = \mathrm{e}^{v/4m} dv \quad &\Rightarrow \quad v_K = +4m\mathrm{e}^{+v/4m} = 4m\, V \ . \end{aligned} \tag{27.156}$$

In particular, the coordinates (u_K, v_K) are now dimensionful (with dimension of length) while the coordinates (U, V) are dimensionless. In terms of (u_k, v_K) the metric takes the form

$$ds^2 = -\frac{2m}{r}\mathrm{e}^{-r/2m} du_K dv_K + r^2 d\Omega^2 \quad , \tag{27.157}$$

but then factors of $(16m^2)$ will reappear in other places. We will mostly work with the dimensionless coordinates (U, V) in the following.

8. It is also possible to further scale u_K and v_K so that the exponential prefactor has the form $\exp(-(r-2m)/2m)$ (e.g. by defining r^* with a different integration constant, $r^* \to r^* - 2m$), so that for $r \to 2m$ the metric tends to $ds^2 \to -du_k dv_k$ without additional numerical factors like e^{-1} but, as I said, this is pure cosmetics. In any case, as we will learn later on, U and V are coordinates that are naturally defined only up to affine transformations, so one choice is as good as any other.

9. Writing the implicit relation (27.147) between r and (U, V) as

$$-UV/e = (r/2m - 1)\mathrm{e}^{(r/2m-1)} \quad , \tag{27.158}$$

one sees that r can be expressed in terms of the *Lambert function* W,

$$r(U, V) = 2m\,(W(-UV/e) + 1) \tag{27.159}$$

defined by

$$x = W(x)\mathrm{e}^{W(x)} \tag{27.160}$$

($W(x) \geq W(-1/e) = -1$, $W(0) = 0$ is the horizon).[86]

10. Note that the transformation (27.156) is strictly identical to the transformation (3.97) between Minkowski and Rindler null coordinates, with the identification $a = 1/4m$. This is much more profound than it sounds at first. In particular, we will learn in section 27.10 that $\kappa = 1/4m$ is the *surface gravity* of a black hole, i.e. a measure of the gravitational acceleration at the horizon. It is therefore this acceleration that plays the same role in the black hole context as the acceleration called a in the Rindler context. Moreover, this identity is at the heart of the deep relation between the so-called *Unruh Effect* (thermal nature of a Minkowski QFT vacuum state when seen by an accelerating observer, already mentioned in section 7.8) and the *Hawking Effect* (quantum radiation of black holes).[87]

[86] For some applications of the Lambert function in this context see e.g. K. Lake, *Some notes on the Kruskal - Szekeres completion*, `arXiv:1002.3600 [gr-qc]`, *Some further notes on the Kruskal - Szekeres completion*, `arXiv:1202.0860 [gr-qc]`.

[87] For an introduction to QFT in curved spacce-time and a discussion of these effects, see e.g. the monographs N. Birrell, P. Davies, *Quantum Fields in Curved Space*, V. Mukhanov, S. Winitzki, *Quantum Effects in Gravity*, or the on-line lecture notes J. Traschen, *An Introduction to Black Hole Evaporation*, `arXiv:gr-qc/0010055`, T. Jacobson, *Introduction to Quantum Fields in Curved Spacetime and the Hawking Effect*, `arXiv:gr-qc/0308048`, C. Krishnan, *Quantum Field Theory, Black Holes and Holography*, `arXiv:1011.5875`.

11. Further insight into the nature of Kruskal-Szekeres coordinates will be provided in section 27.10 (where we elucidate the physical significance of the coordinates (U, V)) and in section 31.9, where we outline the construction of such coordinates for more general black hole metrics.

27.8 Maximal Extension of Schwarzschild: the Kruskal Diagram

Now that we have the coordinates X and T, we can let them range over all the values for which the metric is non-singular. The only remaining singularity is at $r = 0$, which corresponds to the two sheets of the hyperboloid

$$r = 0 \quad \Leftrightarrow \quad T^2 - X^2 = 1 \quad \Leftrightarrow \quad T = \pm\sqrt{1 + X^2} \ . \tag{27.161}$$

That $r = 0$ is indeed a real singularity that cannot be removed by a coordinate transformation can be shown by calculating some invariant of the curvature tensor, like the Kretschmann scalar $K = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$ (8.61). On purely dimensional grounds, from

$$\left(\partial_r^2(1 - 2m/r)\right)^2 \sim m^2/r^6 \ , \tag{27.162}$$

say, one would expect K to be proportional to m^2/r^6, the crucial feature being that the constant of proportionality is not zero, explicit calculations (this is a doable but thoroughly unenlightning exercise) showing that

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = 48\frac{m^2}{r^6} \ . \tag{27.163}$$

Thus the geometry is genuinely singular at $r = 0$. Nevertheless, since the metric is non-singular for all values of (X, T) subject to the constraint $r > 0$ or $T^2 - X^2 < 1$, there is no physical reason to exclude the regions in the other quadrants also satisfying this condition.

By including them, we obtain the *Kruskal-Szekeres Extension* of the Schwarzschild metric, displayed in the *Kruskal diagram* in Figure 20. This extension of the Schwarzschild metric was discovered independently by Kruskal and Szekeres in 1960 and presents us with an amazingly rich and complex picture of what originally appeared to be a rather simple (and perhaps even dull) solution to the Einstein equations. It can be shown that this represents the maximal analytic extension of the Schwarzschild metric in the sense that every affinely parametrised geodesic can either be continued to infinite values of its parameter or runs into the singularity at $r = 0$ at some finite value of the affine parameter.[88]

In addition to the Schwarzschild patch, quadrant I, we have three other regions, living in the quadrants II, III, and IV, each of them having its own peculiarities. Note that

[88] One can also show that any spherically symmetric solution of the vacuum Einstein equations is locally isometric to some domain of the Schwarzschild-Kruskal solution. For a proof of this variant of a generalised Birkhoff theorem see e.g. N. Straumann, *General Relativity*.

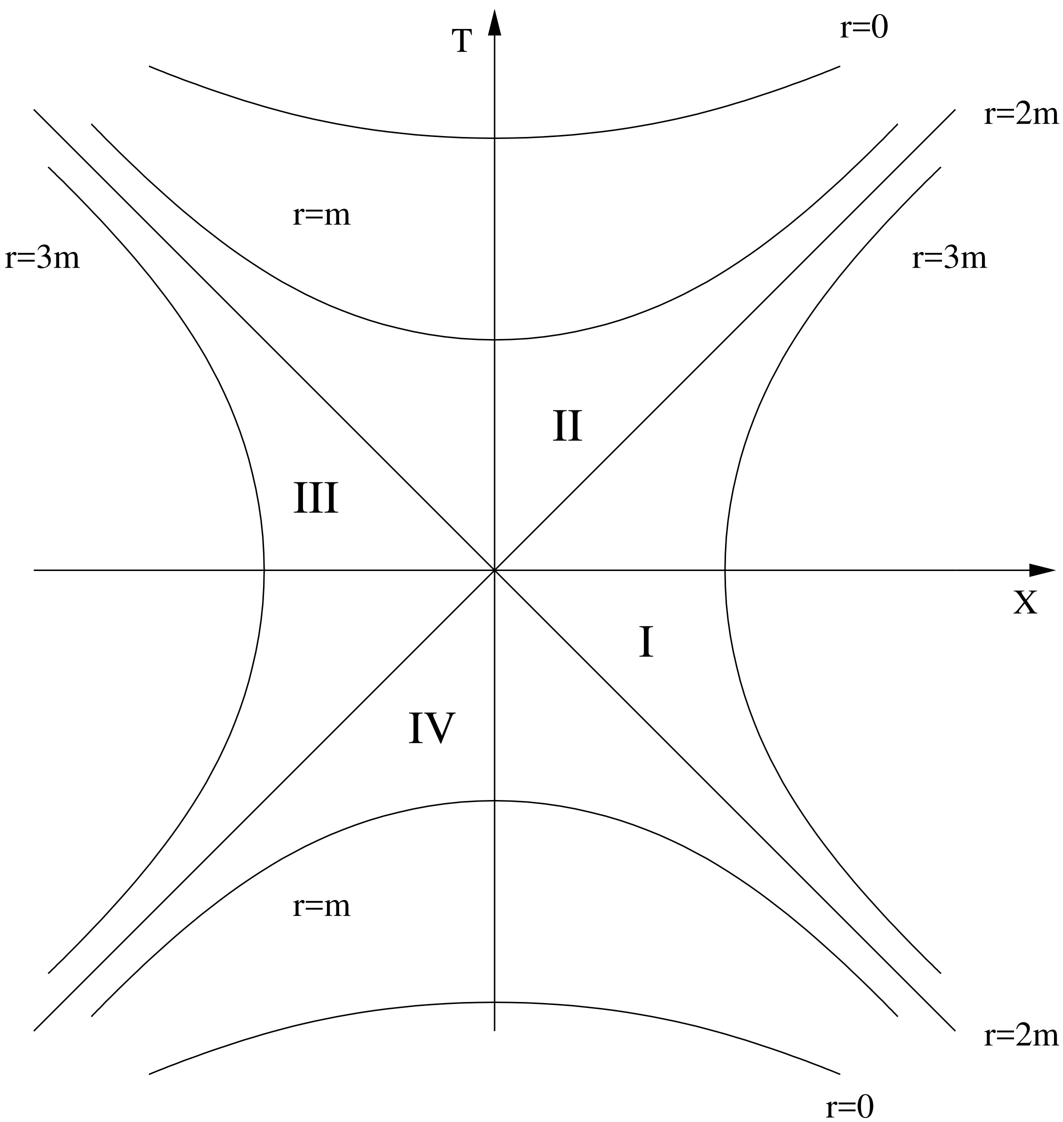


Figure 20: Complete Kruskal-Szekeres universe. Diagonal lines are null, lines of constant r are hyperbolas. Region I is the Schwarzschild patch, separated by the horizon from regions II and IV. The Eddington-Finkelstein coordinates (v, r) cover regions I and II, (u, r) cover regions I and IV. Regions I and III are filled with lines of constant $r > 2m$. They are causally disconnected. Observers in regions I and III can receive signals from region IV and send signals to region II. An observer in region IV can send signals into both regions I and III (and therefore also to region II) and must have emerged from the singularity at $r = 0$ at a finite proper time in the past. Any observer entering region II will be able to receive signals from regions I and III (and therefore also from IV) and will reach the singularity at $r = 0$ in finite time. Events occuring in region II cannot be observed in any of the other regions.

obviously the conversion formulae from $(r,t) \to (X,T)$ in the quadrants II, III and IV differ from those given above for quadrant I. E.g. in region II one can use Schwarzschild(-like) coordinates in which the metric reads

$$ds^2 = \left(\frac{2m}{r} - 1\right) dt^2 - \left(\frac{2m}{r} - 1\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{27.164}$$

(these are not the same coordinates as those in patch I, as we have seen we cannot continue the Schwarzschild coordinates across the horizon), and in this quadrant (where r is a time coordinate etc.) the relation between Schwarzschild and Kruskal coordinates is

$$\begin{aligned} X &= (1 - r/2m)^{1/2} \mathrm{e}^{r/4m} \sinh t/4m \\ T &= (1 - r/2m)^{1/2} \mathrm{e}^{r/4m} \cosh t/4m \ . \end{aligned} \tag{27.165}$$

To get acquainted with the Kruskal diagram, let us note the following basic facts (some of which we had already noted for region I in the previous section).

1. Null lines are diagonals $X = \pm T + const.$, just as in Minkowski space. This greatly facilitates the exploration of the causal structure of the Kruskal-Szekeres metric.

2. In particular, the horizon corresponds to the two lines $X = \pm T$.

3. Lines of constant r are hyperbolas. For $r > 2m$ they fill the quadrants I and III, for $r < 2m$ the other regions II and IV.

4. In particular, the singularity at $r = 0$ is given by the two sheets of the hyperbola $T^2 - X^2 = 1$.

5. Notice in particular also that in regions II and IV worldlines with $r = const.$ are no longer timelike but spacelike. Including the transverse 2-sphere, this should be rephrased as either "lines of constant (r, θ, ϕ) are spacelike for $r < 2m$" or "the surfaces of constant r are spacelike for $r < 2m$".

6. Lines of constant Schwarzschild time t are straight lines through the origin. E.g. in region I one has $X = (\coth(t/4m))T$, with the future horizon $X = T$ corresponding, as expected, to $t \to \infty$.

7. The Eddington-Finkelstein coordinates (v, r) cover the regions I and II, the coordinates (u, r) the regions I and IV.

8. Quadrant III is completely new and is separated from region I by a spacelike distance. That is, regions I and III are causally disconnected. This is the region that already prematurely made a brief appearance when we analysed the coordinate transformation between Schwarzschild and isotropic coordinates at the end of section 24.5. Thus isotropic coordinates actually cover the regions I and III.

Now let us see what all this tells us about the physics of the Kruskal-Szekeres metric.

- An observer in region I (the familiar patch) can send signals into region II and receive signals from region IV.
- The same is true for an observer in the causally disconnected region III.
- Once an observer enters region II from, say, region I, he cannot escape from it anymore and he will run into the catastrophic region $r = 0$ in finite proper time.
- As a reward for his or her foolishness, between having crossed the horizon and being crushed to death, our observer will for the first time be able to receive signals and meet observers emerging from the mirror world in region III.
- Events occurring in region II cannot be observed anywhere outside that region (black hole).
- Finally, an observer in region IV must have emerged from the (past) singularity at $r = 0$ a finite proper time in the past and can send signals and enter into either of the regions I or III.

An even better visualisation of the causal *and global* structure of the maximally extended Schwarzschild solution is provided by its *Carter-Penrose Conformal Diagram* (or *Penrose diagram* for short), given in Figure 21. Its construction (and the notation used here) will be explained in section 28.

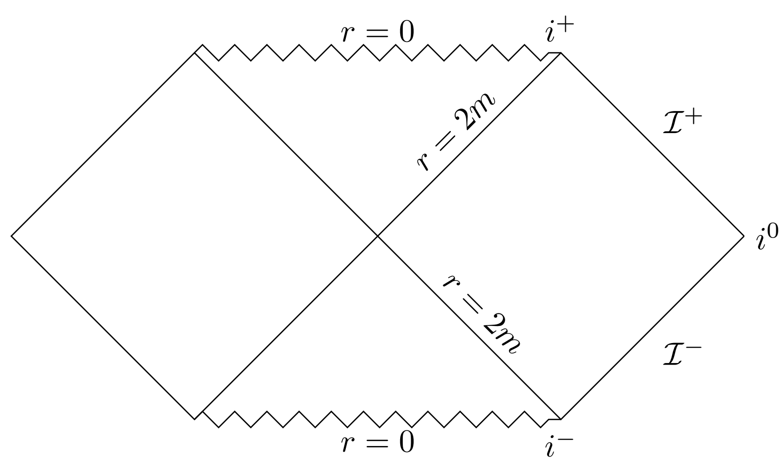


Figure 21: Carter-Penrose Conformal Diagram of the maximal Kruskal-Szekeres extension of the Schwarzschild space-time. For more details, see section 28.

All in all, the picture that we have uncovered of the complete maximally extended Schwarzschild space-time is quite intricate and rich, but also somewhat peculiar (to say the least). In particular, the mirror region III and the white hole region IV appear to be quite unphysical. As we will see in section 29.3, reassuringly the existence of these regions is an artefact of an eternal black hole solution and these regions do not actually exist for astrophysical black hole solutions that are formed by gravitational collapse.

27.9 Properties of the Asymptotically Timelike Killing Vector

An interesting aspect of the Kruskal-Szekeres geometry is its dynamical character. This may appear to be a strange thing to say since we explicitly started off with a static metric, but this statement applies only to region I (and its mirror III). An investigation of the behaviour of spacelike slices analogous to that we performed for region I (see Figure 11) reveals a dynamical picture of continuing gravitational collapse in region II. In simple terms, the loss of staticity can be understood by noting that the timelike Killing vector field $\xi = \partial_t$ of region I, when expressed in terms of Kruskal coordinates, becomes null on the horizon and spacelike in region II.

Indeed it is easy to check from (27.147) and (27.148) that the time-translation symmetry $(t,r) \to (t+c,r)$ of the Schwarzschild patch corresponds to the transformation

$$U \to \mathrm{e}^{-c/4m}U \qquad V \to \mathrm{e}^{c/4m}V \ . \tag{27.166}$$

This is a boost in the (U,V) or (T,X)-plane which leaves the entire Kruskal metric invariant since

$$dUdV = dT^2 - dX^2 \tag{27.167}$$

is invariant and $r = r(U,V)$ depends on U and V only via the boost-invariant (time-independent) quantity UV (27.147),

$$UV = T^2 - X^2 = -(r/2m - 1)\mathrm{e}^{r/2m} \quad \Rightarrow \quad r = r(T^2 - X^2) \ . \tag{27.168}$$

Thus this symmetry is generated by the Killing vector

$$\xi = (V\partial_V - U\partial_U)/4m = (X\partial_T + T\partial_X)/4m \ . \tag{27.169}$$

It follows that ξ has norm proportional to $T^2 - X^2$,

$$||\xi||^2 = \frac{2m}{r}\mathrm{e}^{-r/2m}(T^2 - X^2) \ . \tag{27.170}$$

There are thus three different cases to consider, each of them interesting in its own right:

- ξ timelike

  ξ is timelike in the original region I (and in the mirror region III), with (27.151)

  $$I: \quad ||K||^2 = -\frac{2m}{r}\mathrm{e}^{-r/2m}(\frac{r}{2m} - 1)\mathrm{e}^{r/2m} = -(1 - \frac{2m}{r}) \ , \tag{27.171}$$

  confirming that $\xi = \partial_t$ in region I. We thus recover the statement that the Schwarzschild metric in the Schwarzschild patch is static.

- ξ spacelike

  ξ is spacelike in region II. Thus region II has no timelike Killing vector field, therefore cannot possibly be static, but has instead an additional spacelike Killing vector field - cf. the Remark 1 in section 24.6 in connection with Birkhoff's theorem.

  Related to this is the fact, already mentioned above, that in regions II and IV the slices of constant r are no longer timelike but spacelike surfaces. Thus they are analogous to, say, constant t or T slices for $r > 2m$. Just as it does not make sense to ask "where is the slice $t = 1$?" (say), only "when is $t = 1$?", or "where is $r = 3m$?", in these regions it makes no sense to ask "where is $r = m$?", only "when is $r = m$?".

- ξ null

  ξ is null on the horizon. This turns out to be the most interesting case, and turns out to be one of the characteristic features of static black holes in general, and therefore I will elaborate on this a bit in the following.

27.10 First Encounter with Killing Horizons and Surface Gravity

In this section, we will perform a rather pedestrian analysis of some properties of the horizon which are related to the fact that the Killing vector ξ becomes null on this null surface. These properties can be derived and understood without a general knowledge of the geometry of null hypersurfaces, as discussed in section 17, but it is useful to keep in mind that they are just special cases of the properties of null hypersurfaces, and of more general Killing horizons (to be discussed later on in section 32.5).

In particular, from section 17.2 one knows on general grounds that the integral curves of the normal vector ξ are (possibly non-affinely parametrised) null geodesics, but it is instructive to rederive this here from scratch, and from a slightly different perspective.

ξ is null on (and tangent to) the horizon $T = \pm X$ or $UV = 0$. In fact it reduces to $\xi = (V/4m)\partial_V = \partial_v$ on the horizon $U = 0$ (and to $\xi = \partial_u$ on $V = 0$). As a consequence, due to the characteristic peculiarities of null hypersurfaces (section 17), ξ is also normal to the horizon. In this context (or from this perspective) the horizon is known as a *Killing horizon* (the locus where a Killing vector becomes null and orthogonal to a null surface).

Therefore ξ generates translations along the horizon, $v \to v + c$, and is called the *null generator* of (this branch of) the horizon. Thus v can naturally be used as a coordinate there.

Evidently, ξ vanishes at the "point" $U = V = 0$ where the other horizon $V = 0$ branches off. Actually this is of course not a point but a perfectly respectable 2-sphere of radius $r = 2m$ known as the *bifurcation surface* of the Killing horizon of the Schwarzschild geometry, and ξ vanishing means that this 2-sphere is invariant under the time-translation generated by ξ, something that is also evident from (27.166). On $U = 0$ this branching point lies at $V \to 0 \Rightarrow v \to -\infty$, so the Eddington-Finkelstein coordinate $v \in (-\infty, +\infty)$ only covers the half-line $U = 0, V > 0$ of the horizon.

On the other hand the line $U = 0$ is itself an "outgoing" (radial) null geodesic, but in light of the above v cannot possibly be an affine parameter along that geodesic (the affine parameter should not reach infinite values half-way along the geodesic). The failure of v to be an affine parameter on the horizon can be quantified by calculating the acceleration $\nabla_\xi \xi = \nabla_{\partial t}\partial_t$ (5.99) of the (integral curves of the) Killing vector and then taking the limit $r \to 2m$. This calculation is quite painless in Eddington-Finkelstein coordinates (v, r) in which $\xi = \partial_v$ everywhere,

$$\xi = \partial_t = (\partial_t v)\partial_v + (\partial_t r)\partial_r = \partial_v \tag{27.172}$$

($\xi = \partial_v$ is evidently a Killing vector because the components of the metric do not depend on v in Eddington-Finkelstein coordinates). Then the acceleration is

$$\nabla_\xi \xi \equiv (\nabla_\xi \xi)^\alpha \partial_\alpha = \Gamma^\alpha_{vv}\partial_\alpha = (f'/2)(f\partial_r + \partial_v) \tag{27.173}$$

where $f(r) = 1 - 2m/r$. Thus for $r \to 2m$ one finds

$$\lim_{r\to 2m} \nabla_\xi \xi = \left(\tfrac{1}{2} f'(r)|_{r=2m}\right)\xi = \left(\frac{m}{r^2}|_{r=2m}\right)\xi = \tfrac{1}{4m}\xi \quad . \tag{27.174}$$

<u>Remarks:</u>

1. Since we have $\nabla_\xi \xi \sim \xi$ on the horizon, this shows first of all that there it generates a non-affinely parametrised geodesic (see (2.136) and (5.101)). Moreover, we see that one interpretation of the ubiquitous factor

$$\kappa = \frac{1}{4m} \tag{27.175}$$

   is that it measures the *inaffinity*, i.e. the failure of v to be an affine parameter on the horizon,

$$\lim_{r\to 2m} \nabla_\xi \xi = \kappa\xi \quad . \tag{27.176}$$

2. Another interpretation brought out by the same calculation (27.174) is that $\kappa = 1/4m$ is the *surface gravity* which measures the strength of the gravitational force (acceleration) $\mathsf{a}(r)$ (26.11) acting on a static observer at the horizon, but as measured at infinity (by taking into account the redshift factor $f(r)^{1/2}$),

$$\kappa = \lim_{r\to 2m} f(r)^{1/2}\mathsf{a}(r) = \tfrac{1}{2} f'(r)|_{r=2m} = 1/4m \quad . \tag{27.177}$$

Anyway, to return to the beginning of this story, we have seen that v is not an affine parameter along the horizon. It turns out, however, that the Kruskal coordinate V is an affine parameter there (and this is one way of understanding why Kruskal coordinates are so natural for exploring the causal structure of the metric), meaning that the null curve

$$x^\alpha(\lambda) = (U(\lambda), V(\lambda), \theta(\lambda), \phi(\lambda)) = (0, \lambda, \theta_0, \phi_0) \tag{27.178}$$

is an affinely parametrised null geodesic. This follows on the nose from (27.174) and the result (2.138) of section 2.7. Noting that κ is constant (actually not just along the geodesic but on the entire horizon, but the former is all we need), (2.138) with $\tau \to \lambda$ and $\sigma \to v$ becomes (dropping integration constants)

$$\frac{d\lambda}{dv} \sim \mathrm{e}^{\kappa v} \quad \to \quad \lambda(v) \sim \kappa^{-1}\mathrm{e}^{\kappa v} \sim V \ . \tag{27.179}$$

Another way to see this, which provides some more insight, is to analyse the geodesic Lagrangian and the conserved quantity associated to ξ in ingoing Eddington-Finkelstein coordinates (one could also work in Kruskal coordinates, but nothing is gained by that). The elementary steps in the calculation are the following:

- From the metric (27.82) we deduce that for a radial null-geodesics one has

  $$-f(r)\dot{v}^2 + 2\dot{v}\dot{r} = 0 \tag{27.180}$$

  where a dot denotes a derivative with respect to the affine parameter λ. The geodesics with $v = \text{const}$ describe ingoing null-geodesics and we are not interested in these, so we have

  $$-f(r)\dot{v} + 2\dot{r} = 0 \ , \tag{27.181}$$

  and since $u = t - r^* = v - 2r^*$ and $dr^*/dr = f(r)^{-1}$, this is equivalent to $\dot{u} = 0$ and, as anticipated, describes outgoing geodesics.

- The conserved quantity E associated to the Killing vector $\xi = \partial_v$ is

  $$f(r)\dot{v} - \dot{r} = E \ . \tag{27.182}$$

  From the two preceding equations we immediately deduce (with a convenient parametrisation of the integration constant)

  $$\dot{r} = E \quad \Rightarrow \quad r(\lambda) = E(\lambda - \lambda_0) + r_s \ . \tag{27.183}$$

  For the null geodesic along the horizon we are ultimately interested in, we have $r(\lambda) = r_s$, i.e. $E = 0$, but we need to approach this with some care, so we keep the general solution for now. Analogously, for v we find the equation

  $$f(r)\dot{v} = 2E \quad \Rightarrow \quad \dot{v} = \frac{2Er(\lambda)}{r(\lambda) - r_s} = 2E + \frac{2r_s}{\lambda - \lambda_0} \ . \tag{27.184}$$

We can now take the limit $E \to 0$ with impunity, and are left with

$$\dot{v} = \frac{2r_s}{\lambda - \lambda_0} \quad \Rightarrow \quad v(\lambda) = 2r_s \ln(\lambda - \lambda_0) + \text{const.} \quad . \tag{27.185}$$

- The prefactor $2r_s = 4m$ is now precisely such that it cancels the factor $1/4m$ in the definition of the Kruskal coordinate V (27.142), so that

$$V(\lambda) = \mathrm{e}^{v(\lambda)/4m} = a(\lambda - \lambda_0) \quad , \tag{27.186}$$

  which is precisely the statement that V is related to λ by an affine transformation. Thus V is an affine parameter along the horizon $U = 0$, as claimed.

- We also see explicitly that for other outgoing null geodesics, i.e. those with $E \neq 0$, and with solution

$$v(\lambda) = 2E\lambda + 4m \ln(\lambda - \lambda_0) + \text{const.} \quad , \tag{27.187}$$

  neither v nor V is an affine parameter. However, it is apparent from the explicit solution $r(\lambda) = E\lambda + b$ given above that, for $E \neq 0$, $r(\lambda)$ is related to λ by an affine transformation, and hence for these geodesics r is an affine parameter.

<u>Remarks:</u>

1. We now have two natural coordinates on the future horizon $U = 0, V > 0$ which we can for instance use to measure the frequency of incoming waves. $\xi = \partial_t = \partial_v$ measures what is commonly called Killing frequency (this requires no further explanation since it is associated to the Killing vector which generates time-translations in the Schwarzschild patch), and is the natural notion of frequency to be used by static observers with 4-velocvity $u^\alpha \sim \xi^\alpha$.

   ∂_V, on the other hand, measures the so-called free fall frequency since, more or less by construction, a freely falling observer in the Schwarzschild geometry near $r = 2m$ will see approximately the Minkowski space-time metric $ds^2 \sim -dUdV$ (recall that the transformation (3.97) between Rindler and Minkowski coordinates is strictly analogous to the transformation (27.156) between Schwarzschild and Kruskal coordinates). The exponential relation (27.142) between them reflects the exponential blue- or redshift we first encountered in section 26.5.

2. As a concluding remark to this section, I cannot resist mentioning that the notion of surface gravity κ plays a crucial role in the analysis of the classical dynamics of general black holes, and even more so in the semi-classical context, since it is directly proportional to the temperature of the famous Hawking radiation of an evaporating black hole,

$$T_H = \frac{\kappa}{2\pi} = \frac{1}{8\pi G_N M} \quad . \tag{27.188}$$

For more details on this and other related advanced topics I am not able to cover or do justice to here, I refer you to the references given in footnote 87 of section 27.7, as well as to the superb Cambridge lecture notes on Black Holes.[89]

27.11 From Eddington-Finkelstein to Israel(-Klösch-Strobl) Coordinates

There is one more coordinate system for the Schwarzschild geometry that I want to mention because it is quite remarkable and, equally remarkably, apparently not widely known or used. It was discovered by W. Israel in 1966, and rediscovered several times since, most recently by T. Klösch and T. Strobl in a different particularly insightful way.[90] I will introduce these coordinates in a way that is complementary to those in these articles (and that is also motivated by certain generalisations, that I will however not discuss here).

Recall that the Kruskal coordinates were based on suitably combining the outgoing and ingoing Eddington-Finkelstein coordinates. Now, more generally one frequently encounters the situation that one knows a solution either in coordinates adapted to ingoing null geodesics, or in coordinates adapted to outgoing null geodesics, but usually not both (and given one constructing the other is usually a hard problem that may have no simple analytical solution).

Thus, let us assume that we are given (or have found) the Schwarzschild metric in outgoing Eddington-Finkelstein coordinates,

$$ds^2 = -f(r)du^2 - 2du\,dr + r^2 d\Omega^2 \quad , \quad f(r) = 1 - 2m/r \ . \tag{27.189}$$

We are happy and proud of this, but we quickly realise that this cannot be the end of the story because the above coordinates provide an incomplete covering of the space-time. The simplest way to discover this is to study radial lightrays or null geodesics, governed by the two equations

$$f(r)\dot{u}^2 + 2\dot{u}\dot{r} = 0 \tag{27.190}$$

(the null condition), and

$$f(r)\dot{u} + \dot{r} = E \tag{27.191}$$

(due to the u-translation invariance). One set of null geodesics is simply given by $\dot{u} = 0$, i.e. $u = u_0$ constant, and for these one has

$$u(\tau) = u_0 \quad \Rightarrow \quad \dot{r} = E \quad \Rightarrow \quad r(\tau) = E\tau + r_0 \ . \tag{27.192}$$

[89] P. Townsend, *Black Holes*, `arXiv:gr-qc/9707012v1`.

[90] W. Israel, *New Interpretation of the extended Schwarzschild Manifold*, Phys. Rev. 143 (1966) 1016-1021; T. Klösch, T. Strobl, *Explicit Global Coordinates for Schwarzschild and Reissner-Nordstroem*, `arXiv:gr-qc/9507011`. It is also mentioned in section 15.4.4 of the directory of exact solutions, H. Stephani, D. Kramer, M. MacCallum, C. Hoenslaers, E. Herlt, *Exact Solutions to Einstein's Field Equations* (2nd Edition). See also K. Lake, *An explicit global covering of the Schwarzschild-Tangherlini black holes*, `arXiv:gr-qc/0306073`, and *Maximally extended, explicit and regular coverings of the Schwarzschild - de Sitter vacua in arbitrary dimension*, `arXiv:gr-qc/0507031` for some generalisations.

For future-oriented null geodesics one needs $E > 0$, and therefore one has $\dot{r} > 0$. These are the outgoing null geodesics to which the outgoing Eddington-Finkelstein coordinate system is adapted. Here r_0 is an integration constant which, by an affine transformation (actually just a shift) of τ, we could e.g. without loss of generality choose to be $r_0 = 2m$. Then the solution describes the outgoing null geodesics that emerge from the past event horizon at $r = 2m$ for $\tau = 0$.

The other set of (thus ingoing) null geodesics has $\dot{u} \neq 0$ and is therefore governed by the equations

$$f(r)\dot{u} + 2\dot{r} = 0 \quad \text{and} \quad f(r)\dot{u} + \dot{r} = E \ . \tag{27.193}$$

Subtracting the two one finds $\dot{r} = -E$, and therefore (again with a convenient choice of integration constant or origin of τ)

$$r(\tau) = 2m - E\tau \quad , \quad u(\tau) = -4m \log|\tau| + 2E\tau + c \qquad (-\infty < \tau < 0) \tag{27.194}$$

The restriction on τ arises because for $\tau \to 0_-$ one has

$$\tau \to 0_- \quad \Rightarrow \quad r \to 2m \quad , \quad u \to \infty \ . \tag{27.195}$$

Thus we discover that we reach $u = +\infty$ in finite (affine) time, we run out of coordinate space as the ingoing null geodesics approach $r = 2m$ (and the coordinate u is evidently not suitable for describing what happens beyond $r = 2m, u = \infty$).

Note that this locus is most certainly not the past event horizon ($r = 2m$, $u = u_0$ finite), as we know that we can only cross that horizon along future-directed paths from smaller to larger values of r (the white hole). Thus we have discovered a new barrier (which also turns out to be a horizon, and which, with hindsight, we know is the future event horizon), and we realise that the outgoing Eddington-Finkelstein coordinates do not cover the whole space-time.

Now let us assume that we do not have the luxury of being able to appeal to ingoing Eddington-Finkelstein coordinates to construct a future extension (and subsequently the maximal Kruskal-Szekeres extension) of this space-time. How could we proceed?

Since τ is finite as $u = +\infty$ is reached, it is natural to introduce something like τ as a new coordinate. Looking at the leading term of $u(\tau)$ for $\tau \to 0$, one is then led to introduce a new coordinate x through

$$u(\tau) \approx -4m \log|\tau| \quad \Rightarrow \quad u(x) = -4m \log|x| \ , \tag{27.196}$$

so that the complete range of u, $u \in (-\infty, +\infty)$, is covered as x runs over the interval $x \in (-\infty, 0)$. This clearly now permits to continue the ingoing null geodesics beyond $u = +\infty$. However, if one just replaces $u \to x$, the metric appears to be singular at $x = 0$. This can be rectified by introducing a new coordinate y through the relation

$$r - 2m = xy \ , \tag{27.197}$$

its range subject to the condition $r > 0$. To better understand the rationale for this change of variables, note that as $u \to \infty$ one has $r - 2m \sim \tau \sim x$ so that $(r-2m)/x$ remains finite in the limit - and can therefore be used as a new coordinate, the one that we have called y.

It is straightforward to see that in these coordinates the metric takes the form

$$ds^2 = \frac{8my^2}{xy+2m}dx^2 + 8m\, dx\, dy + (xy+2m)^2 d\Omega^2 \quad . \tag{27.198}$$

This is the Schwarzschild metric in *Israel coordinates*. Before discussing some of it most important properties, note that by the simple scaling $u_I = 4mx$, so that

$$u_I = 4mx = -4m\mathrm{e}^{-u/4m} \equiv u_K \tag{27.199}$$

is just the "canonically normalised" Kruskal coordinate introduced in (27.156), and the relabelling $y \to v_I$, one can put the metric into the somewhat more common "canoncially normalised" form

$$ds^2 = \frac{v_I^2}{2m((u_I/4m)v_I + 2m)}du_I^2 + 2\, du_I\, dv_I + ((u_I/4m)v_I + 2m)^2 d\Omega^2 \quad . \tag{27.200}$$

Here are some of the key-properties of this metric (to describe these I will continue to use the coordinates (x, y), i.e. the form (27.198) of the metric):

1. First of all, one sees that the metric is completely non-singular as long as $xy > -2m$, and one can therefore let x and y run over all the values for which this condition is satisfied.

2. The horizon is given by the degenerate hyperboloid

$$r = 2m \quad \Leftrightarrow \quad xy = 0: \quad \{x = 0\} \cup \{y = 0\} \tag{27.201}$$

3. Then one sees immediately that this space-time covers four distinct "patches":

   - For $x < 0$ and $y < 0$ one has $r > 2m$: this is the original Schwarzschild patch.
   - There are two regions for which x and y have opposite sign, subject to the condition $xy > -2m$: these are the white hole region behind the past horizon ($x < 0, y > 0$), covered by the original outgoing Eddington-Finkelstein coordinates, and the new region beyond the future event horizon with $x > 0, y < 0$.
   - Finally, for $x > 0$ and $y > 0$ one also has $r > 2m$: this is a distinct region isometric to the Schwarzschild patch, the mirror region.

4. Thus the Israel coordinates provide not only a future extension of the Schwarzschild metric in Schwarzschild or outgoing Eddington-Finkelstein coordinates but actually a complete covering of the maximal Kruskal-Szekeres extension of the Schwarzschild space-time.

5. The usual Kruskal-Szekeres coordinates (U, V) are related to the Israel coordinates by
$$U \sim x \quad , \quad V \sim y\, \mathrm{e}^{xy/2m + 1} \quad . \tag{27.202}$$

6. The main difference to (and advantage compared with) Kruskal-Szekeres coordinates is that for the Israel coordinates the coordinate transformation to Schwarzschild coordinates and its inverse are completely explicit (whereas the radial coordinate r is only given implicitly in terms of the Kruskal-Szekeres coordinates). E.g. for $x < 0$ one has
$$\left.\begin{array}{ll} u & = -4m \log(-x) \\ r & = xy + 2m \end{array}\right\} \quad \Leftrightarrow \quad \left\{\begin{array}{ll} x & = -\mathrm{e}^{-u/4m} \\ y & = -(r-2m)\mathrm{e}^{u/4m} \end{array}\right. \tag{27.203}$$

7. By undoing the coordinate transformation $u = t - r^*$ from Schwarzschild to Eddington-Finkelstein coordinates, one can also explicitly write the transformation from Schwarzschild coordinates (t, r) to the Israel coordinates,
$$r(x,y) = xy + 2m \quad , \quad t(x,y) = xy + 2m + \log|y/2mx| \quad . \tag{27.204}$$

8. By inspection, the metric depends only on products like xy, $dxdy$ etc. Thus it has the isometry
$$(x,y) \to (\lambda x, \lambda^{-1} y) \tag{27.205}$$
which corresponds (and reduces) to the time-translation symmetry in the Schwarzschild patch, generated by
$$\partial_t = (y\partial_y - x\partial_x)/4m \quad . \tag{27.206}$$

9. Radial null geodesics are thus goverened by the two equations
$$\frac{y^2}{xy+2m}\dot{x}^2 + \dot{x}\dot{y} = 0 \tag{27.207}$$
(the null condition), and
$$\frac{2y^2}{xy+2m}x\dot{x} + \dot{y}x - \dot{x}y = c \quad . \tag{27.208}$$
(from time-translation invariance). They are therefore of two kinds:
   - One has the curves with x (as well as, of course, (θ, ϕ)) constant, so that these are straight lines in an (x, y)-diagram. For these curves, (27.208) implies that $\dot{y}$ is constant, so that y an affine parameter along these null geodesics.
   - The other null geodesics are given by the relation
$$\frac{y^2}{xy+2m}\dot{x} + \dot{y} = 0 \quad . \tag{27.209}$$

Multiplying by x and subtracting twice that from (27.208), one finds

$$\dot{x}y + x\dot{y} = c \quad \Rightarrow \quad x(\tau)y(\tau) = c\tau + d \ . \tag{27.210}$$

Using this to eliminate x and $\dot{x}$ from (27.209), the equation of motion for y reduces to

$$\dot{y} = -(c/2m)y \quad \Rightarrow \quad y(\tau) = y_0 \mathrm{e}^{-(c/2m)\tau} \ , \tag{27.211}$$

or $c\tau = -2m\log(y(\tau)/y_0)$. Substituting this in (27.210), one finds that these radial null geodesics are given by the curves

$$xy = -2m\log(Cy) \tag{27.212}$$

in the (x, y)-plane, with C an arbitrary constant.

Armed with this information, it is straightforward to draw an Israel analogue of the Kruskal diagram, containing an equivalent amount of information. All in all, Israel coordinates are an attractive and easy to construct and understand alternative to Kruskal-Szekeres coordinates.

27.12 Appendix: Summary of Schwarzschild Coordinate Systems

Here is, to wrap up this section, a list of the coordinate systems that we have used to gain insight into the properties of the Schwarzschild metric. [As region III is isometric to region I, this doubling of possibilities (e.g. Schwarzschild coordinates cover I *or* III, etc.) has not been indicated in the last column. The perfectly valid but somewhat un-insightful option to use Schwarzschild coordinates only in region II, say, and variants thereof, have also not been indicated in the last column.]

Name	Line Element	Adapted to	Covers
Schwarzschild	$-f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2$	Stationary Observers	I
Tortoise	$f(r)(-dt^2 + (dr^*)^2) + r(r^*)^2 d\Omega^2$	In & Out Null Geodesics	I
Isotropic	$-(g_-(\rho)/g_+(\rho))dt^2 + g_+(\rho)^2(d\rho^2 + \rho^2 d\Omega^2)$	Stationary Observers	I,III
Painlevé - Gullstrand	$-f(r)d\tau^2 \pm 2\sqrt{\frac{2m}{r}}dr d\tau + (dr^2 + r^2 d\Omega^2)$	In/Out Timelike Geodesics	I,II / I,IV
Lemaître	$-d\tau^2 + (2m/r(\tau,\rho))d\rho^2 + r(\tau,\rho)^2 d\Omega^2$	In Timelike Geodesics	I,II
Novikov	$-d\tau^2 + \frac{R^2+1}{R^2}(r')^2 dR^2 + r(\tau,R)^2 d\Omega^2$	In Timelike Geodesics	I-IV
Eddington - Finkelstein	$-f(r)d(v/u)^2 \pm 2d(v/u)dr + r^2 d\Omega^2$	In/Out Null Geodesics	I,II / I,IV
Kerr-Schild	$-d\tilde{t}^2 + dr^2 + r^2 d\Omega^2 + \frac{2m}{r}(d(\tilde{t} \pm r))^2$	In/Out Null Geodesics	I,II / I,IV
Kruskal - Szekeres	$F(r(T,X))(-dT^2 + dX^2) + r(T,X)^2 d\Omega^2$	In & Out Null Geodesics	I-IV
Israel	$F_I(x,y)dx^2 + 8m\, dx dy + (xy+2m)^2 d\Omega^2$	In/Out Null Geodesics	I-IV

Abbreviations:

$$
\begin{aligned}
f(r) = 1 - 2m/r \quad &, \quad g_\pm(\rho) = (1 \pm m/2\rho)^2 \\
F(r) = 32m^3 r^{-1} \mathrm{e}^{-r/2m} \quad &, \quad r(\tau,\rho) = [3\sqrt{2m}(\rho - \tau)/2]^{2/3} \\
F_I(x,y) = 8my^2/(xy+2m) \quad &, \quad r(r^*) \text{ from } r^* = r + 2m\ln(r/2m - 1) \\
r(\tau,R) \quad \text{and} \quad r' = \partial r/\partial R \quad & \text{from} \quad (27.59) \\
r(T,X) \quad & \text{from} \quad X^2 - T^2 = (r/2m - 1)\mathrm{e}^{r/2m}
\end{aligned}
\tag{27.213}
$$

28 Interlude: Carter-Penrose Conformal Diagrams

28.1 Introduction

Quite generally, the ability to visualise or depict complex situations plays an important role in developing physical intuition in such a setting. However, clearly curved 4-dimensional space-times provide a challenge for every-day visualisation techniques, and even relatively simple and highly symmetric space-times are often difficult to visualise in a reliable way (just think of the rich structure that we uncovered when analysing the Schwarzschild geometry and its Kruskal-Szekeres maximal extension). This is true in particular for asymptotic or global aspects of a space-time (after all, in all the pictures we have drawn so far this asymptotic region is infinitely far away).

An extremely useful (and widely and commonly used) method to visualise both the causal and the global structure of a (sufficiently symmetric) space-time is that of *Carter-Penrose Conformal Diagrams* (or *Penrose Diagrams* for short in the following, with apologies to B. Carter), already briefly alluded to at the end of section 27 (see Figure 21).

In this section I will introduce and explain these Penrose diagrams by way of some elementary examples.[91] I will not, however, enter into the underlying (and highly technical) issues regarding the proper definition of *(weakly) asymptotically simple* or *asymptotically flat* space-times, say.[92] Other examples will appear later on here and there in these notes.

In order to be able to convey complete information about a 4-dimensional space-time with a single (1+1)-dimensional diagram, we need to be able to suppress 2 dimensions without loss of information. This is e.g. the case for spherically symmetric space-times where very little information is lost by suppressing the angular directions, and we will focus on this case in the following (for less symmetric space-times one would need to look at different 2-dimensional slices of the space-time).

In order to capture both the global and the causal structure of such a space-time in a (1+1)-dimensional diagram of finite extent, the basic idea is to find a coordinate transformation

- such that "infinity" lies at a finite coordinate distance
- and such that (radial) lightrays are always at $\pm 45°$ (as e.g. in the Kruskal diagram)

[91] For introductions to Penrose diagrams, see e.g. Appendix H of S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, or section 2.4 of P. Townsend, *Black Holes*, `arXiv:gr-qc/9707012v1`, or section 2 of A. Strominger, *Les Houches Lectures on Black Holes*, `arXiv:hep-th/9501071`.

[92] The mathematical aspects are discussed in detail in S. Hawking, G. Ellis, *The large scale structure of space-time* and in section 11 of R. Wald, *General Relativity*.

How to implement the 1st requirement is already nicely illustrated by the *conformal compactification* of the Euclidean plane mentioned in section 11.3 (see the discussion around (11.65)):

Starting with the Euclidean metric in Cartesian or polar coordinates,

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2 \tag{28.1}$$

and introducing the new radial coordinate θ through

$$r = \tan\theta/2 \ , \tag{28.2}$$

the metric is mapped to

$$dr^2 + r^2 d\phi^2 = \frac{1}{4\cos^4\theta/2}(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{28.3}$$

Thus the metric is conformal to the standard metric on the 2-sphere, the entire range $0 \leq r < \infty$ of r is mapped to the finite range $0 \leq \theta < \pi$, and infinity $r = \infty$ is mapped to the south pole $\theta = \pi$ of the sphere.

The conformal prefactor diverges as $\theta \to \pi$, as required by the fact that infinity in $\mathbb{R}^2$ is indeed at infinite proper distance. However, if one is willing to sacrifice an accurate representation of distances (and this sacrifice is clearly required if one wants to bring infinity to a finite distance), then one may just as well remove the conformal factor and consider the rescaled metric

$$d\tilde{s}^2 = 4(\cos^4\theta/2)ds^2 = d\theta^2 + \sin^2\theta d\phi^2 \ . \tag{28.4}$$

With respect to the new metric with line element $d\tilde{s}^2$, the point $\theta = \pi$ is now not only at finite coordinate distance but also at finite (affine) geodesic distance, and adding it conformally compactifies $\mathbb{R}^2$ to S^2.

Analogous coordinate transformations (involving the tan function and related objects) and conformal rescalings of the metric are commonly used to "compactify" the range of non-compact coordinates of some space-time metric and to then construct Penrose diagrams. In this case, however, we also need to pay attention to the 2nd requirement, related to the causal structure.

28.2 Causal Structure and Conformal Rescalings of the Metric

Regarding the relation between the causal structure and conformal (or Weyl) rescalings of the metric,

$$g_{\alpha\beta}(x) \to \tilde{g}_{\alpha\beta}(x) = \Omega(x)^2 g_{\alpha\beta}(x) \tag{28.5}$$

or

$$ds^2 \to d\tilde{s}^2 = \Omega(x)^2 ds^2 \ , \tag{28.6}$$

we just note the following facts:

1. The causal nature of a vector field or curve is invariant under conformal rescalings, i.e. a vector field is spacelike with respect to $\tilde{g}_{\alpha\beta}$ iff it is spacelike with respect to $g_{\alpha\beta}$, a curve is everywhere timelike with respect to $\tilde{g}_{\alpha\beta}$ iff it is everywhere timelike with respect to $g_{\alpha\beta}$, etc.

2. In particular, conformal rescalings preserve the lightcones ($ds^2 = 0$) and thus the causal structure of the space-time encoded in the structure and behaviour of lightcones.

3. In general, even though timelike or spacelike paths are mapped into timelike or spacelike paths, timelike or spacelike geodesics are *not* mapped into each other. However, the paths that are traced out by null geodesics *are* mapped into each other under conformal rescalings (albeit with respect to different, and therefore in general non-affine, parametrisations).

The first 2 assertions are obvious. Thus the only one that may require an explanation is the 3rd. This follows directly from the relation between the Christoffel symbols of the 2 metrics which is easily seen to be

$$\tilde{\Gamma}^{\alpha}_{\ \gamma\delta} = \Gamma^{\alpha}_{\ \gamma\delta} + \Omega^{-1}(\delta^{\alpha}_{\gamma}\nabla_{\delta}\Omega + \delta^{\alpha}_{\delta}\nabla_{\gamma}\Omega - g_{\gamma\delta}\nabla^{\alpha}\Omega) \equiv \Gamma^{\alpha}_{\ \gamma\delta} + \Delta^{\alpha}_{\ \gamma\delta} \ . \tag{28.7}$$

This implies that for a curve $x^{\alpha} = x^{\alpha}(\tau)$ one has

$$\ddot{x}^{\alpha} + \tilde{\Gamma}^{\alpha}_{\ \gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta} = \ddot{x}^{\alpha} + \Gamma^{\alpha}_{\ \gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta} + 2\Omega^{-1}\dot{\Omega}\dot{x}^{\alpha} - \Omega^{-1}\nabla^{\alpha}\Omega(g_{\gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta}) \ , \tag{28.8}$$

where

$$\dot{\Omega} = \dot{x}^{\gamma}\partial_{\gamma}\Omega = d\Omega/d\tau \ . \tag{28.9}$$

In particular, if $x^{\alpha}(\tau)$ is an affinely parametrised geodesic for the metric $g_{\alpha\beta}(x)$, then it also satisfies the equation

$$\ddot{x}^{\alpha} + \tilde{\Gamma}^{\alpha}_{\ \gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta} = 2\Omega^{-1}\dot{\Omega}\dot{x}^{\alpha} - \Omega^{-1}\nabla^{\alpha}\Omega(g_{\gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta}) \ . \tag{28.10}$$

If $g_{\gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta} \neq 0$ (i.e. for timelike or spacelike geodesics), this is *not* the geodesic equation for the metric $\tilde{g}_{\alpha\beta}(x)$, the last term on the right hand side playing the role of a force term. For null geodesics, on the other hand, one has

$$\ddot{x}^{\alpha} + \tilde{\Gamma}^{\alpha}_{\ \gamma\delta}\dot{x}^{\gamma}\dot{x}^{\delta} = 2\Omega^{-1}\dot{\Omega}\dot{x}^{\alpha} \ , \tag{28.11}$$

and since the right hand side is proportional to $\dot{x}^{\alpha}$, this *is* the geodesic equation, albeit with respect to a non-affine parametrisation (cf. (2.136) and the discussion in section 2.7), with inaffinity

$$\kappa(\tau) = 2\Omega^{-1}\dot{\Omega} = (d/d\tau)(\log\Omega^2) \ . \tag{28.12}$$

This null geodesic will then be affinely parametrised with respect to the parameter $\tilde{\tau}$ determined along the null geodesic by the relation (2.138)

$$\frac{d\tilde{\tau}}{d\tau} = \mathrm{e}^{\int^{\tau} dt\ \kappa(t)} = \Omega(x(\tau))^2 \tag{28.13}$$

Note that this relation $d\tilde{\tau} = \Omega^2 d\tau$ between the null affine parameters is not what one might have naively (but incorrectly) expected or extrapolated from the relation $d\tilde{s} = \Omega ds$ between the proper spatial distance or proper time intervals in the 2 metrics.

Somewhat more covariantly and succinctly, the above disucssion can be rephrased as follows:

$$
\begin{array}{rrcl}
 & \tilde{g}_{\alpha\beta} = \Omega^2 g_{\alpha\beta} & \Rightarrow & \tilde{\Gamma}^{\alpha}_{\ \gamma\delta} = \Gamma^{\alpha}_{\ \gamma\delta} + \Delta^{\alpha}_{\ \gamma\delta} \\
 & g_{\alpha\beta}\ell^\alpha \ell^\beta = 0 & \Rightarrow & \ell^\alpha \tilde{\nabla}_\alpha \ell^\beta = \ell^\alpha \nabla_\alpha \ell^\beta + 2\Omega^{-1}\dot{\Omega}\,\ell^\alpha \qquad (\dot{\Omega} \equiv \ell^\beta \partial_\beta \Omega) \\
 & \tilde{\ell}^\alpha = \Omega^{-2}\ell^\alpha & \Rightarrow & \tilde{\ell}^\alpha \tilde{\nabla}_\alpha \tilde{\ell}^\beta = \Omega^{-4}\ell^\alpha \nabla_\alpha \ell^\beta \\
\Rightarrow & \ell^\alpha \nabla_\alpha \ell^\beta = 0 & \Leftrightarrow & \tilde{\ell}^\alpha \tilde{\nabla}_\alpha \tilde{\ell}^\beta = 0
\end{array}
\tag{28.14}
$$

28.3 Penrose Diagram for (3+1) Minkowski Space

We will now see how to accomplish the desiderata laid out at the beginning of this section in the simplest example, namely (3+1)-dimensional Minkowski space-time. At first sight the (1+1)-dimensional case may appear to be an even simpler example. However, because of the absence of an honest spatial radial direction in that case, it turns out to be somewhat atypical, and therefore does not constitute the optimal starting point. I will therefore discuss the (1+1)-dimensional case separately in section 28.4 below.

In order to exhibit the spherical symmetry of (3+1)-dimensional Minkowski space, and to isolate the spherical part of the metric, we start off with the Minkowski metric written in spatial spherical coordinates,

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega^2 \ , \tag{28.15}$$

with $-\infty < t < +\infty$ and $0 \leq r < +\infty$.

One simple-minded way to map the infinite coordinate ranges to a finite range would be to introduce, in analogy with the Euclidean case above, a new radial coordinate R related to r via a tan function, and likewise for t, along the lines of

$$t = \tan T \quad , \quad r = \tan R \qquad (???) \ . \tag{28.16}$$

However, while this accomplishes the 1st desideratum (finite range of coordinates), it fails to satisfy the 2nd requirement (lightcones at 45°). Indeed, using $dt = dT/\cos^2 T$ etc., one finds that the (t, r)-part of the metric takes the form

$$-dt^2 + dr^2 = -dT^2/\cos^4 T + dR^2/\cos^4 R \ . \tag{28.17}$$

Thus radial lightrays would be described by

$$dT/dR = \pm \cos^2 T/\cos^2 R \ , \tag{28.18}$$

and evidently have a slope that depends on (T, R) (whereas we would like $dT/dR = \pm 1$).

In order to rectify this, we will first introduce coordinates that are adapated to radial in- and outgoing lightrays, namely the coordinates

$$u = t - r \quad , \quad v = t + r \quad , \tag{28.19}$$

in terms of which the metric has the "double-null" form

$$ds^2 = -du\,dv + ((v-u)^2/4)d\Omega^2 \quad , \tag{28.20}$$

with $-\infty < u, v < \infty$ and (because of $r \geq 0$) $u \leq v$, i.e.

$$-\infty < u \leq v < +\infty \quad . \tag{28.21}$$

Radial lightrays are described by

$$du\,dv = 0 \quad \Rightarrow \quad u = u_0 \quad \text{or} \quad v = v_0 \quad . \tag{28.22}$$

Lines of constant u describe outgoing radial lightrays while lines of constant v describe ingoing radial lightrays.

In terms of these and the original coordinates we can now identify different "infinities' (asymptotic regions) of Minkowski space-time, namely (in standard notation)

- i^+ (*future timelike infinity*): where one asymptotes to when one takes $t \to +\infty$ at fixed r
- i^- (*past timelike infinity*): where one asymptotes to when one takes $t \to -\infty$ at fixed r
- i^0 (*spacelike infinity*): where one asymptotes to when one instead takes $r \to \infty$ at fixed t
- $\mathcal{I}^+$ (*future null infinity*): where outgoing radial lightrays asymptote to in the future, i.e. one takes $v \to \infty$ at fixed u
- $\mathcal{I}^-$ (*past null infinity*): where ingoing radial lightrays asymptote to in the past, i.e. one takes $u \to -\infty$ at fixed v

Here $i^\pm, i^0$ are pronounced as "eye-plus, eye-minus, eye-zero", while $\mathcal{I}^\pm$ are pronounced as "screye-plus, screye-minus" (with "screye" derived from "script I"). As mentioned before, defining these objects properly, however, even in the case at hand and *a fortiori* in somewhat more generality, i.e. for appropriately defined asymptotically flat space-times, requires significantly more care (cf. the references given at the beginning of this section in footnote 92).

In order to be able to indicate these different asymptotic regions in a diagram, we now introduce new coordinates U and V. The double-null structure of the metric (and its associated simple description of in- and outgoing lightrays) is preserved by arbitrary transformations

$$u = u(U) \quad , \quad v = v(V) \tag{28.23}$$

of the lightcone coordinates. One possible choice which maps the symptotic regions to finite values of the new coordinates is the (by now unsurprising) tan transformation

$$u = \tan U \quad , \quad v = \tan V \quad . \tag{28.24}$$

The range of the coordinates (U, V) is

$$-\pi/2 < U \leq V < +\pi/2 \quad , \tag{28.25}$$

so this definitely describes a finite region in the (U, V)-plane, namely a triangle. Infinity corresponds to the locus $|U| \to \pi/2$ and/or $|V| \to \pi/2$.

Another (and equally good) choice could have been

$$\tilde{U} = \tanh u \quad , \quad \tilde{V} = \tanh v \quad , \tag{28.26}$$

say, with

$$-1 < \tilde{U} \leq \tilde{V} < +1 \tag{28.27}$$

but let us continue to work with the coordinates (U, V) defined in (28.3). In terms of these the metric (after some elementary trigono-gymnastics) takes the form

$$ds^2 = \frac{1}{4\cos^2 U \cos^2 V}\left(-4dU\; dV + \sin^2(V - U)d\Omega^2\right) \quad . \tag{28.28}$$

Note in particular that the prefactor diverges as one approaches infinity, in agreement with the evident fact that with respect to this metric infinity is at infinite proper distance even though it is at finite coordinate distance.

However, if our interest is in the global and causal structure of the metric, while disregarding the proper distance structure also encoded in the metric, we can just remove this prefactor. This will not change the fact that in/out radial lightrays are described by lines of constant V or U, but allows us to extend the metric to include the boundary points at which the prefactor diverges. Thus we consider the metric

$$d\tilde{s}^2 = (4\cos^2 U \cos^2 V)ds^2 = -4dU\; dV + \sin^2(V - U)d\Omega^2 \quad . \tag{28.29}$$

This can be put into a more familiar form by replacing the lightcone coordinates U and V by new time and radial coordinates T and R via the analogue of $u = t - r, v = t + r$ (28.19), namely

$$T = U + V \quad , \quad R = V - U \geq 0 \quad . \tag{28.30}$$

Then the metric is

$$d\tilde{s}^2 = -dT^2 + dR^2 + \sin^2 R\; d\Omega^2 \quad . \tag{28.31}$$

and the combined transformation from the original coordinates (t, r) to these coordinates (T, R) is

$$t \pm r = \tan \tfrac{1}{2}(T \pm R) \quad . \tag{28.32}$$

Before proceeding to draw the appropriate (1+1)-dimensional diagram for this (by suppressing the spherical / angular directions), let me make some comments on this (3+1)-dimensional metric.

<u>Remarks:</u>

1. If T had the range $-\infty < T < +\infty$ and R were a standard polar angular coordinate ψ, then this would be the standard metric on $\mathbb{R} \times S^3$, a space-time given by the direct product of the time direction and a spatial 3-sphere of constant unit radius,

$$d\tilde{s}^2 = -dT^2 + d\Omega_3^2 \quad , \tag{28.33}$$

with (2.18)

$$d\Omega_3^2 = d\psi^2 + \sin^2 \psi \, d\Omega_2^2 \quad . \tag{28.34}$$

2. This (in the present context unphysical) metric, regarded as a solution of the Einstein equations, happens to have a name, namely the *Einstein Static Universe* (ESU), and happens to be of some historical interest (because finding such a static "cosmological" solution motivated Einstein to introduce the infamous cosmological constant into his equations in the first place). For this reason, we will briefly discuss this solution in the context of cosmology in section 37.2. However, for present purposes this is just an unnecessary distraction.

3. In the current case of interest this is in any case not the range of the coordinates. Rather, the triangular bound (28.25) on the coordinates U, V translates into the conditions

$$|T| + R < \pi \quad , \quad 0 \leq R < \pi \tag{28.35}$$

on the range of the coordinates T, R. Thus, if one likes one can think of Minkowski space as being conformally equivalent to the subspace of $\mathbb{R} \times S^3$ defined by these conditions. Combined with the previous comment one can thus think of Minkowski space as conformally equivalent to a subspace of the ESU, and pictorial representations of this (with the ESU represented by the cylinder $\mathbb{R} \times S^1$) can be found in many places (including all but one of the references in footnote 91). I have never found this particularly illuminating, however.

For this reason we will now just focus on the (1+1)-dimensional metric

$$d\tilde{s}^2 = -4dU \, dV = -dT^2 + dR^2 \tag{28.36}$$

with the coordinate ranges given in (28.25) and (28.35) respectively. In a (U, V)-diagram (with the U-axis vertical, say, and the V-axis horizontal), this is evidently just the lower right triangular half of a square of length π centered at the origin. In terms of (T, R), the apex of this triangle at $U = \pi/2, V = -\pi/2$ is mapped to $R = V - u = \pi$ and $T = V + U = 0$. Thus this corresponds to a counter-clockwise rotation of the triangle by $\pi/4$, and we therefore obtain Figure 22.

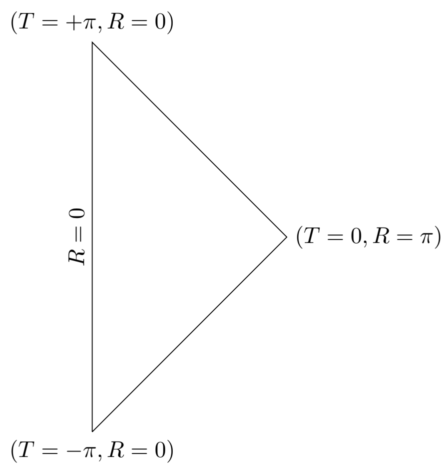


Figure 22: Towards the Penrose Diagram for Minkowski space: Minkowski space corresponds to the interior of the triangle, including the line $R = 0$ but excluding the diagonal boundary lines and their endpoints.

As indicated there, Minkowski space is conformally equivalent to the interior of the triangle (including the vertical line $R = 0 \leftrightarrow r = 0$). All points in the interior (except at $r = 0$) represent 2-spheres. The other boundaries are in precise correspondence with the "infinities" discussed before.

For example, the boundary point $U = -\pi/2, V = \pi/2$ or $T = 0, R = \pi$ corresponds to

$$(T \to 0, R \to \pi) \quad \Leftrightarrow \quad (u \to -\infty, v \to +\infty) \quad \Leftrightarrow \quad (t \text{ finite}, r \to \infty) \tag{28.37}$$

and thus to spacelike infinity i^0,

$$i^0 : \quad (T = 0, R = \pi) \ . \tag{28.38}$$

Likewise, the point $U = V = \pi/2$ or $R = 0, T = \pi$ corresponds to

$$(T \to \pi, R \to 0) \quad \Leftrightarrow \quad (u \to +\infty, v \to +\infty) \quad \Leftrightarrow \quad (t \to +\infty, r \text{ finite}) \tag{28.39}$$

and thus to future timelike infinity i^+,

$$i^+ : \quad (T = \pi, R = 0) \ , \tag{28.40}$$

and likewise for past timelike infinity i^-,

$$i^- : \quad (T = -\pi, R = 0) \ , \tag{28.41}$$

Future null infinity $\mathcal{I}^+$ is characterised by

$$(u \text{ finite}, v \to \infty) \quad \Leftrightarrow \quad (|U| < \pi/2, V = \pi/2) \quad \Leftrightarrow \quad T + R = 2V = \pi \tag{28.42}$$

which is the upper diagonal line in Figure 22, and likewise for past null infinity,

$$\mathcal{I}^{\pm}: \quad \{(T,R) : R \pm T = \pi, 0 < R < \pi\} \ . \tag{28.43}$$

By simply adding these regions and this information to the diagram in Figure 22, we obtain our final version of the Penrose diagram of Minkowski space, Figure 23.

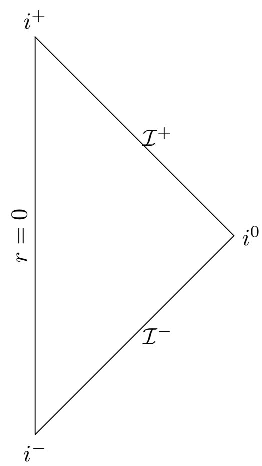


Figure 23: Penrose Diagram for Minkowski space

To get acquainted with this diagram (and with Penrose diagrams in general), let us note the following facts:

1. radial null geodesics / lightcones are at $\pm 45°$
2. all points in the diagram except those at $r = 0$, and $i^{\pm}, i^0$ represent 2-spheres
3. $i^{\pm}, i^0$, on the other hand, are really points because the "radius" $\sin R$ of the 2-sphere vanishes at the poles $R = 0, \pi$;
4. $\mathcal{I}^{\pm}$ are null hypersurfaces with topology $S^2 \times \mathbb{R}$;
5. all (infinitely extended) timelike geodesics begin at i^- and end at i^+;
6. all (infinitely extended) spacelike geodesics begin at i^0, pass through ("are reflected" at) $r = 0$ and end again at i^0;
7. all (infinitely extended) null geodesics begin on $\mathcal{I}^-$, are reflected at $r = 0$ and end at $\mathcal{I}^+$.

While we cannot expect to learn too much about Minkowski space from this diagram that we did not already know, understanding the asymptotic structure of Minkowski space-time will be useful in the following, because any reasonable definition of an *asymptotically flat* space-time representing the gravitational field of an isolated object should be such that asymptotically it looks like Minkowski space, i.e. its Penrose diagram should asymptotically resemble that of Figure 23.

Moreover, this pictorial representation is also interesting in its own right since it makes causal information easily accessible and, in particular, makes two features of Minkowski space manifest that are not shared by all space-times:

1. Any timelike geodesic observer will eventually be able to see all of Minkowski space (since eventually, at i^+, the past lightcone of the observer covers all of Minkowski space).
2. Past and future lightcones of any 2 events intersect. In particular, any 2 events in Minkowski space were causally connected at some time in the past.

The reasons for emphasising these properties here are that

- characteristically, black hole space-times are such that observers at infinity do not have access to all of space-time since they cannot see behind the event horizon (cf. the discussion in section 28.5 below, and the more general discussion in section 32),
- cosmological space-times typically also fail to have at least one of these properties, and this is ususally characterised in terms of so-called cosmological particle and event horizons (cf. the discussion in section 36.7).

28.4 Penrose Diagram for (1+1) Minkowski Space and Rindler Observers

Let us now look at (1+1)-dimensional Minkowski space, with metric

$$ds^2 = -dt^2 + dx^2 \quad , \tag{28.44}$$

and its Penrose diagram. The only (but crucial) difference to the radial part of (3+1)-dimensional Minkowski space,

$$ds^2 = -dt^2 + dr^2 \quad , \tag{28.45}$$

is of course that the coordinate range of x is $-\infty < x < +\infty$ while that of r is $0 \leq r < +\infty$. In particular, there are now two (right and left) spacelike infinities, i^0_R and i^0_L, corresponding to $x \to \pm\infty$. Moreover, while as before we can introduce lightcone coordinates, which we now call $x^\pm$, via

$$x^\pm = t \pm x \quad , \tag{28.46}$$

their interpretation is now different: instead of describing in- and outgoing lightrays, lines of constant x^- describe *right-moving* lightrays while lines of constant x^+ describe *left-moving* lightrays. Thus there are corresponding left and right future and past null infinities $\mathcal{I}_L^\pm$ and $\mathcal{I}_R^\pm$. Common sense and/or the analogue

$$x^\pm = \tan X^\pm = \tan \tfrac{1}{2}(T \pm X) \tag{28.47}$$

of the coordinate transformation (28.32) now shows that this space-time can be represented in a Penrose diagram by doubling the triangle of Figure 23 to a diamond with left and right asymptotic regions (Figure 24).

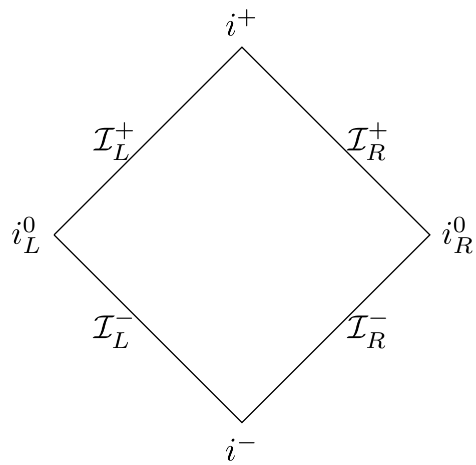


Figure 24: Penrose Diagram for (1+1)-dimensional Minkowski space

It is instructive to reconsider the constantly accelerating Rindler observer, discussed in detail in sections 1.3 and 3.4, from this perspective. In section 1.3 we had introduced Rindler coordinates adapated to such an observer which cover the right (and/or left) Rindler wedge of Minkowski space. The right Rindler wedge is the grey shaded area in Figure 25. Also indicated there is the worldline of a Rindler observer.

This diagram also illustrates that non-geodesic timelike worldlines that are not geodesics do not necessarily end up at future timelike infinity i^+ but can end anywhere on future null infinity $\mathcal{I}^+$ (provided that there is enough acceleration).

28.5 Penrose Diagram for Schwarzschild (Kruskal-Szekeres)

We now come to the Schwarzschild metric. We already know, from our detailed investigations in section 27, that a very convenient global picture of the Schwarzschild space-time is provided by the maximal Kruskal-Szekeres extension of the Schwarzschild metric and the resulting Kruskal diagram (sections 27.7 and 27.8). And indeed a convenient starting point for constructing the Penrose version of the Kruskal diagram is the

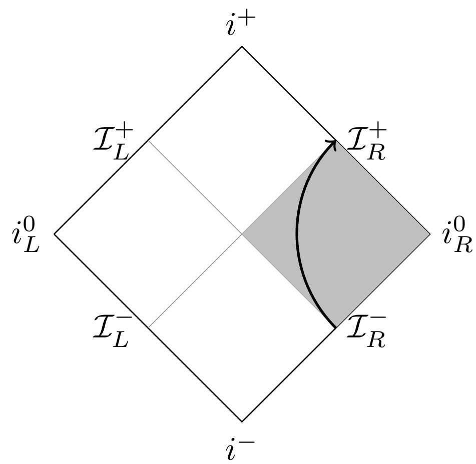


Figure 25: Penrose Diagram for (1+1)-dimensional Minkowski space, showing the right Rindler wedge (the grey shaded area) and the worldline of a Rindler observer.

"double-null" form (27.146)

$$ds^2 = -\frac{32m^3}{r}\mathrm{e}^{-r/2m}dU\,dV + r(U,V)^2 d\Omega^2 \quad , \tag{28.48}$$

of the Schwarzschild metric in Kruskal coordinates.

We will come back to this below. However, it is instructive to go back a step and first start with the more modest aim of constructing a Penrose representation of the "Schwarzschild patch" (region I) of the space-time. As in the case of Minkowski space, we start by introducing coordinates that are adapted to radial lightrays. These are the advanced and retarded Eddington-Finkelstein coordinates u and v, related to the Schwarzschild time coordinate t and the tortoise coordinate r^* (section 26.7) by (27.78)

$$u = t - r^* \quad , \quad v = t + r^* \quad . \tag{28.49}$$

In terms of these the (t, r)-part of the Schwarzschild metric can be written as (27.141)

$$ds^2 = -(1 - 2m/r)du\,dv \tag{28.50}$$

where $r = r(u, v)$. In these coordinates, the asymptotic regions $v \to +\infty$ at fixed u and $u \to -\infty$ at fixed v have the same interpretation as in Minkowski space, namely as future and past null infinity $\mathcal{I}^\pm$,

$$\begin{aligned} \mathcal{I}^+ : &\quad (u \text{ finite}, v \to +\infty) \\ \mathcal{I}^- : &\quad (u \to -\infty, v \text{ finite}) \quad . \end{aligned} \tag{28.51}$$

Because the range of the tortoise coordinate r^* is

$$-\infty < r^* < +\infty \tag{28.52}$$

as r ranges over $2m < r < +\infty$, there are also two other "asymptotic" regions, much as in the case of (1+1)-dimensional Minkowski space discussed above. Here, however, crucially and characteristically, their interpretation is quite different. Namely, as we have seen in sections 27.5 and 27.7, these are the future and past *horizons* of the Schwarzschild black hole at $r = 2m$, now denoted by $\mathcal{H}^{\pm}$,

$$
\begin{aligned}
\mathcal{H}^+ : &\quad (u \to +\infty, v \text{ finite}) \\
\mathcal{H}^- : &\quad (u \text{ finite}, v \to -\infty) \quad .
\end{aligned}
\tag{28.53}
$$

We can now map the entire (u,v)-plane to a finite (diamond) region by introducing the coordinates $\tilde{U}, \tilde{V}$ (not to be confused with the Kruskal coordinates U, V) via

$$
u = \tan \tilde{U} \quad , \quad v = \tan \tilde{V}
\tag{28.54}
$$

with $|\tilde{U}| < \pi/2$ and $|\tilde{V}| < \pi/2$, and we can then depict the Schwarzschild patch as in Figure 26.

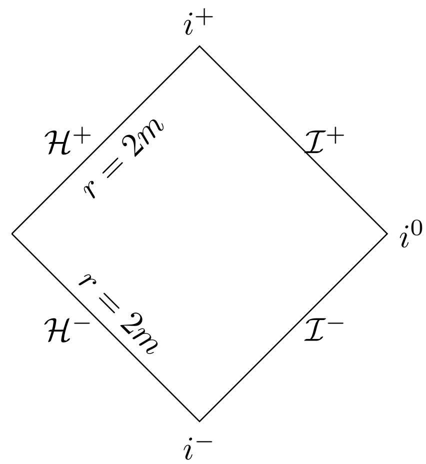


Figure 26: Penrose-diagrammatic representation of the Schwarzschild patch.

This already teaches us what the Schwarzschild patch of the Kruskal diagram (Figure 20) will look like in a Penrose diagram of the maximal Kruskal-Szekeres extension of the Schwarzschild geometry. In order to extend this description beyond the horizons $\mathcal{H}^{\pm}$, one can of course now switch to Kruskal coordinates (U, V) and introduce new coordinates $\hat{U}, \hat{V}$ from the Kruskal coordinates U, V via the (by now familiar) transformation

$$
U = \tan \hat{U} \quad , \quad V = \tan \hat{V} \quad ,
\tag{28.55}
$$

and one can follow this up by the (by now equally familiar, cf. (28.30)) transformation to new time and radial coordinates.

$$
\hat{T} = \hat{U} + \hat{V} \quad , \quad \hat{R} = \hat{V} - \hat{U}
\tag{28.56}
$$

One can of course work also out explicitly the metric in these coordinates. This is straightforward but not really neccessary and I will spare you (and me) the details of this, and will just add some remarks and explanations below. The upshot of this is that then the Penrose $(\hat{T}, \hat{R})$ diagram takes the form displayed in Figure 27.

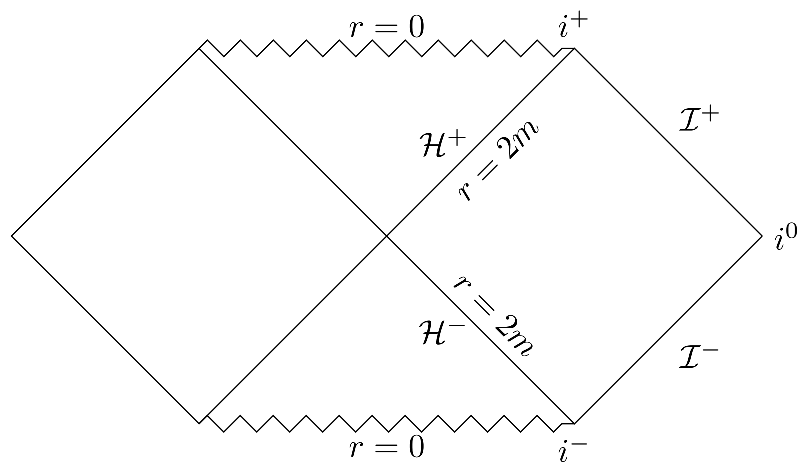


Figure 27: Penrose Diagram of the maximal Kruskal-Szekeres extension of the Schwarzschild space-time.

<u>Remarks:</u>

1. In this diagram I have only labelled the various boundaries and horizons on the right half of the diagram. Evidently the same labels can be pasted onto the mirror left half.

2. The asymptotic structure (in the Schwarzschild patch and its mirror) is precisely that of Minkowski space, in agreement with our intuition that the Schwarzschild metric is asymptotically flat.

3. $\mathcal{I}^+$ corresponds to $\hat{V} = \pi/2$, and $\mathcal{I}^-$ to $\hat{U} = -\pi/2$.

4. The future / past horizons $\mathcal{H}^\pm$ are at $\hat{U} = 0$ and $\hat{V} = 0$ respectively.

5. The future horizon $\mathcal{H}^+$ is (now manifestly) the boundary of the region from which signals can escape to future null infinity $\mathcal{I}^+$.

6. Another way of phrasing this is in terms of the *causal past* of future null infinity $\mathcal{I}^+$, the union of all the space-time points that lie in the past lightcone of some point on $\mathcal{I}^+$: from this perspective, the horizon $\mathcal{H}^+$ is the boundary of (the closure of) the past of future null infinity. We will come back to in section 32.

7. The singularities at $r = 0$ are (here, and typically in general) indicated by wavy or zig-zag lines. In Kruskal coordinates, they correspond to the hyperbolae $UV = 1$,

and thus to

$$\tan \hat{U} \tan \hat{V} = 1 \quad \Leftrightarrow \quad \cos(\hat{U} + \hat{V}) = 0 \ . \tag{28.57}$$

Thus in terms of the time and radial coordinates $\hat{T}$ and $\hat{R}$ (28.56) this is simply the locus

$$\cos(\hat{T}) = 0 \quad \Leftrightarrow \quad \hat{T} = \pm\pi/2 \ . \tag{28.58}$$

This accounts for the fact that in the Penrose diagram the singularities are now represented by straight horizontal lines.

It is instructive to compare the Penrose diagram for the Schwarzschild metric with that for the negative mass $m = -|m|$ Schwarzschild metric, which we write as

$$ds^2 = -\left(1 + \frac{2|m|}{r}\right) dt^2 + \left(1 + \frac{2|m|}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \ . \tag{28.59}$$

In this case the singularity at $r = 0$ is timelike, not hidden behind an event horizon, and therefore naked. The coordinates are valid all the way to $r = 0$, and thus the Penrose diagram (Figure 28) looks deceptively like that of Minkowski space (Figure 23), the crucial difference of course being that the vertical line $r = 0$ now represents a singularity, visible all the way to $\mathcal{I}^+$.

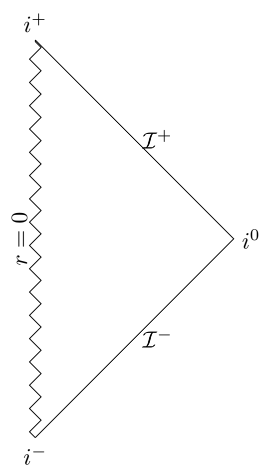


Figure 28: Penrose Diagram for the negative mass Schwarzschild metric. The singularity at $r = 0$ is timelike and not hidden behind an event horizon.

28.6 Penrose Diagrams: List of Other Examples

Here is a list of other examples of Penrose diagrams that appear elsewhere in these notes:

- Figure 29 in section 29.1: worldline of a collapsing null shell in Minkowski space

- Figure 30 in section 29.1: worldline of a collapsing null shell in the maximally extended Schwarzschild geometry
- Figure 31 in section 29.1: collapse of a thin null shell to a black hole
- Figure 33 in section 29.3: gravitational collapse of a star to a black hole
- Figure 34 in section 31.6: maximal analytic extension of the Reissner-Nordstrøm metric
- Figure 35 in section 31.8: regimes of validity of Eddington-Finkelstein coordinates for the Reissner-Nordstrøm metric
- Figure 36 in section 31.8: extremal Reissner-Nordstrøm metric
- Figure 37 in section 31.10: regime of validity of Kruskal-Szekeres coordinates for the Reissner-Nordstrøm metric
- Figure 38 in section 32.4: essential (future horizon) part of a (Schwarzschild) black hole
- Figure 39 in section 32.4: illustration of the definition of an event horizon
- Figure 40 in section 32.10: event horizon vs apparent horizon for the Vaidya metric
- Figure 41 in section 32.11: event horizon vs apparent horizon for the collapsing null shell
- Figure 42 in section 32.12: event horizon vs apparent horizon for the Oppenheimer-Snyder gravitational collapse
- Figure 43 in section 32.13: location of (outer) trapped surfaces for the collapsing null shell
- Figure 47 in section 36.6: spatially flat cosmological solution with constant expansion velocity.
- Figure 48 in section 36.6: spatially flat decelerating and accelerating cosmological solutions.
- Figure 51 in section 37.3: conformal diagram of the $k = +1$ matter dominated universe.
- Figure 52 in section 37.4: conformal diagram of the $k = +1$ radiation dominated universe.
- Figure 53 in section 38.3: illustration of the horizon problem in cosmology.
- Figures 54 - 57 in section 39.2.2: various Penrose diagrams for de Sitter space.

- Figure 58 in section 39.3.2: Penrose diagram for anti-de Sitter space.
- Figures 59 - 61 in section 42.3: Penrose diagrams for the linear mass Vaidya metric.

29 Black Holes III: Simple Models of Gravitational Collapse

Now you may well wonder if all this talk about white holes and mirror regions in the previous sections is for real or just science fiction. Clearly, if an object with $r_0 < 2m$ (figuratively speaking) exists and is described by the Schwarzschild solution, then we will have to accept the conclusions of the previous section.

However, this requires the existence of an eternal black hole (in particular, eternal in the past) in an asymptotically flat space-time, and this is not very realistic. If this were the only way to obtain black holes, then one might be justified in simply regarding them as a mathematical oddity, an unphysical feature permitted by the Einstein equations (much like general relativity does not rule out closed timelike curves and other peculiarities) but having nothing to do with the real world.

Non-eternal black holes are believed to exist, however, because they are believed to form as a consequence of e.g. the gravitational collapse of a star whose nuclear fuel has been exhausted (and which is so massive that it cannot settle into a less singular final state like a *White Dwarf* or *Neutron Star*).

Before trying to understand how we could model such a gravitational collapse of a star (without having to worry about astrophysical issues), we briefly consider a very simple toy model of gravitational collapse and black hole formation.

29.1 Collapse of a Shell of Radiation

The arguably simplest (and very instructive, but highly idealised and unrealistic) model of gravitational collapse to a black hole is provided by a collapsing thin (very thin!) sphericall shell of null matter (radiation) in Minkowski space. Other simple toy-models of gravitational collapse can be based on considering collapsing shells of non-null matter.[93]

Thus we consider the situation where we have an implosion of an infinitely thin spherical shell of radiation in an otherwise empty space-time. This requires somewhat of a conspiracy, of course, but let us assume that we have been nasty enough to arrange this. The assumption that the shell is infinitely thin (delta-function localised) is of course a mathematical idealisation and should be regarded as an approximation to a shell with finite thickness.

In order to describe such incoming radiation (along ingoing null geodesics), it is natural to work in ingoing coordinates that are adapted to such null geodesics. Thus Minkowski space in ingoing coordinates $(v = t + r, r)$ (and the usual coordinates on the sphere)

[93] See e.g. E. Poisson, *A Relativist's Toolkit* or R. Adler, D. Bjorken, P. Chen, J. Liu, *Simple Analytic Models of Gravitational Collapse*, `arXiv:gr-qc/0502040` for pedagogical discussions.

takes the form

$$ds^2 = -dv^2 + 2dv\,dr + r^2 d\Omega^2 \tag{29.1}$$

and the Schwarzschild metric (with $v = t + r^*$) has the ingoing Eddington-Finkelstein form

$$ds^2 = -f(r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} \quad . \tag{29.2}$$

In both metrics, ingoing lightrays are described by lines of constant v.

The relevance of these two metrics for the probem at hand arises from the fact that in the two vacuum regions inside and outside the shell one will have (essentially by Birkhoff's theorem)

- the flat Minkowski geometry inside the shell
- and the Schwarzschild metric outside the shell.

In these adapted ingoing coordinates, we can assume that the shell moves along the ingoing null trajectory $v = v_0$, as viewed from both the internal Minkowski geometry and the external Schwarzschild geometry.

Naively, one can then simply attempt to describe the metric in ingoing Minkowski / Eddington-Finkelstein coordinates by

$$ds^2 = -f(v,r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m_f}{r}\Theta(v - v_0) \ , \tag{29.3}$$

where m_f (or m_f/G_N) is the final / total mass and $\Theta(v)$ is the step function. See section 29.2 for a slightly more detailed justification for this ansatz.

This metric has the form of an ingoing *Vaidya metric*

$$ds^2 = -f(v,r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m(v)}{r} \quad , \tag{29.4}$$

in this particular case with the distributional mass function

$$m(v) = m_f \Theta(v - v_0) \ . \tag{29.5}$$

Vaidya metrics will be briefly mentioned in section 30.2 in an overview of black hole solutions, see (30.17), and then again in sections 32.8 and 32.9 in the context of the discussion of black hole horizons, and will be be discussed in some detail in sections 40 - 42.

Calculating the Einstein tensor of this metric (or using e.g. (40.9)), one finds that this is a solution of the Einstein equations with an energy-momentum tensor whose only non-vanishing component is

$$T_{vv} = \frac{1}{4\pi G_N}\frac{m_f}{r^2}\delta(v - v_0) \ . \tag{29.6}$$

This describes purely ingoing radiation, localised along the null world volume of the shell, with constant total mass $M = m_f/G_N$, as desired and expected,

It is clear that at some point the radius of the shell (moving along the line $v = v_0$ in the direction of decreasing r) will reach and then cross its Schwarzschild radius. Once that has happened, the exterior Schwarzschild geometry (covering the Schwarzschild patch as well as the region outside the shell but inside the Schwarzschild radius) describes a black hole with a future event horizon. However, there is no trace here of either the mirror region III or the white hole region IV.

This is best understood by looking at the Penrose diagram for this solution. Let us first just draw the null worldline of the shell in the Penrose diagram of Minkowski space (Figure 23). This is given in Figure 29.

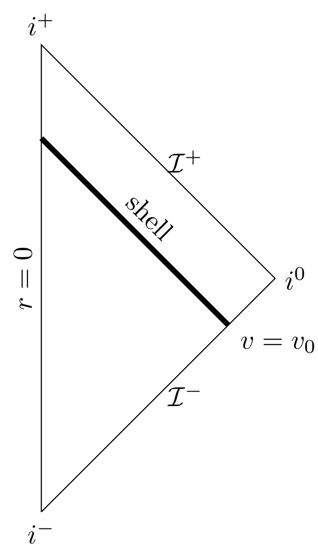


Figure 29: Penrose Diagram indicating the worldline of a null shell in Minkowski space. The worldline of the shell is given by the line $v = v_0$. Only the interior part (below the line $v = v_0$) is displayed correctly in this diagram.

However, this does not yet describe correctly the gravitational field / geometry of this situation. Inside the shell (i.e. below the line $v = v_0$), the geometry is indeed that of Minkowski space, so this part of the diagram is correct. However, outside of the shell (above the line $v = v_0$) the geometry is the Schwarzschild geometry, and there is a singularity when the shell collapses to zero size (reaches $r = 0$).

To see what this amounts to we can also add the worldline of the shell to the Penrose diagram version of the Kruskal diagram (Figure 27). In this way we arrive at Figure 30.

In this diagram, only the geometry outside the shell (i.e. above the line $v = v_0$) is displayed correctly (as that of the Schwarzschild space-time), while the inside should be replaced by Minkowski space. The correct diagram is thus obtained by gluing the

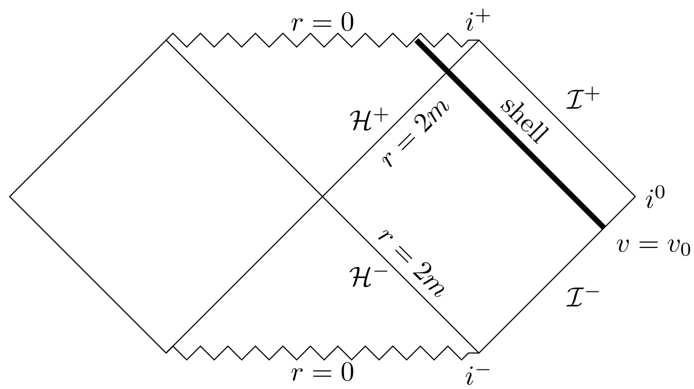


Figure 30: Penrose Diagram indicating the worldline of a null shell in the maximally extended Schwarzschild geometry. The worldline of the shell is given by the line $v = v_0$. Only the exterior part (above the line $v = v_0$) is displayed correctly in this diagram.

two Penrose diagrams in Figures 29 and 30 together along the worldline of the shell. As a consequence, the mirror region III and the white hole region IV get exorcised from the diagram (as well as part of the black hole region II). In this way one arrives at the diagram in Figure 31.

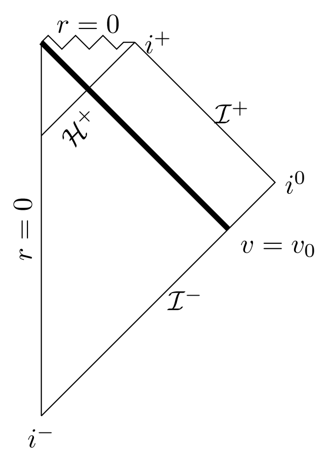


Figure 31: Penrose Diagram for the collapse of a thin null shell to a black hole. The worldline of the shell is given by the line $v = v_0$. In the region $v < v_0$ inside the shell the geometry is that of Minkowski space; the geometry outside the shell is Schwarzschild. Formation of the black hole occurs when the shell crosses the event horizon $\mathcal{H}^+$.

<u>Remarks:</u>

1. Notice that, as indicated in the figure, the event horizon $\mathcal{H}^+$ starts growing / expanding from $r = 0$ a long time before the shell arrives or crosses its Schwarzschild radius. This apparently acausal / prescient behaviour is a peculiar, but

very characteristic feature of the event horizon. This will be discussed further in section 32.

2. This example is also easily generalised to the description of the collapse of a shell onto a pre-existing Schwarzschild black hole with mass m_i by choosing the mass function to be

$$m(v) = m_i\Theta(v_0 - v) + m_f\Theta(v - v_0) = m_i + (m_f - m_i)\Theta(v - v_0) \ . \tag{29.7}$$

3. As an aside: in this case, the (apparently somewhat ham-handed and cavalier) procedure with distributional curvatures leading to (29.6) gives the correct (Barrabes-Israel[94]) surface energy density

$$\rho_\Sigma(r) = \frac{1}{4\pi G_N}\frac{m_f}{r^2} \tag{29.8}$$

of the collapsing null shell (null world volume Σ), with constant total mass $M = m_f/G_N$. This is due to the fact that we have worked from the outset in what are known as adapted coordinates, in this case in particular with the ingoing coordinate v. I will very briefly come back to this in section 29.2 below. In general, however, much more care is required to identify correctly the (distributional) components of the stress tensor of a thin shell.[95]

29.2 Some Comments on Null Shells and Adapted Coordinates

Here is a quick and rough (and by no means indispensable) explanation of what I meant by adapted coordinates in the last remark of the previous section.

To that end, let us describe a bit more carefully the situation we are trying to model. Thus let Σ be the null hypersurface describing the worldvolume of the shell. This divides the space-time into two parts, the inside (past) $\mathcal{V}_-$ and the outside (future) $\mathcal{V}_+$. Let us at first try to model this situation by using the standard Schwarzschild coordinates outside the shell (and correspondingly standard radial Minkowski coordinates inside the shell). This will of course not allow us to describe the geometry inside the horizon, so it is clear that this is not the optimal choice of coordinates, but this is not the issue here (which, as we will see, also manifests itself outside the horizon).

Thus on these two parts of space-time we have the metrics

$$\begin{aligned} \mathcal{V}_- : \quad & ds_-^2 = -dt_-^2 + dr^2 + r^2 d\Omega^2 \\ \mathcal{V}_+ : \quad & ds_+^2 = -f(r)dt_+^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \ , \end{aligned} \tag{29.9}$$

[94] C. Barrabes, W. Israel, *Thin shells in general relativity and cosmology: The lightlike limit*, Phys. Rev. D43 (1991) 1129-1143.

[95] A nice and characteristically lucid dicussion of the general formalism for null and non-null shells can be found in E. Poisson, *A Relativist's Toolkit: the Mathematics of Black Hole Mechanics*.

where $f(r) = 1-2m/r$ (I now write m instead of m_f). Here I have already identified the radial and angular coordinates across the shell (this is possible), but have been careful to introduce two different time-coordinates $t_\mp$. The reason for this is that these two time-coordinates cannot be identified.

Indeed, in terms of the internal Minkowski coordinates, the trajectory of the shell, i.e. the ingoing lightray, is described by

$$\Sigma_- : \quad t_- + r = 0 \quad , \tag{29.10}$$

say, whereas in terms of the external Schwarzschild coordinates it is described by

$$\Sigma_+ : \quad t_+ + r^* = 0 \quad , \tag{29.11}$$

or $t_+ + r^* = C$ for some constant C, with $r^* = r + 2m \log |r/2m - 1|$ the usual tortoise coordinate. Thus t_- and t_+ satisfy different equations, and can therefore not be identified across the shell). In this prototypical situation one can then not write a joint metric for the interior and exterior regions in terms of a single set of corrdinates (t, r, θ, ϕ), i.e. something like

$$ds^2 = -f(r)dt_+^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \tag{29.12}$$

with

$$f(r) \overset{???}{=} 1 - \frac{2m}{r}\Theta(r - r_{\text{shell}}(t)) \tag{29.13}$$

say, with $r_{\text{shell}}(t)$ (supposedly) describing the location of the shell, does not even make sense. In such a situation one has to appeal to the general Barrabes-Israel formalism (footnotes 94 and 95) to determine the surface energy-momentum tensor (and the correct "junction conditions").

Let us now look at the same problem in ingoing and outgoing coordinates. In ingoing coordinates

$$v_- = t_- + r \quad , \quad v_+ = t_+ + r^* \tag{29.14}$$

the metric on the two sides of the shell is

$$\begin{aligned} \mathcal{V}_- : \quad ds_-^2 &= -dv_-^2 + 2dv_- dr + r^2 d\Omega^2 \\ \mathcal{V}_+ : \quad ds_+^2 &= -f(r)dv_+^2 + 2dv_+ dr + r^2 d\Omega^2 \quad , \end{aligned} \tag{29.15}$$

with the location of the shell described by $v_+ = C_+$ outside the shell, and $v_- = C_-$ inside the shell, for some constants $C_\pm$. Just by shifting $v_\pm$ appropriately, we can arrange that $C_\pm = 0$, so that from both sides the shell is described by

$$\Sigma_\mp : \quad v_\mp = 0 \quad , \tag{29.16}$$

with $\mathcal{V}_\mp$ corresponding to $v_- < 0$ and $v_+ > 0$ respectively. Thus in this case

- one can identify $v_- = v_+$,

- it makes sense to write the metric collectively as in (29.3),

$$ds^2 = -f(v,r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m}{r}\Theta(v) \ , \tag{29.17}$$

- and one can calculate the surface energy-momentum tensor by determining the bulk Einstein tensor in these adapted coordinates (as done e.g. in the appendix of the Barrabes-Israel article cited in footnote 94).

As a final variation of this theme, let us look at what happens when one attempts to describe the ingoing shell in outgoing coordinates, i.e. in coordinates $(u_\pm, r, \theta, \phi)$ with

$$u_- = t_- - r \quad , \quad u_+ = t_+ - r^* \ . \tag{29.18}$$

Now the metric on the two sides of the shell takes the form

$$\begin{aligned} \mathcal{V}_- : \quad ds_-^2 &= -du_-^2 - 2du_- dr + r^2 d\Omega^2 \\ \mathcal{V}_+ : \quad ds_+^2 &= -f(r)du_+^2 - 2du_+ dr + r^2 d\Omega^2 \ , \end{aligned} \tag{29.19}$$

with the location of the shell described by

$$\Sigma_- : \quad u_- + 2r = C_- \qquad \Sigma_+ : \quad u_+ + 2r^* = C_+ \ . \tag{29.20}$$

Thus we are again in a situation where $u_\mp$ satisfy different equations and can hence not be identified across the shell (however, this would be the right choice of adapted coordinates for describing outgoing (exploding) shells). Therefore also in this case one would then need to appeal to the general Barrabes-Israel formalism to determine the surface energy-momentum tensor.

In the case at hand, even if for some reason one is interested in the final result in outgoing coordinates (e.g. if one wants to superimpose on this black hole geometry the effect of outgoing Hawking radiation), it is much simpler to first do the calculation in adapted (ingoing) coordinates and to then transform the result back to outgoing coordinates (and then this needs to be done separately for the interior and exterior regions). This option may not always be available, however (e.g. it is an analytically non-trivial problem to transform a general ingoing Vaidya metric to outgoing coordinates, say).

29.3 Qualitative Aspects of the Gravitational Collapse of a Star

To see how we could picture the situation of gravitational collapse (without having to address astrophysical questions and thus without trying to understand why this collapse occurs in the first place), let us estimate the average density ρ of a star whose radius r_0 is equal to its Schwarzschild radius. For a star with mass M we have

$$r_s = \frac{2MG_N}{c^2} \quad \text{and} \quad M \approx \frac{4\pi r_0^3}{3}\rho \ . \tag{29.21}$$

Therefore, setting $r_0 = r_s$, we find that

$$\rho = \frac{3c^6}{32\pi G_N^3 M^2} \approx 2 \times 10^{16} \ \mathsf{g/cm}^3 \ \left(\frac{M_{\text{sun}}}{M}\right)^2 \quad . \tag{29.22}$$

For stars of a few solar masses, this density is huge, roughly that of nuclear matter. In that case, there will be strong non-gravitational forces and hydrodynamic processes, significantly complicating the description of the situation. The situation is quite simple, however, when an object of the mass and size of a galaxy ($M \sim 10^{10} M_{\text{sun}}$) collapses. Then the critical density (29.22) is approximately that of air, $\rho \sim 10^{-3}$ $\mathsf{g/cm}^3$, non-gravitational forces can be neglected completely, and the collapse of the object can be approximated by a free fall. The Schwarzschild radius of such an object is of the order of light-days ($\sim 10^5\mathsf{s}$).

Under these circumstances, a more realistic Kruskal-like space-time diagram of a black hole would be the one depicted in Figure 32. We assume that at time $t = 0$ (equivalently, Kruskal time $T = 0$) we have a momentarily static mass configuration with (initial) radius $r_0 \gg 2m$ and mass M which then starts to collapse in free fall, the surface of the star described by a function $R_0 = R_0(t)$, say, or by $R_0 = R_0(\tau)$, where τ is the proper time of a freely falling particle on the surface of the star. In order to actually describe the crossing of the Schwarzschild radius by the surface of the star, it will evidently be more informative to use the parametrisation $R_0 = R_0(\tau)$.

Neglecting radiation-effects, the mass M of the star (galaxy) remains constant so that the exterior of the star, $r > R_0(\tau)$, is described by the corresponding subset of region I, and subsequently (once $R_0(\tau) < 2m$) also region II, of the Kruskal-Szekeres metric. Note that regions III and IV no longer exist because the region $r < R$ is simply not at all described by the Schwarzschild solution, but should be described by a solution of the Einstein equations appropriate for the interior of the collapsing star (in particular, this better be a solution of the non-vacuum Einstein equations, and we will describe such solutions later on in this section).

Schematic Kruskal and Penrose diagrams for this process are given in Figures 32 and 33. As the Penrose diagram shows, much like in the case of collapsing null shells the event horizon starts growing before the star has crossed its Schwarzschild radius. This is not our main concern here, but we will analyse this in some detail in section 32.12.

To model this, we start with the Schwarzschild metric in the region outside the star. For $t \leq 0$ this is the region $r > r_0$, while for $t > 0$ this is the region $r > R_0(t)$ where $R_0(t)$ is the radius of the star at time t, and $R_0(t)$ describes the radial free fall (geodesic motion) of the points on the surface towards the center discussed in section 26.3. By continuity of the metric, the space-time metric induced by the exterior metric on the

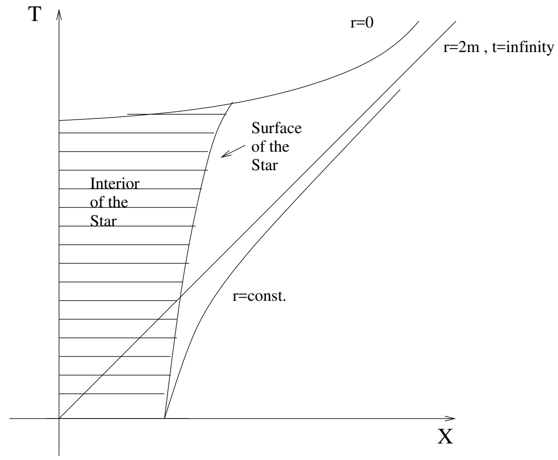


Figure 32: Kruskal diagram of a gravitational collapse. The surface of the star is represented by a timelike geodesic, modelling a star (or galaxy) in free fall under its own gravitational force. The surface will reach the singularity at $r = 0$ in finite proper time whereas an outside observer will never even see the star collapse beyond its Schwarzschild radius. However, as discussed in the text, even for an outside observer the resulting object is practically 'black'.

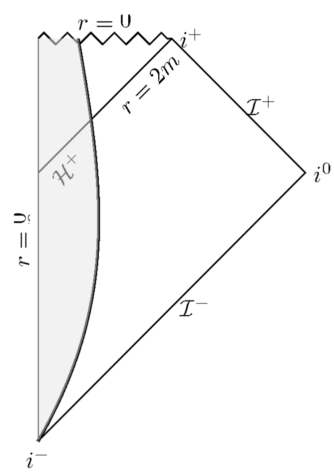


Figure 33: Penrose Diagram for the collapse of a star to a black hole (schematic). The shaded region indicates the interior of the star.

(2+1)-dimensional worldvolume Σ_{ext} of the surface of the star is then given by

$$
\begin{aligned}
ds^2_{\Sigma_{ext}} &= \left(-f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2\right)_{r=R_0(t)} \\
&= -[(f(R_0(t)) - f(R_0(t))^{-1}(dR_0/dt)^2]dt^2 + R_0(t)^2 d\Omega^2
\end{aligned}
\tag{29.23}
$$

(the subscript "ext" on Σ_{ext} is used to indicate that this is the metric induced on the surface of the star by the exterior metric, i.e. the metric outside the star). Expressed in terms of proper time τ, this becomes (writing now $R_0 = R_0(\tau)$, and $\dot{t} = dt/d\tau, \dot{R}_0 = dR_0/d\tau$, as usual)

$$ds^2_{\Sigma_{ext}} = -[(f(R_0(\tau))\dot{t}^2 - f(R_0(\tau))^{-1}\dot{R}_0^2]d\tau^2 + R_0(\tau)^2 d\Omega^2 \ . \tag{29.24}$$

Because $(t(\tau), R_0(\tau))$ parametrise radial timelike geodesics (for each value of the angular coordinates (θ, ϕ)), one has

$$-f(R_0(\tau))\dot{t}^2 + f(R_0(\tau))^{-1}\dot{R}_0^2 = -1 \ , \tag{29.25}$$

and therefore one finds

$$ds^2_{\Sigma_{ext}} = -d\tau^2 + R_0(\tau)^2 d\Omega^2 \ , \tag{29.26}$$

with $R_0(\tau)$ given implicitly by (26.28) and (26.29) (with $r_i \to r_0$),

$$\begin{aligned} R_0(\eta) &= \tfrac{1}{2} r_0 \ (1 + \cos\eta) \\ \tau(\eta) &= \left(\frac{r_0^3}{8m}\right)^{1/2} (\eta + \sin\eta) \ . \end{aligned} \tag{29.27}$$

<u>Remarks:</u>

1. This simple form of the metric is due to the fact that the radially falling particles remain at fixed values of the angular coordinates (so these are again *comoving coordinates*), and that τ is the corresponding proper time, so that one has $ds^2 = -d\tau^2$.

   Indeed, we can think of the induced metric $ds^2_{\Sigma_{ext}}$ as the restriction of the Novikov metric (27.62)

   $$ds^2 = -d\tau^2 + f(r_i)^{-1} r'(\tau, r_i)^2 dr_i^2 + r(\tau, r_i)^2 d\Omega^2 \ , \tag{29.28}$$

   to the comoving radial coordinate $r_i = r_0$,

   $$\begin{aligned} ds^2_{\Sigma_{ext}} &= -d\tau^2 + R_0(\tau)^2 d\Omega^2 \equiv -d\tau^2 + r(\tau, r_0)^2 d\Omega^2 \\ &= \left[-d\tau^2 + f(r_i)^{-1} r'(\tau, r_i)^2 dr_i^2 + r(\tau, r_i)^2 d\Omega^2\right]_{r_i = r_0} \ . \end{aligned} \tag{29.29}$$

2. From (29.27) one sees that R_0 takes its initial (maximal) value $R_0 = r_0$ at $\eta = 0$ or $\tau = 0$, will inevitably cross the Schwarzschild radius of the star at some finite value of τ, and will reach $r = 0$ at $\eta = \pi$ after the finite proper time

   $$\tau_{r_0 \to 0} = \left(\frac{r_0^3}{8m}\right)^{1/2} (\pi + \sin\pi) = \pi \left(\frac{r_0^3}{8m}\right)^{1/2} \ . \tag{29.30}$$

   For an object the size of the sun (for which our free-fall approximation is, however, not really adequate) this would be of the order of one hour, and correspondingly somewhat larger for larger, more massive and less dense, objects.

3. As an aside, note also that this implies that when freely falling radially into a black hole, the proper time it takes to reach the singularity at $r = 0$ once one has crossed the Schwarzschild radius is

$$\tau_{r_s \to 0} = \pi \left(\frac{r_s^3}{8m} \right)^{1/2} = \pi m \quad , \tag{29.31}$$

or, restoring c,

$$\tau_{r_s \to 0} = \pi G_N M / c^3 \quad . \tag{29.32}$$

4. For an observer remaining outside the collpasing star, say at the constant value $r = r_\infty$, in principle the situation (not unexpectedly by now) presents itself in a rather different way. Up to a constant factor $(1 - 2m/r_\infty)^{1/2}$, his proper time equals the coordinate time t. As the surface of the collapsing galaxy crosses the horizon at $t = \infty$, strictly speaking the outside observer will never see the black hole form.

   However, we had also seen that this period is accompanied by an infinite and exponentially growing gravitational redshift (26.49), $z \sim \exp t/4m$ for radially emitted photons. Therefore the luminosity L of the star decreases exponentially, as a consequence of this gravitational redshift and the fact that photons emitted at equal time intervals from the surface of the star reach the observer at greater and greater time intervals. It can be shown that

$$L \sim \mathrm{e}^{-t/3\sqrt{3}m} \quad , \tag{29.33}$$

   so that the star becomes very dark very quickly, the characteristic time being of the order of

$$3\sqrt{3}m \approx 2,5 \times 10^{-5}\mathsf{s} \; \left(\frac{M}{M_{\text{sun}}} \right) \quad . \tag{29.34}$$

   Thus, even though for an outside observer the collapsing star never disappears completely, for all practical intents and purposes the star is black and the name 'black hole' is justified.

5. Since only regions I and II of the Kruskal diagram are relevant for gravitational collapse, and for black holes arising from gravitational collapse, for most practical purposes Kruskal-Szekeres coordinates are not required and it is sufficient to consider coordinates that cover these two regions, such as Painlevé-Gullstrand (section 27.2) or Eddington-Finkelstein coordinates (section 27.4).

6. Note that, even if the free fall (geodesic) approximation is no longer justified at some point, once the surface of the star has crossed the Schwarzschild horizon, nothing, no amount of pressure, can stop the catastrophic collapse to $r = 0$ because, whatever happens, points on the surface of the star will have to move within their forward lightcone and will therefore inevitably end up at $r = 0$ in finite proper time.

In order to substantiate this claim, note that since timelike geodesics *maximise* proper time, any non-geodesic radial attempt to avoid hitting $r = 0$ will only get you there even quicker.[96]

Also, trying to somehow pick up some angular momentum will not help, because for $r < 2m$ the attractive general relativistic correction term in the effective potential (25.33) dominates over the repulsive angular momentum barrier term,

$$r < 2m \quad \Rightarrow \quad \frac{L^2}{2r^2} < \frac{mL^2}{r^3} \quad . \tag{29.35}$$

7. In interpreting the collapse to $r = 0$, it should be kept in mind that the Schwarzschild metric was never meant to be valid at $r = 0$ anyway (as it is supposed to describe the exterior of a gravitating body). Nevertheless, just being close enough to $r = 0$, without actually reaching that point is more than sufficient to crush any kind of matter. Indeed, (27.163) and the geodesic deviation equation (section 8.3) show that the force needed to keep neighbouring particles apart is proportional to r^{-3}. Thus the tidal forces within arbitrary objects (be they solids or elementary particles) eventually become infinitely big so that these objects will be crushed or torn apart completely. In that sense, the physics becomes hopelessly singular even before one reaches $r = 0$ and there seems to be nothing to prevent a collapse of such an object to $r = 0$ and infinite density.

8. In sections 29.4 - 29.8 below we will construct a matching interior solution to the Einstein equations describing a collapsing star, the Oppenheimer-Snyder collapse solution. It shows that this singularity is akin to a cosmological (Big Bang, or rather Big Crunch in the present context) singularity, and confirms that in the interior there is a genuine singularity in the form of a diverging matter density.

   Certainly classical general relativity (and even current-day quantum field theory) are inadequate to describe this situation (and if or how a theory of quantum gravity can deal with these matters remains to be seen).

It is fair to wonder at this point if the above conclusions regarding the collapse to $r = 0$ are only a consequence of the fact that we assumed exact spherical symmetry. Would the singularity be avoided under more general conditions? The answer to this is, somewhat surprisingly and shockingly, a clear 'no'.

It has e.g. been shown that the gravitational field of a static vacuum black hole, even without further symmetry assumptions, is necessarily given by the spherically symmetric Schwarzschild metric and is thus characterised by the single parameter M (Israel, 1967).

[96] This conclusion is a bit hasty. For a slightly more careful discussion of this issue, see e.g. A. Toporensky, O. Zaslavskii, *On strategies of motion under the black hole horizon*, `arXiv:1905.02150 [gr-qc]`, and the non-technical account A. Toporensky, S. Popov, *How to delay death and look further into the future if you fall into a black hole*, `arXiv:2209.04823 [physics.pop-ph]`.

This was the first of a series of remarkable black hole uniqueness (or "no hair") theorems which I will briefly come back to in section 30.1 below. Curiously, initially the result by Israel was interpreted by many as confirming that such singularities could only occur in exactly spherically symmetric situations.[97] It turned out, however, that what this theorem actually implies is that higher multipole moments will have to be radiated away during gravitational collapse.

Moreover, there are very general *singularity theorems*, due to Penrose, Hawking and others, which all state in one way or another that if Einstein's equations hold, the energy-momentum tensor satisfies some kind of positivity condition, and there is a regular event horizon, then some kind of singularity will appear (typically in some form of "geodesic incompleteness", i.e. in the existence of geodesics that cannot be extended to arbitrary values of their affine parameter). These theorems do not rely on any symmetry assumptions.[98]

In this sense, therefore, singularities appear to be unavoidable in classical general relativity, and the theory predicts and points to its own incompleteness ("it's singular" can hardly be considered to be a satisfactory answer . . .).

29.4 Oppenheimer-Snyder Set-Up: Geometry and Matter Content

In section 24.7 we had described the general set-up (as well as a special solution) for the interior solution of a static spherically symmetric star, and in section 29.3 above we have described the exterior (Schwarzschild) geometry of a collapsing star. We will now attempt to find an idealised description of the interior of a star during the time-dependent phase of gravitational collapse. This interior of a star will be modelled on a (bounded subset of a) gravitationally collapsing cosmological model, in particular that of a collapsing "dust"-filled universe. The exact solutions to the Einstein (Friedmann-Lemaître) equations for this case are derived in section 37.3.

We will also make sure that the exterior and interior descriptions of this gravitational collapse match at the surface of the star.

If we assume that matter inside the spherically symmetric star can be modelled by a perfect fluid with spatially uniform energy density $\rho = \rho(t)$ and pressure $p = p(t)$, then the spatial geometry is locally both isotropic and homogeneous. Thus (see section 14.1) the spatial geometry of the star is that of a (bounded subspace of a) maximally

[97] I owe this remark to A. Ashtekar, *The Last 50 Years of General Relativity and Gravitation: From GR3 to GR20 Warsaw Conference*, `arXiv:1312.6425 [gr-qc]`.

[98] See e.g. S. Hawking and G. Ellis, *The large scale structure of space-time* and R. Wald, Chapter 9 of *General Relativity* for textbook accounts, chapter 1 of S. Hawking, *The Nature of Space and Time*, `arXiv:hep-th/9409195` for an introduction, and J. Senovilla, *Singularity Theorems in General Relativity: Achievements and Open Questions*, `arXiv:physics/0605007v1 [physics.hist-ph]` for an overview.

symmetric space, and solutions are then governed by the Friedmann equations (section 35.7), i.e. by the Einstein equations specialised to this situation. Some familiarity with sections 34.4, 35.7 and 37.3 will therefore be necessary (and assumed) in the following.

Let us first address the geometry of this problem. The spherically symmetric star is a 3-ball B^3, i.e. a 3-dimensional space with boundary a 2-sphere S^2. Its 2-dimensional counterpart would usually be called a disc (or 2-disc) D^2, a surface with boundary a circle S^1. A priori, one could model the geometry of this disc e.g. as the subset of the Euclidean plane (with its induced maximally symmetric flat metric),

$$ds^2(D^2) = dr^2 + r^2 d\phi^2 \qquad (r \leq r_0) \tag{29.36}$$

or as the cap of a sphere (with its induced maximally symmetric positive curvature metric),

$$ds^2(D^2) = d\theta^2 + \sin^2\theta \, d\phi^2 \qquad (\theta \leq \theta_0) \tag{29.37}$$

or even as its negative curvature counterpart, say the Poincaré disc model of the hyperbolic plane, given in polar coordinates in (11.67), i.e.

$$ds^2(D^2) = 4\frac{dr^2 + r^2 d\phi^2}{(1-r^2)^2} \qquad (\, 0 \leq r \leq r_0 \,) \ . \tag{29.38}$$

Likewise, one can model the 3-ball geometry of a spherically symmetric star in terms of bounded subspaces of any of the $k = 0, \pm 1$ 3-geometries that we have been considering, e.g. the spatially flat 3-ball or 3-disc for $k = 0$ or the cap of a 3-sphere for $k = +1$,

$$\begin{aligned} k = 0: \quad & ds^2(B^3) = dr^2 + r^2 \, d\Omega_2^2 && (r \leq r_0) \\ k = +1: \quad & ds^2(B^3) = d\psi^2 + \sin^2\psi \, d\Omega_2^2 && (\psi \leq \psi_0) \end{aligned} \tag{29.39}$$

(we had already encountered the $k = +1$ cap/disc/3-ball as the spatial geometry underlying a static spherically symmetric star in section 24.7).

As far as the matter content is concerned, a spatially constant non-zero pressure would in particular lead to a non-zero pressure at the surface of the star. This would need to be compensated by a non-zero surface tension, a further contribution to the energy-momentum tensor, δ-function localised on the surface of the star. In order not to have to deal with this situation, we will only consider the simplest possibility, namely that of pressureless dust, $p = 0$ (Oppenheimer and Snyder, 1939).

In this case, the interior solution is provided by the cosmological solutions of the matter dominated era derived in section 37.3, with the radial coordinate of the star restricted to run over a finite range, as in (29.39). The exterior solution would then, as in our discussion of section 29.3, be given by the Schwarzschild metric, and one thing we need to do is make sure that the exterior and interior descriptions of the surface of the star agree.

In order to understand the role of the spatial curvature k and how to glue the interior and exterior solutions together, recall that we had already seen in section 35.5, in (35.78), that pressureless dust necessarily moves along geodesics of the space-time geometry. In the present case these are the geodesics of comoving observers (dust particles) and the cosmological time t is their proper time. In particular, if we think of $a(t)$ as (proportional to) the time-dependent radius of the collapsing star, then $a(t)$ describes the geodesic trajectories of particles at the surface of the star.

However, these surface particles should follow geodesics as if the total mass of the star were concentrated at the center of the star, i.e. they should also move along geodesics of the outside Schwarzschild geometry with that mass. Thus we are led to the, a priori perhaps somewhat surprising, statement that the Friedmann equations for dust must agree with the geodesic equation for radially freely falling particles in the Schwarzschild geometry. This is indeed the case:

- On the one hand, the evolution of the cosmic scale factor is governed by the Friedmann equation (37.23),
$$\dot{a}^2 + k = \frac{C_m}{a} \quad . \tag{29.40}$$
- On the other hand, according to the results of section 26.3, radial free fall is governed by the equation (26.25),
$$\dot{r}^2 + \frac{2m}{r_i} = \frac{2m}{r} \quad , \tag{29.41}$$
where r_i is the radius where the particle is initially at rest, $\dot{r}(r = r_i) = 0$.

We see that these really do have the same form (and we will match them more precisely below). We also see that the choice of pressureless matter, i.e. the equation of state parameter $w = 0$, for the interior of the star is essential for this (other values of w leading to other powers of a on the right-hand side of (29.40)).

This similarity of the equations in the case $w = 0$ is also reflected in the explicit solutions of the geodesic and Friedmann equations, as for example in the solution (29.27) of the radial geodesic equation in the Schwarzschild geometry,

$$\begin{aligned} R_0(\eta) &= \tfrac{1}{2} r_0 \,(1 + \cos\eta) \\ \tau(\eta) &= \left(\frac{r_0^3}{8m}\right)^{1/2} (\eta + \sin\eta) \quad , \end{aligned} \tag{29.42}$$

and the recollapsing solution (37.34) of the Friedmann equations for a spatially closed dust-filled universe,

$$\begin{aligned} a(\eta) &= \frac{a_{max}}{2}(1 - \cos\eta) \\ t(\eta) &= \frac{a_{max}}{2}(\eta - \sin\eta) \quad . \end{aligned} \tag{29.43}$$

From this we see that a finite r_i or r_0 corresponds to a $k = +1$ interior solution. The spatially flat $k = 0$ 3-disc geometry, on the other hand, corresponds to $r_i = \infty$, i.e. the case where the surface of the star behaves as if it had been released from rest at infinity, as can be seen by comparing the explicit solutions for radial geodesics (26.24) and the cosmic scale factor (37.25) in this case,

$$r(\tau) \sim (\tau_0 - \tau)^{2/3} \quad , \quad a(t) \sim (t_f - t)^{2/3} \quad . \tag{29.44}$$

Thus a matching with the exterior geometry of the collapsing star discussed in section 29.3 (where we assumed free fall from rest from a finite radius r_0) requires $k = +1$. However, the $k = 0$ solution is also instructive in its own right, and we will analyse both possibilities below.

In order to match the exterior and interior geometries, we should start by matching the coordinates used for the two solutions. In both cases, due to spherical symmetry (and due to having chosen coordinates that make this symmetry manifest), there is a transverse 2-sphere with line-element $d\Omega_2^2 = d\theta^2 + \sin^2\theta d\phi^2$, and we will simply identify the coordinates (θ, ϕ) of the two solutions.

This leaves us with the temporal and radial directions. The time coordinate of the cosmological (interior) metrics is the proper time of comoving observers (in particular those on the surface of the star), and this is a natural choice which we will maintain. It follows that also for the exterior Schwarzschild geometry we should choose coordinates such that the time-coordinate is the proper time of these comoving = freeely falling observers, and we have already constructed various such coordinate systems in section 27. In particular, we could use

- either Painlevé-Gullstrand coordinates (section 27.2), adapted to $r_i = \infty$, i.e. $k = 0$ (or their $r_i < \infty$ counterparts which we did not discuss explicitly - see the reference in footnote 82), and their cosmological counterpart (which we had already briefly introduced in section 34.4);
- or comoving Lemaître ($k = 0$) or Novikov ($k = +1$) coordinates (section 27.3, the latter already featured in our discussion of the exterior solution for gravitational collapse in section 29.3) and their cosmological counterpart (which are just the standard comoving coordinates of the Robertson-Walker metrics).

In order to illustrate the procedure, we will pursue both options and discuss the case $k = 0$ in terms of PG-like coordinates and the case $k = +1$ in terms of comoving coordinates.

29.5 $k = 0$ Collapse and Painlevé-Gullstrand Coordinates

We model the exterior geometry by the Schwarzschild metric in PG coordinates (section 27.2)

$$ds^2 = -dT^2 + (dr + \sqrt{2m/r}dT)^2 + r^2 d\Omega^2 \quad . \tag{29.45}$$

This is adapted to radially infalling observers with $dr = -\sqrt{2m/r}dT$ (so that $dT = d\tau$ is proper time). These are the radial geodesics with $E = 1 \leftrightarrow r_i = \infty$. We assume that from the exterior point of view particles on the surface of the star follow such geodesics, so that the surface of the star is described by $r = R_0(\tau)$, with

$$\dot{R}_0(\tau) = -\sqrt{2m}(R_0(\tau))^{-1/2} \quad \Rightarrow \quad R_0(\tau) = (9m/2)^{1/3}(\tau_0 - \tau)^{2/3} \ . \tag{29.46}$$

We model the interior geometry by the spatially flat $k = 0$ Robertson-Walker metric in PG-like coordinates (34.43),

$$ds^2 = -d\tau^2 + (d\tilde{r} - \tilde{r}H(\tau)d\tau)^2 + \tilde{r}^2 d\Omega^2 \ . \tag{29.47}$$

Here we have already identified the cosmological time $t = \tau$ as the proper time of comoving observers, $H(\tau) = \dot{a}(\tau)/a(\tau)$ is the Hubble parameter, and the radial coordinate $\tilde{r}$ is related to the usual comoving radial coordinate of the Robertson-Walker metric (now denoted r_c, to avoid confusion with the radial coordinate of the Schwarzschild or PG metric)

$$ds^2 = -d\tau^2 + a(\tau)^2(dr_c^2 + r_c^2 d\Omega^2) \tag{29.48}$$

by

$$\tilde{r} = a(\tau)r_c \ . \tag{29.49}$$

This form of the metric is adapted to comoving observers (fixed r_c), which obey the Hubble relation $d\tilde{r}/d\tau = \tilde{r}H(\tau)$. We assume that the surface of the star has fixed (comoving) radial coordinate $r_{c,0}$, and is thus described by $\tilde{r} = \tilde{R}_0(\tau)$,

$$\tilde{R}_0(\tau) = a(\tau)r_{c,0} \quad \Rightarrow \quad \dot{\tilde{R}}_0 = \dot{a}r_{c,0} = H\tilde{R}_0 \ . \tag{29.50}$$

Its time-dependence (i.e. the collapse of the star) is governed by the negative square-root of the $k = 0$ Friedmann equation for pressureless matter,

$$\begin{aligned} \dot{a}(\tau) = -\sqrt{C_m}a(\tau)^{-1/2} \quad &\Rightarrow \quad \dot{\tilde{R}}_0(\tau) = -\sqrt{C_m}(r_{c,0})^{3/2}\tilde{R}_0(\tau)^{-1/2} \\ &\Rightarrow \quad \tilde{R}_0(\tau) = r_{c,0}(9C_m/4)^{1/3}(\tau_0 - \tau)^{2/3} \ , \end{aligned} \tag{29.51}$$

with

$$H(\tau) = -2/3(\tau_0 - \tau) \ . \tag{29.52}$$

Thus more explicitly the interior metric can now be written as

$$\begin{aligned} ds^2 &= -d\tau^2 + (d\tilde{r} + 2\tilde{r}d\tau/3(\tau_0 - \tau))^2 + \tilde{r}^2 d\Omega^2 \\ &= -d\tau^2 + (d\tilde{r} + \sqrt{2\tilde{m}(\tau, \tilde{r})/\tilde{r}}d\tau)^2 + \tilde{r}^2 d\Omega^2 \ , \end{aligned} \tag{29.53}$$

where

$$\tilde{m}(\tau, \tilde{r}) = \frac{2\tilde{r}^3}{9(\tau_0 - \tau)^2} \ . \tag{29.54}$$

Comparison of the two metrics, the exterior Schwarzschild metric in Painlevé-Gullstrand form (29.45) and the interior metric in Painlevé-Gullstrand-like form (29.53), makes it manifest that we should identify not only PG time T with the cosmological time $t = \tau$ (as already anticipated above), but also the PG-Schwarzschild radial coordinate r with the cosmological radial coordinate $\tilde{r} = a(\tau) r_c$.

A seamless matching of the two metrics then requires the identification of the location of the surface of the star from the two sides, through

$$R_0(\tau) = \tilde{R}_0(\tau) \ , \tag{29.55}$$

or, equivalently, that the mass function $\tilde{m}(\tau, \tilde{r})$ (29.54) evaluated on the surface of the star $\tilde{r} = \tilde{R}_0(\tau)$, agrees with the constant Schwarzschild mass m of (29.45),

$$\tilde{m}(\tau, \tilde{R}_0(\tau)) = m \quad \Leftrightarrow \quad \tilde{R}_0(\tau) = (9m/2)^{1/3}(\tau_0 - \tau)^{2/3} = R_0(\tau) \ . \tag{29.56}$$

Note that this necessarily leads to the requirement (that we had already imposed) that the interior of the star is described by pressureless dust, in order to reproduce the characteristic $\tau^{2/3}$-behaviour of the geodesic.

Either from the explicit expression for the two solutions, or from comparing the geodesic equation in (29.46) with the Friedmann equation in (29.51), one finds that $R_0(\tau) = \tilde{R}_0(\tau)$ is equivalent to

$$R_0(\tau) = \tilde{R}_0(\tau) \quad \Leftrightarrow \quad (9m/2)^{1/3} = r_{c,0}(9C_m/4)^{1/3} \quad \Leftrightarrow \quad 2m = C_m r_{c,0}^3 \ . \tag{29.57}$$

This resulting condition relating the parameters of the exterior and interior solutions can be demystified by recalling the definition (36.14) of C_m as the constant

$$C_m = \frac{8\pi G_N}{3}\rho(\tau)a(\tau)^3 \ . \tag{29.58}$$

Then (29.57) becomes

$$2m = C_m r_{c,0}^3 \quad \Leftrightarrow \quad M \equiv \frac{m}{G_N} = \frac{4\pi}{3}(a(\tau)r_{c,0})^3\rho(\tau) = \frac{4\pi}{3}\tilde{R}_0(\tau)^3 \ \rho(\tau) \ , \tag{29.59}$$

which is simply the statement that at all times the Schwarzschild gravitational mass-energy M of the star felt by the freely-falling particles on the star's surface is precisely the total mass-energy (density times volume) of the star. This encapsulates the essence of the Oppenheimer-Snyder construction.

With this identification, the induced metric on the surface Σ of the star satisfies

$$\begin{aligned}
ds^2_{\Sigma_{ext}} &\equiv \left[-dT^2 + (dr + \sqrt{2m/r}dT)^2 + r^2 d\Omega^2\right]_{T=\tau, r=R_0(\tau)} \\
&= -d\tau^2 + R_0(\tau)^2 d\Omega^2 \\
&= -d\tau^2 + \tilde{R}_0(\tau)^2 d\Omega^2 \\
&= \left[-d\tau^2 + (d\tilde{r} - \tilde{r}H(\tau)d\tau)^2 + \tilde{r}^2 d\Omega^2\right]_{\tilde{r}=\tilde{R}_0(\tau)} \equiv ds^2_{\Sigma_{int}} \ .
\end{aligned} \tag{29.60}$$

29.6 Synopsis of the Oppenheimer-Snyder Construction

As the above construction involved a number of different ingredients, from Schwarzschild geodesics to collapsing solutions of the Friedmann equations, one runs the risk of not seeing the forest for the trees. Thus it may be useful to provide a brief summary / synopsis of what we have done so far.

Dropping all tildes and other now (in retrospect) irrelevant decorations, and choosing without loss of generality the instant of total collapse of the star to be at $\tau_0 = 0$, the set-up and results can be summarised as follows:

1. The construction turns out to be particularly simple in Painlevé-Gullstrand-like coordinates.[99] In particular,

   - the exterior Schwarzschild metric is described in terms of Painlevé-Gullstrand coordinates,

   $$ds^2 = -d\tau^2 + (dr + \sqrt{2m/r}d\tau)^2 + r^2 d\Omega^2 \ ; \tag{29.61}$$

   - the interior (cosmological) metric is described in terms of Painlevé-Gullstrand-like cosmological coordinates

   $$ds^2 = -d\tau^2 + (dr - rH(\tau)d\tau)^2 + r^2 d\Omega^2 \ . \tag{29.62}$$

   where $H(\tau)$ is the Hubble parameter for a dust-filled contracting universe.

2. The surface of the star is described by an equation of the form

   $$r = R(\tau) = C(-\tau)^{2/3} \quad \Rightarrow \quad H(\tau) = -\tfrac{2}{3}(-\tau)^{-1} < 0 \ . \tag{29.63}$$

   Here C is given in terms of the total mass m (the parameter characterising the exterior Schwarzschild geometry) by

   $$C = (9m/2)^{1/3} \ , \tag{29.64}$$

   and $R(\tau)$ in (29.63) describes equivalently

   - either radial infalling geodesics in the exterior mass m Schwarzschild geometry with $E = 0$ (i.e. which would have started off at infinity with zero initial velocity),
   - or a collapsing dust-sphere solution of the Friedmann equations.

[99] More recently this was also recognised in Y. Kanai, M. Siino, A. Hosoya, *Gravitational collapse in Painlevé-Gullstrand coordinates*, arXiv:1008.0470 [gr-qc].

3. Jointly the exterior and interior metrics can be written compactly as

$$ds^2 = -d\tau^2 + \left(dr + \sqrt{\frac{2m(\tau, r)}{r}} d\tau\right)^2 + r^2 d\Omega^2 \tag{29.65}$$

with

$$m(\tau, r) = \begin{cases} m & r > R(\tau) = C(-\tau)^{2/3} \\ 2r^3/9(-\tau)^2 & r < R(\tau) = C(-\tau)^{2/3} \end{cases} \tag{29.66}$$

This solution describes a collapsing dust star for $\tau < \tau_0 = 0$, collapsing to zero radius at time $\tau = \tau_0 = 0$.

Alternatively, it is occasionally convenient to write the metric more explicitly in terms of the radial freefall velocities

$$v(\tau, r) = \sqrt{\frac{2m(\tau, r)}{r}} \tag{29.67}$$

as

$$ds^2_\pm = -d\tau^2 + (dr + v_\pm(\tau, r) d\tau)^2 + r^2 d\Omega^2 \tag{29.68}$$

where in the exterior (+) and interior (−) regions one has

$$v_+(\tau, r) = v_+(r) = \frac{2m}{r} \quad , \quad v_-(\tau, r) = -rH(\tau) = \frac{2r}{3(-\tau)} \quad . \tag{29.69}$$

In particular, continuity of the metric across the surface of the star is now expressed by the fact that on the surface of the star one has $v_+ = v_-$, with

$$v_+(R(\tau)) = v_-(\tau, R(\tau)) = \left(\frac{4m}{3(-\tau)}\right)^{1/3} \quad . \tag{29.70}$$

29.7 Continuity of the Normal Derivatives of the Metric

While the metric is now certainly continuous across the surface of the star, in order to complete the story one should also check that the first derivatives of the metric match on the two sides as well. Indeed, by the Einstein equations in order for the energy-momentum tensor to only exhibit a finite jump as one crosses Σ, rather than a δ-function localised surface energy-momentum tensor on Σ, also the 1st derivative of the metric induced on Σ should be continuous. This is automatic for derivatives tangent to the surface, and thus this continuity requirement boils down to the requirement that the normal derivatives of the metric (i.e. derivatives in the direction orthogonal to Σ) agree on Σ, and we will come back to this issue below.

Second derivatives, however, will not and cannot be continuous across the surface, because the energy momentum tensor has spatially constant density inside the star and is identically zero outside the star, so that by the Einstein equations also the Einstein tensor necessarily has a discontinuity across the surface of the star.

In order to address the issue of continuity of the normal derivatives of the induced metric, first of all we need to determine the (normalised) normal vectors on the two sides. To that end we note first that the tangent directions to the surface of the star are the two spacelike directions tangent to the 2-sphere S^2, as well as the timelike direction u^α spanned by the geodesics describing the free-fall motion of the surface. Thus the normal vector N^α is a spacelike radial vector orthogonal to u^α, determined (up to a choice of sign) by the conditions

$$u^\alpha N_\alpha = 0 \quad , \quad N^\alpha N_\alpha = +1 \ . \tag{29.71}$$

This is particularly simple in Painlevé-Gullstrand coordinates (on both sides), in which one has

$$u_\alpha = -\partial_\alpha \tau = (-1,0,0,0) \quad \Rightarrow \quad u^\alpha = (1,-v_\pm,0,0) = (\dot{\tau},\dot{r},\dot{\theta},\dot{\phi}) \ . \tag{29.72}$$

Thus N^α, with $N^\alpha u_\alpha = 0$ and radial, is necessaarily proportional to $(0,1,0,0)$, i.e. $N^r \neq 0$, $N^\alpha = 0$ otherwise, and since $g_{rr} = 1$ the correctly normalised choice is

$$N^\alpha = (0,1,0,0) \quad \Leftrightarrow \quad N = N^\alpha \partial_\alpha = \partial_r \ , \tag{29.73}$$

which is evidently continuous across the surface Σ (and we have chosen $N = \partial_r$ to be outwards pointing).

Returning now to the issue at hand, namely the continuity of the normal derivative of the metric, recall first of all that we had already checked the continuity of the metric on Σ, a condition that we can express in terms of the induced metric

$$h_{\alpha\beta} = g_{\alpha\beta} - N_\alpha N_\beta \tag{29.74}$$

or its equivalent $h_{ab} = E^\alpha_a E^\beta_b h_{\alpha\beta}$ as the statement that

$$h^+_{\alpha\beta} = h^-_{\alpha\beta} \quad \Leftrightarrow \quad h^+_{ab} = h^-_{ab} \ , \tag{29.75}$$

where $h^\pm_{ab}$ denotes the metric on Σ induced from the exterior / interior geometry respectively. Now recall that in section 18.2 we introduced the extrinsic curvature $K_{\alpha\beta}$ precisely as the tangential projection of the normal derivative of the induced metric (18.24),

$$\begin{aligned} K_{\alpha\beta} &= \tfrac{1}{2} h^\gamma_\alpha h^\delta_\beta L_N g_{\gamma\delta} = h^\gamma_\alpha h^\delta_\beta \nabla_\gamma N_\delta \\ \Leftrightarrow \quad K_{ab} &= E^\alpha_a E^\beta_b \nabla_\alpha N_\beta \ . \end{aligned} \tag{29.76}$$

In view of this it is entirely plausible that we can formulate the condition for the continuity of the normal derivative of the metric across the surface of the star, and the absence of distributionally localised energy-momentum at the surface of the star, as the condition that the interior and exterior extrinsic curvatures be equal,

$$K^+_{\alpha\beta} = K^-_{\alpha\beta} \quad \Leftrightarrow \quad K^+_{ab} = K^-_{ab} \ . \tag{29.77}$$

This is indeed the correct condition and together the conditions (29.75) and (29.77) are known as the *Israel(-Darmois) junction condition.*[100]

We will now verify (29.77). Since N^α has the simple form ∂_r (and the same form both outside and inside the star), while for its associated covector one has the (marginally more complicated) expression

$$N_\alpha = (v_\pm, 1, 0, 0) \tag{29.78}$$

(of course with $v_+ = v_-$ on Σ), it is convenient to write (29.76) in terms of the covariant derivative of N^α as

$$K_{\alpha\beta} = h_\alpha^\gamma h_{\beta\delta} \nabla_\gamma N^\delta|_\Sigma = h_\alpha^\gamma h_{\beta\delta}\, \Gamma^\delta_{\gamma\mu} N^\mu|_\Sigma = h_\alpha^\gamma h_\beta^\delta\, \Gamma_{\delta\gamma r}|_\Sigma \ \ . \tag{29.79}$$

By construction, $N^\alpha K_{\alpha\beta} = K_{\alpha\beta}N^\beta = 0$, and therefore we only need to analyse the tangential (angular and along u^α) components of $K_{\alpha\beta}$:

1. Let $y^k = (\theta, \phi)$ denote the angular coordinates. Then one has

   $$h_k^\gamma = \delta_k^\gamma \tag{29.80}$$

   because N_α has no angular components. Thus for the angular components one simply needs

   $$g_{ik} dy^i dy^k = r^2 d\Omega^2 \quad \Rightarrow \quad \Gamma_{ikr} = \tfrac{1}{2}\partial_r g_{ik} = r(d\theta_i d\theta_k + \sin^2\theta d\phi_i d\phi_k) \tag{29.81}$$

   and therefore

   $$K^+_{ik} = K^-_{ik} = R(\tau)(d\theta_i d\theta_k + \sin^2\theta d\phi_i d\phi_k) \ \ . \tag{29.82}$$

2. For the mixed components, say $K_{i\beta}u^\beta$ one needs

   $$u^\beta \Gamma_{i\beta r} = \tfrac{1}{2}(u^\beta \partial_r g_{i\beta} + u^\beta \partial_\beta g_{ir} - u^\beta \partial_i g_{\beta r}) \ \ . \tag{29.83}$$

   and all three terms are individually zero: (i) $(\partial_r g_{i\beta})u^\beta = 0$ because u^β has no component in an angular direction, (ii) $g_{ir} = 0$, (iii) $g_{\beta r}$ components are independent of y^i. Therefore one has

   $$K^+_{i\beta}u^\beta = K^-_{i\beta}u^\beta = 0 \ \ . \tag{29.84}$$

3. It remains to show that the $u - u$ components $K^\pm_{\alpha\beta}u^\alpha u^\beta$ are equal. This can be shown by explicit calculation, but it is more enlightning to note that this follows in general from (18.43), because u^α is geodesic. Thus

   $$K^+_{\alpha\beta}u^\alpha u^\beta = K^-_{\alpha\beta}u^\alpha u^\beta = 0 \ \ . \tag{29.85}$$

[100] For a detailed discussion of these junction conditions see E. Poisson, *A Relativist's Toolkit: the Mathematics of Black Hole Mechanics*, sections 3.7 - 3.11.

Putting (1), (2) and (3) together, we have thus established

$$K^+_{\alpha\beta} = K^-_{\alpha\beta} \tag{29.86}$$

and therefore the continuity of the first derivatives of the metric across the surface of the star. We also learn that this is essentially due to the fact that we have matched the two space-times along geodesics (something that would not have been necessary if we had just wanted to have a continuous metric).

This is also easy to understand intuitively: one could have of course tried to force our dust star to collapse at a different rate, i.e. not in free fall (or e.g. tried to force a star with a different interior to collapse as if it were in free fall). In either case, however, this would require introducing some pressure / surface tension localised at the surface of the star to make the star do this. Within the Israel junction condition formalism such a surface energy-momentum tensor is precisely equivalent to a discontinuity of the extrinsic curvature.

29.8 $k = 1$ Collapse and Comoving Coordinates

We now consider the collapse of a star whose surface is initially at rest at some finite radius. We model the exterior geometry by the Schwarzschild metric in comoving Novikov coordinates (27.62),

$$ds^2 = -d\tau^2 + f(r_i)^{-1} r'(\tau, r_i)^2 dr_i^2 + r(\tau, r_i)^2 d\Omega^2 \ , \tag{29.87}$$

r_i is a comoving coordinate, and we assume that particles on the the surface of the star move along the geodesics with $r_i = r_0$, r_0 labelling the maximal radius of the star. As already described in section 29.3, the restriction of the Novikov metric (27.62) to the surface of the star, i.e. to the comoving radial coordinate $r_i = r_0$, is

$$\begin{aligned} ds^2_{\Sigma_{ext}} &= \left[-d\tau^2 + f(r_i)^{-1} r'(\tau, r_i)^2 dr_i^2 + r(\tau, r_i)^2 d\Omega^2\right]_{r_i = r_0} \\ &= -d\tau^2 + R_0(\tau)^2 d\Omega^2 \end{aligned} \tag{29.88}$$

where $R_0(\tau) = r(\tau, r_0)$ solves the geodesic equation (29.41) (with $r_i \to r_0$),

$$\dot{R}_0^2 + \frac{2m}{r_0} = \frac{2m}{R_0} \ . \tag{29.89}$$

As in (29.27), we write the solution in parametrised form as

$$\begin{aligned} R_0(\eta) &= \tfrac{1}{2} r_0 \, (1 + \cos\eta) \\ \tau(\eta) &= \left(\frac{r_0^3}{8m}\right)^{1/2} (\eta + \sin\eta) \ , \end{aligned} \tag{29.90}$$

with the collapse beginning at $\tau = 0 \leftrightarrow \eta = 0$, when $R_0(0) = r_0$.

We model the interior geometry by the spatially closed $k = +1$ Robertson-Walker metric in the standard comoving coordinates (where we now, in analogy with the notation of the previous section, write r_c for the comoving "radial" coordinate)

$$ds^2 = -d\tau^2 + a(\tau)^2 (\frac{dr_c^2}{1-r_c^2} + r_c^2 \, d\Omega^2) \quad . \tag{29.91}$$

We assume that from the interior point of view the surface of the star is at the fixed comoving radius $r_c = r_{c,0}$, so that the induced metric on the surface of the star is

$$\begin{aligned} ds^2_{\Sigma_{int}} &= \left[-d\tau^2 + a(\tau)^2 (\frac{dr_c^2}{1-r_c^2} + r_c^2 \, d\Omega^2) \right]_{r_c = r_{c,0}} \\ &= -d\tau^2 + a(\tau)^2 r_{c,0}^2 \, d\Omega^2 \equiv -d\tau^2 + \tilde{R}(\tau)^2 \, d\Omega^2 \quad . \end{aligned} \tag{29.92}$$

Now $a(\tau)$ satisfies the $k = +1$ Friedmann equation

$$\dot{a}^2 + 1 = \frac{C_m}{a} \quad , \tag{29.93}$$

whose solution in parametrised form is (37.34). Shifting η (by π) and $t = \tau$ (by $a_{max}\pi/2$) so that the maximal radius $a_{max} = C_m$ is reached at $\eta = 0, \tau = 0$, the solution takes the form

$$\begin{aligned} a(\eta) &= \frac{C_m}{2}(1 + \cos\eta) \\ \tau(\eta) &= \frac{C_m}{2}(\eta + \sin\eta) \quad . \end{aligned} \tag{29.94}$$

Thus $\tilde{R}_0(\tau) = r_{c,0} a(\tau)$ satisfies the equation

$$\dot{\tilde{R}}(\tau)^2 + r_{c,0}^2 = \frac{r_{c,0}^3 C_m}{\tilde{R}_0} \quad . \tag{29.95}$$

Continuity of the metric across the surface of the star requires $R_0(\tau) = \tilde{R}_0(\tau)$, and comparison of (29.89) and (29.95), say, gives us two conditions. The first,

$$\frac{2m}{r_0} = r_{c,0}^2 \quad \Leftrightarrow \quad r_{c,0} = \sqrt{2m/r_0} < 1 \tag{29.96}$$

just provides us with the relation between the comoving Novikov coordinate r_0 and the comoving Robertson-Walker coordinate $r_{c,0}$. The second,

$$2m = C_m r_{c,0}^3 \quad , \tag{29.97}$$

is identical to the condition (29.57) found in the previous section, in the context of the $k = 0$ PG collapse, with the same consequence

$$M \equiv \frac{m}{G_N} = \frac{4\pi}{3} \rho(\tau) \tilde{R}_0(\tau)^3 \quad . \tag{29.98}$$

The physical content of this equation is again that the (constant) Schwarzschild mass M, i.e. the gravitational mass of the collapsing star as seen from the outside, is at any time τ given by the product of the density and the (coordinate) volume of the star (cf. also the comment on coordinate versus proper volume in this context in section 24.6).

The two conditions (29.96) and (29.97) can also equivalently be written as

$$C_m = (r_0^3/2m)^{1/2} \quad \text{and} \quad r_0 = C_m r_{c,0} \ , \tag{29.99}$$

and with these identification it is now manifest that the solution for $R_0(\tau)$ given in (29.90) is identical to the solution for $\tilde{R}_0(\tau) = r_{c,0}a(\tau)$ obtained from (29.94). Thus the metric is now manifestly continuous across the surface of the star (with analogous comments regarding its 1st and 2nd derivatives as in the case $k = 0$).

30 Black Holes IV: Other Black Hole Solutions (a brief overview)

So far, we have only discussed the Schwarzschild geometry as an example of a black hole solution of the Einstein equations, but this is far from the only solution. While the situation is surprisingly manageable (and classifiable) in the situation that has traditionally attracted the most interest of the general relativity community (namely that of *stationary*, *asymptotically flat* solutions of the *4*-dimensional *Einstein-Maxwell equations*), the situation changes completely if any of the above italicised conditions is or are relaxed.

As a consequence, it is hopeless to attempt to give a reasonably complete overview of other black hole solutions, let alone to discuss them in some detail. Nevertheless, it is good to have a rough idea of what other sorts of black objects may exist. In this section I will therefore give a very brief (and by no means representative) overview of some selected other black hole solutions, with references to the literature for further information.

30.1 Kerr-Newman Family of 4-dimensional Black Holes

While the Schwarzschild solution that we have discussed at length above is not the only black hole solution of the Einstein equations, in 4 space-time dimensions the possibilities are remarkably restricted. In particular, (with some technical assumptions) it can be shown that the most general stationary and asymptotically flat black hole solution of the 4-dimensional vacuum Einstein or Einstein-Maxwell equations (with a regular event horizon) is characterised by just three parameters, namely its mass M, charge Q and angular momentum J.

These *black hole uniqueness theorems* constitute a significant generalisation of the remarkable Israel theorem (1967) on the Schwarzschild solution (briefly already referred to at the end of section 29.3), which states that (under certain technical conditions) the unique regular static black hole solution of the Einstein vacuum equations is the Schwarzschild solution. In particular, under these circumstances staticity implies spherical symmetry (this is not to be confused with the content of Birkhoff's theorem which asserts that spherical symmetry and the vacuum Einstein equations imply staticity, a much more elementary result).

The generalisations of this theorem constituting the black hole uniqueness theorems are colloquially referred to as the fact that *black holes have no hair* or also as the *no-hair theorems*.[101] In terms of gravitational collapse, these theorems roughly amount

[101] For an historical review and account of these developments, see e.g. D.C. Robinson, *Four decades of black hole uniqueness theorems*, available from `http://www.mth.kcl.ac.uk/staff/dc_robinson/papers.html`. For a detailed recent critical review of the status (and limitations) of the assertions of the uniqueness theorems, see P. Chrusciel, J.

to the statement that the only characteristics of a black hole which are not somehow radiated away during the phase of collapse via multipole moments of the gravitational, electro-magnetic, ... fields are those which are protected by some conservation laws.

The two most important examples of black hole solutions generalising the Schwarzschild metric, 2-parameter subfamilies of the complete 3-parameter family of black hole metrics, are the Reissner-Nordstrøm and Kerr metrics.

- Reissner - Nordstrøm Metric

  This is the metric with line element

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2} \ . \tag{30.1}$$

  It is a solution of the coupled Einstein-Maxwell equations describing the exterior geometry of a spherically symmetric electrically charged star. It is characterised by two parameters m and q related to its mass M and charge Q respectively and arises from the Schwarzschild metric by the replacement

$$m \to m - q^2/2r \ , \tag{30.2}$$

  which can be thought of as a "mass renormalisation" due to the electrostatic self-energy.

  Although (astro-)physically perhaps not particularly relevant, it displays a number of interesting and curious features which are of interest in their own right and relatively easy to understand. For this reason we will take a rather detailed look at this solution in section 31.

- Kerr and Kerr-Newman Metrics

  Astrophysical black holes, while they carry negligible charge Q, are expected to typically have a non-zero angular momentum J. Now one no longer has spherical symmetry (because the axis of rotation picks out a particular direction) but only axial symmetry. The situation is thus a priori much more complicated, and the solution was found by Kerr only in 1963, almost fifty years after the Schwarzschild and Reissner-Nordstrøm solutions, through monumental calculations.[102]

  The Kerr metric describing a rotating black hole depends on two parameters m and a related to its mass M and angular momentum J respectively. In *Boyer-Lindquist coordinates* (one of the coordinate systems in which the metric looks

Lopes Costa, M. Heusler, *Stationary Black Holes: Uniqueness and Beyond*, Living Rev. Relativity 15 (2012) 7, `http://www.livingreviews.org/lrr-2012-7`, `arXiv:1205.6112 [gr-qc]`.

[102]See e.g. R. Kerr, *Discovering the Kerr and Kerr-Schild metrics*, `arXiv:0706.1109 [gr-qc]`; M. Visser, *The Kerr spacetime: A brief introduction* `arXiv:0706.0622 [gr-qc]`, S. Teukolsky, *The Kerr Metric*, `arXiv:1410.2130 [gr-qc]`.

"simplest", relatively speaking), the metric takes the form

$$\begin{aligned}
ds^2 &= -dt^2 + \frac{2mr(dt - a\sin^2\theta d\phi)^2}{\rho(r,\theta)^2} + (r^2+a^2)\sin^2\theta d\phi^2 \\
&\quad + \rho(r,\theta)^2(\Delta(r)^{-1}dr^2 + d\theta^2) \\
&= -\left(1 - \frac{2mr}{\rho(r,\theta)^2}\right)dt^2 - \frac{4mra\sin^2\theta}{\rho(r,\theta)^2}dt\, d\phi + \frac{\Sigma(\rho,\theta)}{\rho(r,\theta)^2}\sin^2\theta d\phi^2 \\
&\quad + \rho(r,\theta)^2(\Delta(r)^{-1}dr^2 + d\theta^2) \quad ,
\end{aligned} \tag{30.3}$$

where

$$\begin{aligned}
\Delta(r) &= r^2 - 2mr + a^2 \\
\rho(r,\theta)^2 &= r^2 + a^2\cos^2\theta \\
\Sigma(r,\theta) &= (r^2+a^2)^2 - \Delta(r)a^2\sin^2\theta \quad .
\end{aligned} \tag{30.4}$$

Another useful way of grouping the terms in the metric is as (suppressing now the arguments (r,θ) of the functions Δ, ρ, Σ)

$$ds^2 = -\frac{\rho^2\Delta}{\Sigma}dt^2 + \frac{\Sigma}{\rho^2}\sin^2\theta(d\phi - \omega dt)^2 + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 \quad , \tag{30.5}$$

where

$$\omega = 2mar/\Sigma \tag{30.6}$$

evidently has the interpretation as some kind of angular velocity.

<u>Remarks:</u>

1. This metric is time-independent and axially symmetric, with the two commuting Killing vectors ∂_t and ∂_ϕ.
2. However, it is only stationary, not static (cf. the discussion in section 16.4). In the present adapted coordinates, in which the metric components are independent of t, this amounts to the statement that the metric is invariant under constant time-translations but not under time-reflection $t \to -t$ (because of the rotation term $g_{t\phi}$). More invariantly, this is the statement that the Killing vector ∂_t is not hypersurface-orthogonal, neither to the surfaces of constant t not to any other hypersurface.
3. The Kerr metric fairly manifestly reduces to the Schwarzschild metric for $a = 0$. It also reduces to the Minkowski metric for $m = 0$, as one might expect ("rotating Minkowski space is still Minkowski space"), but this is somewhat less manifest as one obtains the Minkowski metric in some rather obscure coordinates (known as "oblate spheroidal" coordinates).
4. Regardless of how one writes the metric, its singularity and horizon structure and the behaviour of geodesics are much more intricate and intriguing than for the Schwarzschild and Reissner-Nordstrøm solutions. We will take a brief

look at the horizon structure in section 32.3, but for more details (and an analysis of the more intricate singularity structure) I have to refer you to any of the modern standard textbooks on general relativity.

5. Electric charge can be added to this solution by the same replacement $m \to m - q^2/2r$ (30.2) as in the relation between the Schwarzschild and Reissner-Nordstrøm metrics. The resulting (charged Kerr or rotating Reissner-Nordstrøm) metric is known as the *Kerr-Newman metric.*[103]

30.2 Other 4-dimensional Solutions

There exist various generalisations of the Schwarzschild, Reissner-Nordstrøm and Kerr metrics in 4 space-time dimensions (higher-dimensional generalisations will be discussed below), ranging from the reasonably straightforward to the surprising and to the outright weird:

1. Kottler Metric

   The *Kottler metric*

   $$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} - \frac{\Lambda r^2}{3} \tag{30.7}$$

   is the unique spherically symmetric solution of the Einstein vacuum equations with a cosmological constant Λ,

   $$G_{\mu\nu} + \Lambda g_{\mu\nu} = 0 \quad \Leftrightarrow \quad R_{\mu\nu} = \Lambda g_{\mu\nu} \quad . \tag{30.8}$$

   It is also known as the *Schwarzschild - de Sitter metric* for $\Lambda > 0$ and the *Schwarzschild - anti-de Sitter metric* for $\Lambda < 0$. This solution is not asymptotically flat but asymptotically (A)dS, i.e. asymptotic to pure de Sitter or anti-de Sitter space, which is the maximally symmetric solution of the Einstein equation with a positive (negative) cosmological constant - see section 39 for a detailed discussion of these space-times. In particular, we will derive the solution (30.7) in section 39.2 (see equation (39.64)).

   Unsurprisingly, one can also add charge to this solution,

   $$f(r) = 1 - \frac{2m}{r} - \frac{\Lambda r^2}{3} + \frac{q^2}{r^2} \quad , \tag{30.9}$$

   to find an exact charged black hole solution of the Einstein-Maxwell equations with a cosmological constant.

[103] See e.g. T. Adamo, E. Newman, *The Kerr-Newman metric: A Review*, `arXiv:1410.6626 [gr-qc]` for the (pre-)history of this metric (and the Kerr metric) and a discussion of its basic properties.

2. Topological Black Holes

   Remarkably, and more surprisingly, for $\Lambda < 0$ one can replace the "1" in $f(r)$ by a constant $k = 0, \pm 1$,

$$f_k(r) = k - \frac{2m}{r} - \frac{\Lambda r^2}{3} \tag{30.10}$$

   (formally this also works for $\Lambda > 0$, but since $f_k(r)$ is strictly negative in that case, this requires some reinterpretation ...), provided that one also replaces the 2-sphere by $\mathbb{R}^2$ or T^2 for $k = 0$, and the 2-dimensional hyperboloid H^2 for $k = -1$,

$$ds^2 = -f_k(r)dt^2 + f_k(r)^{-1}dr^2 + r^2 d\Omega^2_{(k)} \quad , \quad d\Omega^2_{(k)} = \begin{cases} d\Omega_2^2 & \text{for } k = +1 \\ d\vec{x}^2 & \text{for } k = 0 \\ d\tilde{\Omega}_2^2 & \text{for } k = -1 \end{cases} \tag{30.11}$$

   ($d\tilde{\Omega}_2^2$ denotes the line element of the standard metric on H^2). These solutions describe black holes immersed into AdS space, with horizons with a non-spherical topology. Therefore such solutions are also, somewhat confusingly, known as *topological black holes*.[104] A special case of (30.11) are the metrics (39.142) one obtains for $m = 0$, which describe pure AdS space (no black hole) in different coordinate systems.

3. Black Hole Solutions of the Einstein-Yang-Mills(-Higgs) Equations

   It is natural to consider not just solutions of the Einstein-Maxwell equations, but more generally solutions of the Einstein equations coupled to other fields that appear in the fundamental theories of physics, such as Yang-Mills fields and scalar fields. The study of solutions to these equations brought about some surprises, and the realisation that solutions to the Einstein-Maxwell equations have special properties that are not valid for other matter content. In particular, there appears to be no useful analogue of the "no-hair theorem" for these solutions (e.g. there are static solutions of the Einstein-Yang-Mills equations with non-trivial Yang-Mills fields but vanishing Yang-Mills global charges). Moreover, there are black hole solutions of these equations that are static but not spherically symmetric, so that there is no analogue of the Israel theorem for this matter content either. All this makes the study of solutions to these equations quite interesting and rewarding (but is also an analytically challenging activity).[105]

4. Regular Black Holes

   A curious class of solutions are so-called *regular black hole solutions*, solutions with event horizons but without singularities. Due to the singularity theorems

[104] See R. Mann, *Topological Black Holes – Outside Looking In*, `arXiv:gr-qc/9709039` for a general discussion of these 4-dimensional topological black holes.

[105] Early developments are reviewed in M. Volkov, D. Galtsov, *Gravitating Non-Abelian Solitons and Black Holes with Yang-Mills Fields*, Physics Reports 319 (1999) 1-83, `arXiv:hep-th/9810070`. For a recent assessement and discussion, see section 5 of Chrusciel et al. (footnote 101).

of general relativity, mentioned at the end of section 29.3, which are typically of the form "under some reasonable assumptions, if there is something like an event horizon, there must be something like a singularity", such solutions need to walk a fine line between avoiding the singularity theorems and not being outright unphysical. Usually this is achieved by some weak violation of the (occasionally unreasonably strong) positive energy conditions (cf. section 22.1) entering the singularity theorems, in particular the strong energy condition (SEC).[106]

One of the earliest and simplest solutions of this kind, which is also of the standard simple $f - f^{-1}$-form, is the *Bardeen solution*, a solution of the Einstein equations coupled to (some non-linear version of) Maxwell theory, with metric function

$$f(r) = 1 - \frac{2mr^2}{(r^2+e^2)^{3/2}} \quad . \tag{30.12}$$

For suitable choice of the mass and charge parameters m and e, $f(r)$ possesses simple zeros (the largest zero corresponding to an event horizon). It approaches the Schwarzschild metric for large r,

$$r \to \infty : \quad f(r) \to 1 - \frac{2m}{r} \tag{30.13}$$

but for small r it approaches the de Sitter metric (in the form of the metric (30.7) with $m = 0$),

$$r \to 0 : \quad f(r) \to 1 - (2m/e^3)r^2 \equiv 1 - \Lambda r^2/3 \quad . \tag{30.14}$$

It is thus completely regular at $r = 0$ (the "cosmological constant" near the core providing the required violation of the positive energy conditions, specifically the SEC).

Another popular regular metric is the so-called Hayward metric.[107] In this metric, the singularity at $r = 0$ is regularised by a cut-off length parameter L (which one is invited to think of as something like the Planck length), the function $f(r)$ having the form

$$f(r) = 1 - \frac{2mr^3}{r^4 + 2mrL^2} \quad . \tag{30.15}$$

Again this has the asymptotic Schwarzschild behaviour for $r \to \infty$, and for $r \to 0$ one has

$$r \to 0 : \quad f(r) \to 1 - (r/L)^2 \quad , \tag{30.16}$$

corresponding to a de Sitter metric with curvature radius L.

Many other solutions of this kind and their general properties are known and understood.[108]

[106] See e.g. O. Zaslavskii, *Regular black holes and energy conditions*, `arXiv:1004.2362 [gr-qc]` for a discussion.

[107] S. Hayward, *Formation and evaporation of non-singular black holes*, `arXiv:gr-qc/0506126`.

[108] If you are tempted to open this particular box in Pandora's large collection of boxes better left untouched, I refer you to the recent review article by S. Ansoldi, *Spherical black holes with regular center*, `arXiv:0802.0330 [gr-qc]`, and to the copious references to the original literature therein.

5. Vaidya Metrics

Moving away from time-independent solutions, there is a simple class of time-dependent generalisations of the Schwarzschild metric known as Vaidya metrics. They can be obtained from the Schwarzschild metric written in ingoing or outgoing Eddington-Finkelstein coordinates by replacing the constant mass m by a mass function $m(v)$ or $m(u)$ depending on an advanced or retarded time coordinate. Thus these metrics have the form

$$\begin{aligned} ds^2 &= -f(v,r)dv^2 + 2dvdr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m(v)}{r} \\ ds^2 &= -f(u,r)du^2 - 2dudr + r^2 d\Omega^2 \quad , \quad f(u,r) = 1 - \frac{2m(u)}{r} \quad . \end{aligned} \tag{30.17}$$

These turn out, none too surprisingly, to give rise to solutions to the Einstein equations that describe null dust (or radiation) either entering (falling into) the black hole or star (for $m = m(v)$ a function of the ingoing Eddington-Finkelstein coordinate) or exiting from (or being radiated away by) the black hole or star (for $m = m(u)$ a function of the outgoing Eddington-Finkelstein coordinate). In particular, Vaidya metrics provide one with toy models that allow one to study the formation and evolution of a black hole.

There are a lot of interesting things that one can do with, say about and learn from Vaidya metrics. I will discuss some of them in sections 32 and 40 - 42.

30.3 Higher-dimensional Solutions

The (standard) black hole solutions are also easily generalised to higher dimensions, but in addition to that higher dimensions surprisingly offer many more possibilities that have no 4-dimensional counterpart:

5. Schwarzschild-Tangherlini Solution

The $D = d+1$ dimensional generalisation of the Schwarzschild metric is sometimes also known as the *Schwarzschild-Tangherlini solution*. It has the standard form

$$\begin{aligned} ds^2 &= -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2_{d-1} \\ f(r) &= 1 - \frac{\mu}{r^{d-2}} \equiv 1 - \left(\frac{r_0}{r}\right)^{d-2} \quad . \end{aligned} \tag{30.18}$$

Here μ is again proportional to the mass M of the black hole (cf. also section 24.8), and this solution has an event horizon at $r = r_0$ (with spatial topology S^{d-1}).

There is also a corresponding generalisation of the Reissner-Nordstrøm solution, and these static and asymptotically flat solutions are subject to uniqueness theorems that are pretty much analogous to those in $D = 4$.

6. Topological Black Holes in Higher Dimensions

   There is a corresponding, but much richer, generalisation of the topological black hole solutions,

$$f_k(r) = k - \frac{\mu}{r^{d-2}} \pm \frac{r^2}{\ell^2} \tag{30.19}$$

   where the transverse space $(d-1)$-dimensional space can now be S^{d-1}, $\mathbb{R}^{d-1}$ or H^{d-1}, or any other Einstein manifold with a metric h_{ij} with the same curvature, $R_{ij}(h) = (d-2)kh_{ij}$.[109]

7. Myers-Perry Black Holes

   There are also analogues of the Kerr metric for $D > 4$, known as *Myers-Perry black holes*, characterised by n rotation parameters for $D = 2n+1$ or $D = 2n+2$ (with n being the rank of the spatial rotation group $SO(D-1) = SO(2n)$ or $SO(D-1) = SO(2n+1)$ and thus the maximal number of independent commuting generators of the corresponding Lie algebra).[110]

8. Black Strings and Branes (extended gravitating objects)

   As soon as one moves beyond $D = 4$, it is easy to write down solutions of the (vacuum or matter-) Einstein equations that describe extended black objects. For example, by simply taking the direct product of the usual 4-dimensional Schwarzschild solution and the real line (or a circle) with coordinate y,

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_2^2 + dy^2 \quad , \quad f(r) = 1 - \frac{2m}{r} \quad , \tag{30.20}$$

   one obtains a solution of the 5-dimensional vacuum Einstein equations describing what is known as a *black string* (a black object extended in the y-direction). More generally, and in the same way, the $(d+1)$-dimensional Tangherlini solution (30.18) can be extended to

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_{d-1}^2 + d\vec{y}_p^2 \quad , \quad f(r) = 1 - \frac{\mu}{r^{d-2}} \quad , \tag{30.21}$$

   a solution of the $(d+p+1)$-dimensional vacuum Einstein equations describing what is known as a a *black p-brane* (a black object extended in p spatial directions $\vec{y}$). These are just the first and simplest examples of a bewildering assortment of p-brane solutions of higher-dimensional supergravity theories that play an important role in supergravity and string theory.[111]

[109] See e.g. D. Birmingham, *Topological Black Holes in Anti-de Sitter Space*, `arXiv:hep-th/9808032`, and R. Emparan, *AdS/CFT Duals of Topological Black Holes and the Entropy of Zero-Energy States*, `arXiv:hep-th/9906040` for details and applications of these solutions.

[110] For more, and authoritative, information about these solutions see R. Myers, *Myers-Perry black holes*, `arXiv:1111.1903 [gr-qc]`.

[111] For a gentle first introduction to these topics see H. Elvang, G. Horowitz, *Quantum gravity via supersymmetry and holography*, `arXiv:1311.2489 [gr-qc]`, and for a detailed discussion of p-brane solutions, see T. Ortin, *Gravity and Strings*, in particular sections 18 and 19.

9. Black Rings and other Exotic Black Objects

What is perhaps as remarkable as the uniqueness theorems for rotating black holes in $D = 4$ is the fact that the situation is completely different for $D > 4$, a far cry from the completely orderly and manageable situation in $D = 4$. In particular, in $D = 5$ there are asymptotically flat *black ring* solutions with horizon-topology $S^2 \times S^1$, and even more exotic objects in $D > 5$. The general construction and classification of black solutions in higher dimensions is an active area of research and many open questions remain.[112]

[112] See e.g. R. Emparan, H. Reall, *Black Holes in Higher Dimensions*, `arXiv:0801.3471 [hep-th]` for a general review, as well as S. Hollands, A. Ishibashi, *Black hole uniqueness theorems in higher dimensional spacetimes*, `arXiv:1206.1164 [gr-qc]`, and G. Galloway, *Constraints on the topology of higher dimensional black holes*, `arXiv:1111.5356 [gr-qc]`.

31 Black Holes V: the Reissner-Nordstrøm Solution

The Reissner - Nordstrøm metric is a solution of the coupled Einstein-Maxwell equations describing the exterior geometry of a spherically symmetric electrically charged star or black hole. It has the form (30.1)

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \quad f(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2} \ . \tag{31.1}$$

In this section we will analyse various aspects of this metric. Some of these depend on the specific form of $f(r)$, e.g. the analysis of the motion of (charged) particles in section 31.6. Others, such as the construction of Eddington-Finkelstein and Kruskal-Szekeres coordinates in sections 31.8 and 31.10, are valid more generally for static black holes of the ubiquitous (see e.g. the examples in section 30) $-f(r)dt^2 + f(r)^{-1}dr^2$ form, and these will initially be discussed in this more general context before specialising to the Reissner-Nordstrøm metric.

31.1 Derivation of the Reissner-Nordstrøm Metric

In order to obtain the solution we again start with with the standard form (24.6)

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{31.2}$$

of a spherically symmetric metric, and for the gauge field make the electrostatic ansatz that

$$A_t = A_t(r) \equiv -\phi(r) \quad , \quad A_r = A_\theta = A_\phi = 0 \ . \tag{31.3}$$

with $\phi(r)$ the usual scalar potential. Thus the only non-vanishing component of the field strength tensor is

$$F_{tr} = \partial_t A_r - \partial_r A_t = \phi'(r) = -E_r \ . \tag{31.4}$$

Therefore one has

$$F_{\alpha\beta}F^{\alpha\beta} = -2A(r)^{-1}B(r)^{-1}\phi'(r)^2 \ , \tag{31.5}$$

This implies that the energy-momentum tensor

$$T_{\alpha\beta} = \frac{1}{4\pi}\left(F_{\alpha\gamma}F_\beta^{\ \gamma} - \tfrac{1}{4}g_{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta}\right) \tag{31.6}$$

(here and in the following it is convenient to use Gauss units) has the non-vanishing components

$$(T_{tt}, T_{rr}, T_{\theta\theta}, T_{\phi\phi}) = \left(A(r), -B(r), r^2, r^2\sin^2\theta\right) A(r)^{-1}B(r)^{-1}\frac{\phi'(r)^2}{8\pi} \ , \tag{31.7}$$

or

$$T^\alpha_{\ \beta} = \text{diag}(-1,-1,+1,+1)\,\frac{\phi'(r)^2}{8\pi A(r)B(r)} \ . \tag{31.8}$$

Note that this is traceless, as it should be. Moreover, the combination

$$B(r)T_{tt} + A(r)T_{rr} = 0 \tag{31.9}$$

vanishes identically so that as in the Schwarzschild case (24.26) - (24.28) one concludes that

$$BR_{tt} + AR_{rr} = 0 \quad \Rightarrow \quad A(r) = B(r)^{-1} \equiv f(r) \ . \tag{31.10}$$

This means that the determinant of the metric is the same as the determinant of the flat metric. Recalling that

$$\nabla_\alpha F^{\alpha\beta} = \frac{1}{\sqrt{g}}\partial_\alpha(\sqrt{g}F^{\alpha\beta}) \tag{31.11}$$

this in turn implies that the usual Minkowski space solution of the Maxwell equations

$$\partial^\alpha F_{\alpha\beta} = -4\pi j_\beta \quad , \quad j_\alpha = (Q\delta(r), 0, 0, 0) \tag{31.12}$$

describing the electric field of an electric point charge with charge Q at $r = 0$, namely

$$\phi(r) = Q/r \quad , \quad E_r = Q/r^2 \ , \tag{31.13}$$

is also a solution of the Maxwell equations in the gravitational background we are trying to determine. Thus we can take the matter source of the Einstein equations to be given by the standard electrostatic field $E_r = Q/r^2$ of a point charge. The energy momentum tensor then reduces to

$$(T_{tt}, T_{rr}, T_{\theta\theta}, T_{\phi\phi}) = \left(f(r), -f(r)^{-1}, r^2, r^2\sin^2\theta\right)\frac{Q^2}{8\pi r^4} \tag{31.14}$$

which can also be written as

$$T^\alpha_\beta = \text{diag}(-1, -1, +1, +1)\ \frac{Q^2}{8\pi r^4} \ . \tag{31.15}$$

Note the characteristic negative radial pressure! Note also that it is a fortuitous coincidence (or hindsight, if you will) in this case that, by doing things in the right order, we have been able to more or less decouple the matter and gravitational equations. Usually one cannot just plug one's favourite Minkowski solution to the equations of motion into the energy-momentum tensor and then solve the Einstein equations because there is no guarantee that the initial solution will also be a solution of the matter equations of motion in the resulting curved space-time.

As for Schwarzschild, the fact that $A(r) = B(r)^{-1} = f(r)$ implies (cf. (24.25) and (24.29)) that

$$R_{\theta\theta} = 1 - f - rf' \tag{31.16}$$

and plugging this into the $(\theta\theta)$-component of the Einstein equation one finds

$$1 - f - rf' = 8\pi G_N T_{\theta\theta} = G_N\frac{Q^2}{r^2} \quad \Rightarrow \quad f = 1 + \frac{C}{r} + G_N\frac{Q^2}{r^2} \ . \tag{31.17}$$

One can now verify that this is a solution of the complete set of Einstein(-Maxwell) equations.

Comparison with the Schwarzschild solution, and introducing, in analogy with the gravitational mass radius $m = G_N M$, the gravitational charge radius q via

$$q^2 = G_N Q^2 \tag{31.18}$$

then gives

$$f(r) = \left(1 - \frac{2m}{r} + \frac{q^2}{r^2}\right) \tag{31.19}$$

and finally the Reissner-Nordstrøm solution

$$ds^2 = -\left(1 - \frac{2m}{r} + \frac{q^2}{r^2}\right) dt^2 + \left(1 - \frac{2m}{r} + \frac{q^2}{r^2}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{31.20}$$

The main new features of the Reissner-Nordstrøm metric, compared with the Schwarzschild metric, are all due to the fact that the function $f(r)$ now potentially has two roots, at

$$r_\pm = m \pm \sqrt{m^2 - q^2} \ . \tag{31.21}$$

Thus one has to distinguish the three cases $m^2 - q^2 < 0, = 0, > 0$, corresponding to the three possibilities for the relative size of the gravitational mass radius $m = G_N M$ and the gravitational charge radius $q = \sqrt{G_N} Q$, and we will discuss these three possibilities in turn below. In the following, we will always assume that $m > 0$.

<u>Remarks:</u>

1. Instead of with (31.2) one can also start with the ansatz (24.68)

$$ds^2 = -\mathrm{e}^{2h(t,r)} f(t,r) dt^2 + f(t,r)^{-1} dr^2 + r^2 d\Omega^2 \ , \tag{31.22}$$

for which the Einstein equations take the form (24.71). In particular, one concludes

$$\dot{m}(t,r) = 4\pi G_N \, r^2(-T^r_{\ t}) = 0 \tag{31.23}$$

and

$$h'(t,r) = 4\pi G_N \, r f(t,r)^{-1}(-T^t_{\ t} + T^r_{\ r}) = 0 \ , \tag{31.24}$$

so that, without loss of generality, we can choose $h(t,r) = 0$. The remaining equation for $m(r)$ is

$$m'(r) = 4\pi G_N \, r^2(-T^t_{\ t}) = \frac{G_N Q^2}{2r^2} \ , \tag{31.25}$$

leading to

$$m(r) = m - \frac{G_N Q^2}{2r} = m - \frac{q^2}{2r} \tag{31.26}$$

and thus to (31.19),

$$f(r) = 1 - \frac{2m(r)}{r} = 1 - \frac{2m}{r} + \frac{q^2}{r^2} \ . \tag{31.27}$$

2. As a check on the dimensions note that Newton's constant has dimensions (M mass, L length, T time) $[G_N] = \mathsf{M}^{-1}\mathsf{L}^3\mathsf{T}^{-2}$ so that (24.38)

$$[G_N] = \mathsf{M}^{-1}\mathsf{L}^3\mathsf{T}^{-2} \quad \Rightarrow \quad [m] = [G_N M/c^2] = \mathsf{L} \ , \tag{31.28}$$

while in Gauss units the Coulomb force has no dimensionful factors apart from Q_1Q_2/r^2, and therefore (force $F = ma$ having units $[F] = \mathsf{MLT}^{-2}$)

$$[Q^2] = \mathsf{L}^2(\mathsf{MLT}^{-2}) = \mathsf{ML}^3\mathsf{T}^{-2} \quad \Rightarrow \quad [q^2] = [G_N Q^2/c^4] = \mathsf{L}^2 \ . \tag{31.29}$$

31.2 Basic Properties of the Naked Singularity Solution with $m^2 - q^2 < 0$

In this case of an "overcharged" star (this is not a very realistic situation to put it mildly), $f(r)$ has no real roots, and the coordinate system is valid all the way to $r = 0$ (where there is a curvature singularity). In particular, the coordinate t is always time-like and the coordinate r is always space-like. While this may sound quite pleasing, much less insane than what happens for the Schwarzschild metric, this is actually a disaster.

The singularity at $r = 0$ is now timelike, and it is not protected by an event-horizon. Such a singularity is known as a *naked singularity*. An observer could travel to the singularity and come back again. Worse, whatever happens at the singularity can influence the future physics away from the singularity, but as there is a singularity this means that the future cannot be predicted/calculated in such a space-time because the laws of physics break down at $r = 0$. The Penrose diagram of this space-time looks exactly like that of the negative mass Schwarzschild solution (Figure 28 in section 28.5).

Note that $m^2 - q^2 < 0$ includes as a special case the solution with $m = 0$, supposedly describing the gravitational field of a massless charged object. As shown in section 23.4, m measures the total energy / mass of the system (including, therefore, the positive electrostatic energy of the solution). There is thus clearly something disturbingly unphysical about this solution.

There is a famous conjecture, known as the *Cosmic Censorship Conjecture* (due to Penrose, 1969), which roughly speaking states that the collapse of physically realistic matter configurations will generically not lead to a naked singularity. In spite of a lot of partial results and circumstantial evidence in favour of this conjecture, it is not known if (or in which precise form) it holds in General Relativity.[113]

Since, beyond exhibiting a naked timelike singularity the causal structure of these space-times is not particularly interesting, and they are considered to be unphysical (and, ideally, excluded by cosmic censorship), we shall not discuss this case any further in the following and just note the following curious fact: if one wanted to model elementary

[113] See e.g. R. Wald, *Gravitational Collapse and Cosmic Censorship*, `arXiv:gr-qc/9710068`.

particles such as the electron classically as point particles with a given mass and charge (but one should, in any case, resist that temptation), they would satisfy $q^2 > m^2$ by a wide margin (since the gravitational interaction is completely negligible compared with the Coulomb repulsion).

However, one should not conclude from this that elementary particles should hence be modelled by naked singularities - electrons are essentially quantum mechanical objects and outside the regime of validity / applicability of classical general relativity, and thus outside the regime of validity of the present considerations.

31.3 Basic Properties of the Extremal Solution with $m^2 - q^2 = 0$

In this (astrophysically speaking still unrealistic but theoretically particularly interesting) case, known as the extreme or extremal Reissner-Nordstrøm solution, and characterised by

$$G_N^2 M^2 = G_N Q^2 \quad \Leftrightarrow \quad Q = \pm\sqrt{G_N} M \quad \Leftrightarrow \quad q = \pm m \ , \tag{31.30}$$

the function

$$f(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2} = \left(1 - \frac{m}{r}\right)^2 \tag{31.31}$$

has a double zero at $r_+ = r_- = m$. Thus the metric takes the simple form

$$ds^2 = -\left(1 - \frac{m}{r}\right)^2 dt^2 + \left(1 - \frac{m}{r}\right)^{-2} dr^2 + r^2 d\Omega^2 \ . \tag{31.32}$$

Since $f(r) \geq 0$ everywhere, the singularity is timelike as in the overcharged case discussed above. Thus crossing the horizon one can avoid running into the singularity and turn back. However, since one cannot cross the same horizon in two directions (one should and can substantiate this by constructing ingoing Eddington-Finkelstein coordinates in this case which will exhibit this one-way behaviour via the tilting of the lightcones at the horizon, and we will do this below) this means that on the way out one is really crossing a white hole horizon (use outgoing Eddington-Finkelstein coordinates there) into a new asymptotically flat extremal Reissner-Nordstrøm universe, and this story repeats itself ad infinitum.

It is instructive to write the metric (31.32) in a slightly different form, suggested by

$$\begin{aligned} ds^2 &= -\left(1 - \frac{m}{r}\right)^2 dt^2 + \left(1 - \frac{m}{r}\right)^{-2} \left(dr^2 + \left(1 - \frac{m}{r}\right)^2 r^2 d\Omega^2\right) \\ &= -\left(1 - \frac{m}{r}\right)^2 dt^2 + \left(1 - \frac{m}{r}\right)^{-2} \left(dr^2 + (r-m)^2 d\Omega^2\right) \ . \end{aligned} \tag{31.33}$$

Introducing

$$\tilde{r} = r - m \ , \tag{31.34}$$

one thus has

$$ds^2 = -\left(1 + \frac{m}{\tilde{r}}\right)^{-2} dt^2 + \left(1 + \frac{m}{\tilde{r}}\right)^2 \left(d\tilde{r}^2 + \tilde{r}^2 d\Omega^2\right) \tag{31.35}$$

Since $d\tilde{r}^2 + \tilde{r}^2 d\Omega^2 = d\vec{x}^2$ is the flat Euclidean metric, this is the extremal Reissner-Nordstrøme metic in *isotropic* coordinates (see (24.7), and, for comparison, (24.46) for the Schwarzschild metric in isotropic coordinates). Choosing q positive so that $m = q$, we can write the metric in the suggestive form

$$ds^2 = -\left(1 + \frac{q}{|\vec{x}|}\right)^{-2} dt^2 + \left(1 + \frac{q}{|\vec{x}|}\right)^2 d\vec{x}^2 \quad . \tag{31.36}$$

Remarks:

1. In Euclidean space, $\vec{x} = 0 \leftrightarrow \tilde{r} = 0$ would just be a single point, the origin, but this is clearly not the case here. The isotropic form of the metric only describes the outside of the black hole (the horizon being at $\tilde{r} = 0$). There is also no singularity in the angular part of the metric as $\tilde{r} \to 0$, the radius of the two-sphere as $\tilde{r} \to 0$ being m (as it should be),

$$\lim_{\tilde{r}\to 0} \left(1 + \frac{m}{\tilde{r}}\right)^2 \tilde{r}^2 = m^2 \quad . \tag{31.37}$$

2. The metric (31.36) has a remarkable generalisation, where the function

$$H(\vec{x}) = 1 + \frac{q}{|\vec{x}|} \tag{31.38}$$

describing the gravitational solution for a charge q at the position $\vec{x} = 0$ is replaced by

$$H(\vec{x}) = 1 + \sum_k \frac{q_k}{|\vec{x} - \vec{a}_k|} \tag{31.39}$$

(with all $q_k > 0$, say). This is a special case of the *Majumdar-Papapetrou class of solutions* (characterised in general by an equality $\sqrt{G_N}\rho_m = \rho_e$ between the matter and charge densities), and describes a multi-centered extremal black hole solution of the Einstein-Maxwell equations, with the mutual gravitational attraction precisely balanced by the electrostatic repulsion, and with mass $m = \sum_k q_k$.

Such extremal black holes (typically characterised by the saturation of an inequality between mass and charges) arise naturally as supersymmetric solutions of supergravity theories. As such they enjoy particular stability properties and provide useful toy-models for all kinds of considerations.

31.4 Peculiar Properties of the Extremal Solution with $m^2 - q^2 = 0$

As for the Schwarzschild metric (section 26.6), it is instructive to look at the geometry of the solution in the near-horizon region $\tilde{r} \to 0$. In the Schwarzschild case (with $\tilde{r} = r - 2m$ of course) this gave us

$$ds^2 = -\frac{\tilde{r}}{2m} dt^2 + \frac{2m}{\tilde{r}} d\tilde{r}^2 + (2m)^2 d\Omega^2 \tag{31.40}$$

and then the Rindler-like metric (26.77)

$$ds^2 = -\rho^2 d\eta^2 + d\rho^2 + (2m)^2 d\Omega^2 \quad . \tag{31.41}$$

Here, because of the double pole / zero of the extremal metric at $r = m$, one finds instead that for $\tilde{r} \to 0$ one has

$$ds^2 = -\frac{\tilde{r}^2}{m^2} dt^2 + \frac{m^2}{\tilde{r}^2} d\tilde{r}^2 + m^2 d\Omega^2 \quad . \tag{31.42}$$

In particular, this metric factorises, i.e. has a product structure, the second factor just being the standard metric on the 2-sphere with constant radius m. To identify the first factor, introduce the coordinate $y = m^2/\tilde{r}$. Then one has

$$-\frac{\tilde{r}^2}{m^2} dt^2 + \frac{m^2}{\tilde{r}^2} d\tilde{r}^2 = m^2 \frac{-dt^2 + dy^2}{y^2} \tag{31.43}$$

which is nothing other than the Lorentzian counterpart (11.74) of the constant negative curvature metric on the Poincaré upper-half plane metric, (11.59), also known as the two-dimensional anti-de Sitter metric AdS_2 (with (curvature) radius m). See section 39.3 for more information on AdS metrics and coordinate systems for them.

A further transformation to radial proper distance

$$d\rho = -m\, y^{-1} dy = m\, \tilde{r}^{-1} d\tilde{r} \tag{31.44}$$

puts this metric into the form

$$m^2 \frac{-dt^2 + dy^2}{y^2} = -\mathrm{e}^{2\rho/m} dt^2 + d\rho^2 \quad . \tag{31.45}$$

Thus we can conclude that the near-horizon geometry of the extremal Reissner-Nordstrøm metric is the product geometry

$$\text{Extremal Reissner-Nordstrøm} \stackrel{\text{near horizon}}{\longrightarrow} AdS_2 \times S^2 \quad . \tag{31.46}$$

This is itself a solution of the Einstein-Maxwell equations (evidently with different asymptotics), and a particular case of the *Bertotti-Robinson* family of solutions.

<u>Remarks:</u>

1. Note that even in this near-horizon limit the location of where the horizon used to be, namely at

$$r \to m \quad \Leftrightarrow \quad \tilde{r} \to 0 \quad \Leftrightarrow \quad y \to \infty \quad \Leftrightarrow \quad \rho \sim \log \tilde{r} \to -\infty \quad , \tag{31.47}$$

is at an infinite proper distance from any point outside the horizon. This should then be (and is) a fortiori true in the original extremal Reissner-Nordstrøm metric (31.32). Indeed, proper radial distance in that metric is determined by

$$d\rho = \left(1 - \frac{m}{r}\right)^{-1} dr \quad . \tag{31.48}$$

Up to the replacement $2m \to m$ this is exactly the relation (26.91) that determined the tortoise coordinate r^* (26.92) for the Schwarzschild geometry, so that the solution is (up to a choice of integration constant)

$$\rho = r + m \log(r - m) \quad . \tag{31.49}$$

The crucial difference is that here ρ is a measure of proper distance whereas in the Schwarzschild case r^* was just a coordinate. In any case this result exhibits the logarithmic divergence $\sim \log(r - m)$ as $r \to m$. Nevertheless, the horizon can of course be reached in finite proper time along e.g. timelike geodesics, or in finite affine parameter along null geodesics. In particular, it is easy to see (cf. section 31.6) that radial ingoing lightrays are simply described by $r(\lambda) = r_0 - E\lambda$ (31.100) for a constant (energy) E, so that the affine "time" required to reach the horizon is proportional to the coordinate distance to the horizon, not the proper distance.

2. This, by the way, provides a nice and drastic illustration of the fact that one should not attempt to define something like an average velocity of a particle between two points by dividing proper spatial distance by proper time (which would come out to be not only "superluminal" in this case, but actually infinite). Proper time and proper distance measure distance along completely different paths in space-time, and dividing them is akin to dividing apples by oranges. Don't yield to the temptation to do this (unless you just want a way of quantifying small deviations from Minkowskian physics in a weak gravitational field, say, without insisting on interpreting (apples)/(oranges) as a velocity). Unfortunately, many people do not heed this advice ...

The extremal Reissner-Nordstrøm solution has another curious property, which I mention here for the record, namely a conformal isometry (i.e. isometry up to multiplication by a conformal factor) generated by a spatial inversion.[114] This is a generalisation of the perhaps better known fact that under the inversion transformation

$$r \to z = 1/r \tag{31.50}$$

the Euclidean metric on $\mathbb{R}^3$ becomes

$$dr^2 + r^2 d\Omega^2 = z^{-4}(dz^2 + z^2 d\Omega^2) \quad . \tag{31.51}$$

In the case at hand, starting with the metric in isotropic coordinates (31.35), we perform the transformation

$$y = m^2/\tilde{r} \tag{31.52}$$

[114] W. Couch, R. Torrence, *Conformal invariance under spatial inversion of extreme Reissner-Nordstrøm black holes*, Gen. Rel. Grav. 16 (1984) 789-792.

(note that this is the same coordinate transformation already used above in the analysis of the near-horizon limit of the metric), which exchanges the horizon at $\tilde{r} = 0$ and infinity. Then one readily finds

$$ds^2 = \frac{m^2}{y^2}\left(-\left(1+\frac{m}{y}\right)^{-2} dt^2 + \left(1+\frac{m}{y}\right)^2 \left(dy^2 + y^2 d\Omega^2\right)\right) \tag{31.53}$$

which is indeed again precisely the extremal Reissner-Nordstrøm metric in isotropic coordinates, up to the conformal prefactor m^2/y^2. In terms of the original radial coordinate $r = \tilde{r} + m$, the appropriate transformation is

$$r \to z = \frac{m}{1 - m/r} \quad , \tag{31.54}$$

with

$$-f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 = \frac{m^2}{(z-m)^2}\left(-f(z)dt^2 + f(z)^{-1}dz^2 + z^2 d\Omega^2\right) \quad , \tag{31.55}$$

where (of course) $f(r) = (1 - m/r)^2$.

31.5 Basic Properties of the Non-extremal Solution with $m^2 - q^2 > 0$

This is in some sense the most interesting case, not because we actually expect to find stars carrying a significant amount of electric charge, but rather because the double-horizon structure this solution exhibits is not untypical and also appears in astrophysically more relevant cases like those of rotating black holes such as the Kerr solution (however, due to the lack of spherical symmetry, the Kerr solution is only axially symmetric, the actual horizon and singularity structure of the Kerr solution is more intricate than that of the Reissner-Nordstrøm solution).

There are now two radii

$$r_{\pm} = m \pm \sqrt{m^2 - q^2} \tag{31.56}$$

at which $f(r)$ vanishes, and $f(r)$ is positive for $0 < r < r_-$ and $r > r_+$ and negative in the intermediate region $r_- < r < r_+$. In this case it is more informative to write the function $f(r)$ and the metric as

$$f(r) = \frac{(r-r_+)(r-r_-)}{r^2} \tag{31.57}$$

and

$$ds^2 = -\frac{(r-r_+)(r-r_-)}{r^2}dt^2 + \frac{r^2}{(r-r_+)(r-r_-)}dr^2 + r^2 d\Omega^2 \tag{31.58}$$

The coordinate system in which we have written the metric is valid within each of the three separate regions $r > r_+$, $r_- < r < r_+$ and $r < r_-$ (but (r,t) in one region are of course *not* the same coordinates as (r,t) in another region).

The outer radius $r_+ > m$ is and behaves just like the event horizon of the Schwarzschild metric, to which it tends for $q \to 0$,

$$q \to 0 \quad \Rightarrow \quad r_+ \to 2m \ . \tag{31.59}$$

In the intermediate region $r_- < r < r_+$ any timelike (or null) curve will then have to move from larger to smaller values of r (and will in fact reach r_- in finite proper time). At the inner radius $r_- < m$, which is absent for the Schwarzschild metric since

$$q \to 0 \quad \Rightarrow \quad r_- \to 0 \ , \tag{31.60}$$

there is also just a coordinate singularity and a horizon that reverses the role of radius and time once more so that the singularity is time-like and can be avoided by returning to larger values of r.

<u>Remarks:</u>

1. Again (i.e. as in the extremal case) we can anticipate the appearance of a new white hole region beyond the original inner horizon, through which the particle can pass back across $r = r_-$ to larger values of r, and on across $r = r_+$ to a new asymptotically flat region etc. For a more detailed discussion of this see sections 31.6 - 31.10 below.

2. This turn-around behaviour is unavoidable for timelike geodesics. Indeed we will see below that e.g. neutral massive test particles cannot even ever reach $r = 0$ since there is a repulsive core at the center of the Reissner-Nordstrøm metric. This is essentially due to the fact that the mass function $m(r) = m - q^2/2r$ (31.26) appearing in

$$f(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2} = 1 - \frac{2}{r}\left(m - \frac{q^2}{2r}\right) \tag{31.61}$$

becomes negative for sufficiently small values of r (and thus effectively acts like a negative mass). This critical (or core) radius is

$$m(r_c) = 0 \quad \Leftrightarrow \quad r_c = q^2/2m \ , \tag{31.62}$$

and always lies inside the inner horizon, $r_c < r_-$, because $f(r_c) = 1 > 0$.

3. In the extremal case we saw that the horizon $r = m$ is at an infinite proper distance from any point $r_0 > m$. Let us see what the situation is in the current non-extremal case. Thus we want to calculate the distance between r_+ and $r_0 > r_+$,

$$\rho = \int_{r_+}^{r_0} dr \frac{r}{\sqrt{(r - r_+)(r - r_-)}} = \int_0^{\tilde{r}_0} d\tilde{r} \frac{\tilde{r} + r_+}{\sqrt{\tilde{r}(\tilde{r} + \delta)}} \tag{31.63}$$

where $\tilde{r} = r - r_+$ and $\delta = r_+ - r_-$. For $\delta = 0$ we evidently recover the logarithmic divergence of the extremal case. For $\delta > 0$, on the other hand, the integral is finite (it can be expressed in closed form in terms of some unenelightning arccosh-expression, but we won't need this), the potentially dangerous piece coming from

$$\frac{r_+}{\sqrt{\tilde{r}(\tilde{r}+\delta)}} \to \frac{r_+}{\sqrt{\delta\tilde{r}}} \sim \tilde{r}^{-1/2} \tag{31.64}$$

which has a finite integral near zero.

4. While we are at it, let us perform another similar calculation which will have an amusing and perhaps unexpected consequence. Namely, let us calculate the proper (timelike) distance between the two horizons r_+ and r_-, i.e. the proper time it takes a radially freely falling observer to get from r_+ to r_-. We can use the standard (t, r) coordinates in this region between the horizons (remembering that ∂_r is timelike there) and thus, according to the usual rules, the proper time is

$$\tau = -\int_{r_+}^{r_-} \frac{dr}{\sqrt{-f(r)}} = -\int_{r_+}^{r_-} dr \frac{r}{\sqrt{-(r-r_+)(r-r_-)}} \quad . \tag{31.65}$$

Introducing the coordinate η via

$$\begin{aligned} r &= \tfrac{1}{2}(r_+ + r_-) + \tfrac{1}{2}(r_+ - r_-)\cos\eta \\ &= m + \sqrt{m^2 - q^2}\cos\eta \qquad 0 \leq \eta \leq \pi \end{aligned} \tag{31.66}$$

one has

$$dr = -\tfrac{1}{2}(r_+ - r_-)\sin\eta \, d\eta \tag{31.67}$$

and

$$(r - r_+)(r - r_-) = -\tfrac{1}{4}(r_+ - r_-)^2 \sin^2\eta \ , \tag{31.68}$$

and therefore

$$\tau = \int_0^\pi d\eta \, r(\eta) = \int_0^\pi d\eta \, \left(m + \sqrt{m^2 - q^2}\cos\eta\right) \quad . \tag{31.69}$$

So far so straightforward. The curious thing, however, is that the second term of the integrand $\sim \cos\eta$ integrates to zero over the interval $[0, \pi]$ and therefore does not contribute to the integral at all, leading to the universal result

$$\tau = m\pi \tag{31.70}$$

independent of q. Thus the proper time is $m\pi$ for any $q < m$, and also in the extremal limit $q \to m \Leftrightarrow r_- \to r_+$. In the extremal black hole, on the other hand, with $r_+ = r_-$, the coordinate distance between r_+ and $r_- = r_+$ is clearly zero. What this shows is that the extremal black hole ($q = m$) is perhaps not for all intents and purposes the same thing as the extremal limit ($q \to m$) of a non-extremal black hole, and there are also other contexts in which this distinction appears to play a role.[115]

[115] For further discussion, see e.g. S. Carroll, M. Johnson, L. Randall, *Extremal limits and black hole entropy*, `arXiv:0901.0931 [hep-th]`.

5. As a final variation of this theme, let us ask what is the maximum angular distance that a lightray can travel between r_+ and r_-.[116] Choosing the lightray to move along the θ-direction, we need to solve

$$f(r)^{-1}dr^2 + r^2 d\theta^2 = 0 \quad \Rightarrow \quad \Delta\theta = \int_{r_-}^{r_+} \frac{dr}{r\sqrt{-f(r)}} \quad . \tag{31.71}$$

Using the same parametrisation for r as above, we see that we just get the universal result

$$\Delta\theta = \int_0^\pi d\eta = \pi \quad , \tag{31.72}$$

which does not depend at all on the parameters (m, q) or $r_\pm$ of the solution. In particular there is precisely enough time/space between the two horizons for a lightray to travel from the north-pole to the south-pole of the sphere.

Curiously, this is not true in higher dimensions. The $D = d + 1$ dimensional Reissner-Nordstrøm metric (for $D > 3$) has

$$f(r) = \frac{(r^{D-3} - r_+^{D-3})(r^{D-3} - r_-^{D-3})}{r^{2(D-3)}} \quad . \tag{31.73}$$

With the change of variables

$$r^{D-3} = \tfrac{1}{2}(r_+^{D-3} + r_-^{D-3}) + \tfrac{1}{2}(r_+^{D-3} - r_-^{D-3})\cos\eta \tag{31.74}$$

and

$$r^{D-3}dr = \frac{r}{D-3} d(r^{D-3}) \tag{31.75}$$

the integral for $\Delta\theta$ again trivialises, and one finds

$$\Delta\theta = \frac{1}{D-3}\int_0^\pi d\eta = \frac{\pi}{D-3} \quad . \tag{31.76}$$

As in the case of the Schwarzschild metric, the coordinates (t, r, θ, ϕ) employed so far are not for all purposes the most convenient coordinates (and we will discuss the analogues of Eddington-Finkelstein and Kruskal coordinates in quite some detail in sections 31.7-31.10 below). Here we take a brief look at some other commonly used coordinate systems.

In the extremal case, it turned out to be useful and instructive to write the metric in isotropic coordinates (31.36). Equivalently, it was useful to introduce the coordinate distance $\tilde{r} = r - m$ to the extremal horizon, and the resulting metric turned out to be in isotropic form.

Extending this construction to the non-extremal case, these two strategies result in two different coordinate systems that are occasionally useful, true isotropic coordinates and

[116] See e.g. B. Freivogel, R. Jefferson, L. Kabir, I-S. Yang, *Geometry of the Infalling Causal Patch*, `arXiv:1406.6043 [hep-th]`.

another coordinate system which I will refer to as "brane coordinates" because it the prototype coordinate system in which one usually writes solutions to higher-dimensional (super-)gravity theories describing black spatially extended objects. These are known as black p-branes, with "brane" extracted from "membrane", so that a 2-brane is a membrane and in the case at hand we are dealing with a 0-brane.[117]

1. Isotropic Coordinates

   Recall that for the Schwarzschild metric, the isotropic radial coordinate is (24.45)

   $$r(\rho) = \rho(1+\frac{m}{2\rho})^2 = \rho + m + \frac{m^2}{4\rho} \quad , \tag{31.77}$$

   and that for the extremal Reissner-Nordstrøm metric we found (31.34) (with $\tilde{r} \to \rho$)

   $$r(\rho) = \rho + m \quad . \tag{31.78}$$

   In the present non-extremal case, we introduce ρ via

   $$r(\rho) = \rho + m + \frac{m^2 - q^2}{4\rho} \tag{31.79}$$

   which interpolates nicely between the two previous expressions for $q = 0$ and $q^2 = m^2$. Then it is straightforward to check that

   $$r^2 - 2mr + q^2 = \left(\rho - \frac{m^2-q^2}{4\rho}\right)^2 \equiv \Delta(\rho) \quad . \tag{31.80}$$

   and that

   $$dr = \frac{\Delta(\rho)^{1/2}}{\rho} d\rho \quad . \tag{31.81}$$

   As a consequence, the Reissner-Nordstrøm metric can be written as

   $$ds^2 = -\frac{\Delta(\rho)}{r(\rho)^2} dt^2 + \frac{r(\rho)^2}{\rho^2}(d\rho^2 + \rho^2 d\Omega^2) \quad . \tag{31.82}$$

   which is of the characteristic isotropic form, with spatial metric $\sim d\rho^2 + \rho^2 d\Omega^2$.

2. Brane Coordinates

   It is also possible to write the metric in a form which interpolates nicely between the standard Schwarzschild form of the Schwarzschild metric for $q \to 0$ and the nice isotropic form (31.35) of the extremal metric in the extremal limit $q \to m$. To that end, all we need to do is replace r by $\rho = r - r_-$, which reduces to the isotropic coordinate $\rho \to \tilde{r} = r - m$ in the extremal limit and, moreover, to the standard radial coordinate $\rho \to r$ in the Schwarzschild limit $r_- \to 0$ (unlike the $\tilde{r} = r - r_+$ introduced in remark 3 above, which has the Schwarzschild limit $\tilde{r} \to r - 2m$).

[117] See e.g. T. Ortin, *Gravity and Strings* for a detailed account of these objects and their uses and applications.

In terms of $\rho = r - r_-$, and with

$$\delta = r_+ - r_- = 2\sqrt{m^2 - q^2} \tag{31.83}$$

measuring the deviation from extremality, the metric takes the form

$$ds^2 = -H(\rho)^{-2}F(\rho)dt^2 + H(\rho)^2\left[F(\rho)^{-1}d\rho^2 + \rho^2 d\Omega^2\right] \tag{31.84}$$

with

$$H(\rho) = 1 + \frac{r_-}{\rho} \quad , \quad F(\rho) = 1 - \frac{\delta}{\rho} \quad . \tag{31.85}$$

Observe that this manifestly reduces to (31.35) in the extremal limit,

$$\delta \to 0 \quad \Rightarrow \quad \begin{cases} F(\rho) & \to 1 \\ r_- & \to m \\ H(\rho) & \to 1 + m/\rho \end{cases} \tag{31.86}$$

and to the Schwarzschild metric in the limit $q \to 0 \leftrightarrow r_- \to 0$,

$$q \to 0 \quad \Rightarrow \quad \begin{cases} H(\rho) & \to 1 \\ \delta & \to 2m \\ F(\rho) & \to 1 - 2m/\rho \end{cases} \quad . \tag{31.87}$$

As mentioned above, the form of the metric in (31.84) is representative of the form of a large class of solutions of higher-dimensional (super-)gravity theories describing black holes and black branes (spatially extended black objects). It also nicely illustrates a general recipe:

- Given an extremal solution, one can introduce non-extremality (or temperature) by "dressing" the extremal solution with Schwarzschild-like factors F, as in (31.84).
- Given a neutral (uncharged) solution, one can construct a charged solution by "dressing" the neutral solution with harmonic functions H, as in (31.84).

31.6 Motion of a Charged Particle: the Effective Potential

We now consider the motion of a test particle with mass μ and charge e in the Reissner - Nordstrøm space-time with $m^2 > q^2$. This is evidently described by the geodesic equation modified by the Lorentz-force term,

$$\ddot{x}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{x}^\beta\dot{x}^\gamma = (e/\mu)F^\alpha_{\ \beta}\dot{x}^\beta \quad . \tag{31.88}$$

We will set $\mu = 1$ in the following (the case $\mu \neq 1$ can obviously be recovered from this by scaling $e \to e/\mu$), so that quantities associated to the particle like charge, energy and angular momentum that appear below are, as usual (cf. the discussion in section 3.1) to be thought of as quantities per unit particle mass.

Proceeding in exact analogy with the derivation of the effective potential for geodesics in the Schwarzschild geometry in section 25.3, in order to exploit the symmetries and conserved charges of the system it will be convenient to work at the level of the Lagrangian which we can choose to be

$$\mathcal{L} = \tfrac{1}{2} g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta + eA_\alpha \dot{x}^\alpha \tag{31.89}$$

Plugging in the metric and gauge field (and choosing rightaway equatorial paths at $\theta = \pi/2$ because of spherical symmetry), this Lagrangian becomes more explicitly

$$\mathcal{L} = \tfrac{1}{2}(-f\dot{t}^2 + f^{-1}\dot{r}^2 + r^2\dot{\phi}^2) - eQ\dot{t}/r \quad . \tag{31.90}$$

We now introduce the conserved quantities E (energy) associated to time-translation invariance,

$$E = -\frac{\partial \mathcal{L}}{\partial \dot{t}} = f\dot{t} - eA_t \quad \Leftrightarrow \quad \dot{t} = f^{-1}(E + eA_t) = f^{-1}(E - eQ/r) \tag{31.91}$$

and $L = r^2\dot{\phi}$ (angular momentum) associated to rotation invariance. Note that L has the standard form, and that E is shifted relative to the same quantity for the Schwarzschild solution precisely by the Coulomb electrostatic potential energy, $E \to E - eQ/r$. Plugging this into

$$\epsilon \equiv g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta = -1 \tag{31.92}$$

(the condition that particle trajectories are time-like and parametrised by proper time) one deduces

$$-1 = -f^{-1}(E-eQ/r)^2 + f^{-1}\dot{r}^2 + L^2/r^2 \quad \Leftrightarrow \quad \dot{r}^2 = (E-eQ/r)^2 - f(1+L^2/r^2) \tag{31.93}$$

In particular, for purely radial paths, $L = 0$, one has

$$\dot{r}^2 = (E - eQ/r)^2 - f \quad , \tag{31.94}$$

or, separating out the constant contributions,

$$\tfrac{1}{2}\dot{r}^2 + \left(\frac{eEQ - m}{r} + \frac{q^2 - e^2Q^2}{2r^2}\right) = \tfrac{1}{2}(E^2 - 1) \quad . \tag{31.95}$$

The term in brackets defines the effective potential (for radial motion)

$$V_{eff}(r) = \frac{eEQ - m}{r} + \frac{q^2 - e^2Q^2}{2r^2} = \frac{eEQ - G_N M}{r} + Q^2\frac{G_N - e^2}{2r^2} \quad , \tag{31.96}$$

and the term on the right-hand side defines the effective energy E_{eff}, so that, as in the case of the Schwarzschild metric, we can write the (first integral of the) radial geodesic equation in the suggestive Newtonian form

$$\tfrac{1}{2}\dot{r}^2 + V_{eff}(r) = E_{eff} \quad . \tag{31.97}$$

From this one can then readily deduce the qualitative features of the worldlines of charged and uncharged particles in the Reissner-Nordstrøm geometry.

REMARKS:

1. For $L \neq 0$ this gets modified to

$$V_{eff}(r) \to V_{eff}(r) + f(r)L^2/2r^2 \quad , \tag{31.98}$$

which includes the usual angular mometum barrier term $\sim L^2/r^2$ as well as the familiar attractive general relativistic correction term $\sim mL^2r^{-3}$ and a novel repulsive correction term $\sim q^2L^2r^{-4}$. However, we will focus on radial motion in the following.

2. For massless particles ($\epsilon = 0$), which we will of course also consider to be uncharged ($e = 0$), the effective potential consists of just the above angular momentum term,

$$e = \epsilon = 0: \quad \dot{r}^2 + f(r)L^2/r^2 = E^2 \quad . \tag{31.99}$$

In particular, for radial lightrays the effective potential is zero and $\dot{r} = \pm E$ for outgoing (respectively ingoing) null geodesics, so that

$$r(\lambda) = r_0 \pm E\lambda \quad , \tag{31.100}$$

with λ the affine parameter. Thus massless particles can reach the horizon (and $r = 0$) in finite affine time.

3. For certain purposes it is also useful to write the effective potential in terms of the horizons $r_\pm$ instead of the parameters m and q using the relations

$$r_+ + r_- = 2m \quad , \quad r_+r_- = q^2 \quad . \tag{31.101}$$

Then one has, with $\tilde{e} = e/\sqrt{G_N}$,

$$V_{eff}(r) = \frac{2\tilde{e}E\sqrt{r_+r_-} - (r_+ + r_-)}{2r} + \left(1 - \tilde{e}^2\right)\frac{r_+r_-}{2r^2} \quad . \tag{31.102}$$

4. The interpretation of the first term in (31.96) is pretty clear: it describes the competition between the leading Coulomb electrostatic and Newton gravitational $1/r$ interactions between the charged massive star and the charged massive test particle. The only thing that may require some explanation is the factor of E in the Coulomb interaction. As we know from the discussion of Schwarzschild, E is essentially the special-relativistic γ-factor and thus the substitution $Q \to EQ$ accounts for the Lorentz contraction of the electric field lines as seen by a particle with velocity $\dot{r}_\infty^2 = E^2 - 1$ at $r = \infty$.

5. The second term is more mysterious and interesting in several respects. First of all, for a neutral test particle freely falling (following a geodesic) in the Reissner - Nordstrøm geometry, it provides a repulsive potential at short distances,

$$e = 0 \quad \Rightarrow \quad V_{eff}(r) = -\frac{m}{r} + \frac{q^2}{2r^2} \quad . \tag{31.103}$$

mimicking the angular momentum barrier term $L^2/2r^2$. This inevitably leads to a turning point of the trajectory, and we will see below that this turning point lies inside the inner horizon, i.e. in the region $r < r_-$. A heuristic but not entirely satisfactory explanation for the occurrence of this phenomeneon is that this is due to a *mass renormalisation*

$$m \to m(r) = m - \frac{q^2}{2r} \tag{31.104}$$

required to compensate the infinite electrostatic energy density $\sim q^2/r^4$ of the star. Alternatively one may take this as an indication that the interior of the Reissner-Nordstrøm solution is not particularly physical (we will come back to this below).

6. For a charged particle, the coefficient $q^2 = G_N Q^2$ in the $1/r^2$ term of the potential is replaced by $Q^2(G_N - e^2)$. Recalling that e is really the charge per unit mass and replacing $e \to e/\mu$, one sees that the sign of this term is determined by the sign of $G_N^2\mu^2 - G_N e^2$, i.e. the relative size of the gravitational mass and charge radii of the test particle. Reverting to $\mu = 1$, we will call ordinarily charged particles those for which $e^2 < G_N$, extremal those with $e^2 = G_N$ and overcharged those with $e^2 > G_N$. Thus the $1/r^2$ term in the radial effective potential provides a repulsive potential at short distances for all ordinarily charged particles, but this term becomes attractive for overcharged particles.

   Note that this not somehow an electrostatic effect (in particular since it is independent of the sign of the charge e of the particle) but a purely gravitational effect. I do not have a good heuristic explanation for why overcharged particles all of a sudden experience an attractive $1/r^2$ potential (and would be glad to learn of one ...).

7. If one gives the particle some angular momentum, no matter how tiny, i.e. if there is just the slightest deviation from radial motion, then the term q^2L^2/r^4 will kick in at short distances to yet again provide a repulsive potential as a joint effect of the charge of the black hole and the angular momentum of the particle.

Let us now take a closer look at some of the possible trajectories. First of all, for a given choice of parameters the allowed values of r are constrained by the condition

$$\dot{r}^2 = (E - eQ/r)^2 - f(r)(1 + L^2/r^2) \geq 0 \tag{31.105}$$

and at a turning point r_m (for maximal or minimal radius) of the trajectory one has

$$(E - eQ/r_m)^2 = f(r_m)(1 + L^2/r_m^2) \quad \Rightarrow \quad f(r_m) \geq 0 \quad \Rightarrow \quad r_m \geq r_+ \quad \text{or} \quad r_m \leq r_- \tag{31.106}$$

Thus turning points can only occur either in the region outside the outer horizon or in the region inside the inner horizon.

A simple (but nevertheless for present purposes sufficiently prototypical) example is the radial free fall of an uncharged particle, described by the simple effective potential (31.103). Turning points are determined by

$$-\frac{2m}{r_m} + \frac{q^2}{r_m^2} = E^2 - 1 \quad . \tag{31.107}$$

For a particle initially at rest at infinity, $E = 1$, one immediately reads off that the minimal radius is equal to the core radius r_c (31.62),

$$r_m(E=1) = r_c = \frac{q^2}{2m} = \frac{r_+ r_-}{r_+ + r_-} = \frac{r_-}{1 + (r_-/r_+)} < r_- \quad . \tag{31.108}$$

It is also plausible (and moreover true) that the particle will penetrate slightly deeper into the Reissner-Nordtstrøm core if it initially has a non-zero inward directed velocity, i.e.

$$E > 1 \quad \Rightarrow \quad r_m < r_c \quad , \tag{31.109}$$

but clearly no finite energy particle can overcome the charge barrier to reach $r = 0$. These particles will turn around at $r_m(E)$ and then escape again to infinity in a new branch of the universe (since they clearly can't cross the same inner horizon r_- in both directions).

Particles with $E < 1$ have both a minimum and a maximum radius $r_{m\pm}$, located in the regions $r_{m-} < r_-$ and $r_{m+} > r_+$ respectively. Thus these particles appear to oscillate in and out of the black hole region $r_- < r < r_+$ but this is clearly not possible (if r_+ is a black hole horizon, you cannot just dance around and oscillate in and out of it to your heart's content). What is happening is that, after having reached its inner turning point at $r_{m-} < r_-$, the particle turns around to larger values of r, crosses a new $r = r_-$ horizon into a new (time-reversed) version of the region $r_- < r < r_+$ in which it can only move to larger values of r, crosses a white hole horizon at $r = r_+$ into a new asymptotically flat Reissner-Nordstrøm patch, up to the maximal radius r_{m+} allowed by its energy, and then turns around again to enter another new region etc. etc.

This suggests that somehow the maximally extended Reissner-Norstrøm solution consists of an infinite sequence of such universes patched together along the horizons, and this indeed turns out to be the case. Moreover, none too surprisingly the analysis reveals that, as in the Kruskal diagram for the Schwarzschild metric, there is in addition a mirror region, and therefore also an infinite sequence of such mirror regions. The resulting Penrose diagram, in its full glory (well, almost full glory, I had to truncate it somewhere) is shown in Figure 34. While this looks quite crazy, we will substantiate this picture somewhat below by constructing coordinates that allow us to (patchwise) cover all these regions of the extended space-time.

We can also consider charged particles. Their behaviour depends strongly on whether they are ordinarily charged or overcharged particles (the latter having a regretful suicidal

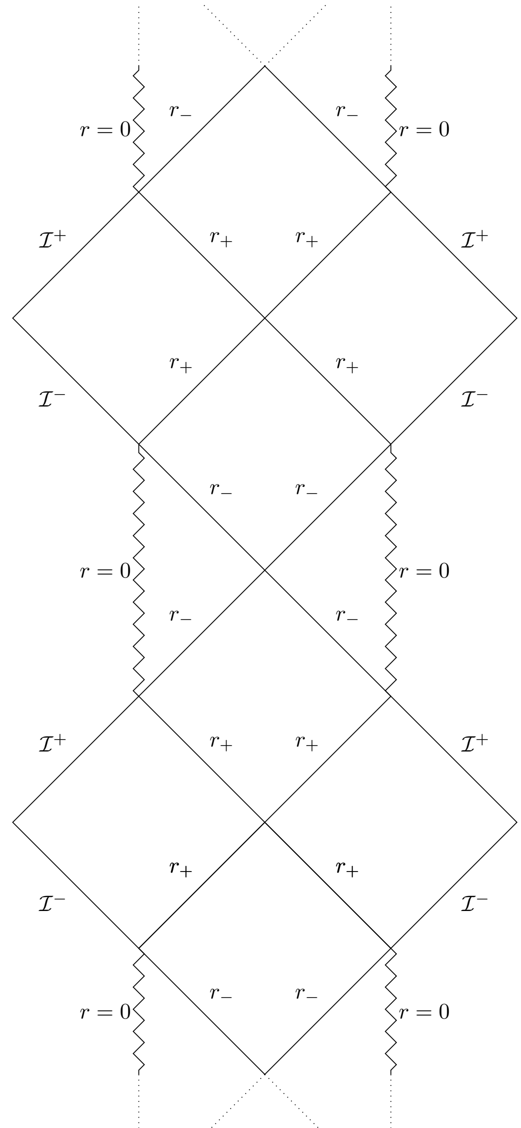


Figure 34: Penrose Diagram of the maximal analytic extension of the non-extremal Reissner-Nordstrøm black hole.

tendency to end up in the singularity at $r = 0$ regardless of the sign of the charge), but also on the energy and on whether the Coulomb or gravitational $1/r$ interaction is dominant. Thus this requires a bit of a case by case analysis which we will not pursue here. Suffice it to say here that for all ordinarily charged particles one finds that the first turning point of a radially infalling particle is located inside the inner horizon (and not outside the outer horizon, which would, in principle, have been the other option).

In order to get a better picture of the Reissner-Nordstrøm geometry, and in order to provide e.g. coordinates for the particles that appear to oscillate in and out of the black hole / white hole regions, we now introduce ingoing and outgoing Eddington-Finkelstein coordinates, and we can do this in exactly the same way as for the Schwarzschild metric.

In fact, the basic construction works for any metric of the form

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \tag{31.110}$$

where we assume $f(r)$ to have a zero at $r = r_h$.

1. We first introduce the (generalised) tortoise coordinate r^* via

$$ds^2 = f(r)[-dt^2 + (dr^*)^2] + r(r^*)^2 d\Omega^2 \qquad dr^* = f(r)^{-1}dr \quad . \tag{31.111}$$

2. We also introduce the retarded and advanced time-coordinates u and v by

$$u = t - r^* \quad , \quad v = t + r^* \quad . \tag{31.112}$$

   Infalling radial null geodesics ($dr^*/dt = -1$) are characterised by $v = const.$ and outgoing radial null geodesics by $u = const.$

3. In terms of (v, r) the metric reads

$$ds^2 = -f(r)dv^2 + 2dv\, dr + r^2 d\Omega^2 \quad . \tag{31.113}$$

   (with an analogous expression for the metric written in terms of (u, r)). This metric is now regular at any zero r_h of f, $f(r_h) = 0$.

4. In these coordinates, the Killing vector $\xi = \partial_t$ of the original space-time, rendering the "outside" region $f(r) > 0$ static, takes the form

$$\partial_t \to \xi = (\partial_t v)\partial_v + (\partial_t r)\partial_r = \partial_v \quad . \tag{31.114}$$

5. Ingoing radial lightrays are characterised by $dv = 0$, and r is an affine parameter along these null geodesics. Outgoing radial lightrays are characterised by $2dr = f(r)dv$, i.e. by

$$2dr = f(r)dv \quad \Leftrightarrow \quad du = d(t - r^*) = d(t + r^*) - 2dr^* = dv - 2f(r)^{-1}dr = 0 \quad . \tag{31.115}$$

   In particular, at r_h the "outgoing" lightrays satisfy $dr = 0$, i.e. the surface $r = r_h$ is lightlike, and ξ becomes tangent to this surface there. So again we have a *Killing horizon* (to be discussed in more detail in section 32.5).

6. If $f(r)$ changes sign from $f(r) > 0$ to $f(r) < 0$ as one moves from $r > r_h$ to $r < r_h$, then the situation is identical to that for the Schwarzschild black hole:

   - the Killing vector ∂_v becomes null on the horizon and spacelike for for $r < r_h$
   - $r = r_h$ is an event horizon and r will decrease along any future-directed causal (timelike or lightlike) path.

   On the other hand, if $f(r)$ is non-negative, with a minimum $f(r_h) = 0$, the situation is qualitatively quite different: the metric possesses an everywhere *causal* Killing vector, null on the horizon and timelike away from the horizon.

31.8 Eddington-Finkelstein Coordinates: the Reissner-Nordstrøm Metric

Now let us specialise to the non-extremal Reissner-Nordstrøm metric, which in Eddington-Finkelstein coordinates has the form

$$ds^2 = -\frac{(r-r_+)(r-r_-)}{r^2}dv^2 + 2dv\,dr + r^2 d\Omega^2 \ . \tag{31.116}$$

In that case $f(r)$ has simple zeros at $r = r_+$ and $r = r_-$, where $f(r)$ changes sign. Crossing r_+ from $r > r_+$ to $r < r_+$ is then like entering the Schwarzschild black hole region, while crossing r_- from $r > r_-$ to $r < r_-$ is (somewhat) like exiting the Schwarzschild white hole region (with the difference that in the present case one is expelled from the region $r_- < r < r_+$ in the direction of decreasing r).

Let us now see what happens to the lightcones, and what are therefore the allowed paths for massless or massive particles, as one enters the Reissner-Nordstrøm black hole through the future outer horizon.

1. The first thing that will happen is, as already discussed above, and exactly as in the Schwarzschild case, that the lightcones tilt over at $r = r_+$. Subsequently both ingoing and (misleadingly still called) “outgoing” lightrays will converge to smaller values of r,

   $$2dr = f(r)dv \quad \Rightarrow \quad dr/dv < 0 \quad \text{for} \quad r_- < r < r_+ \ . \tag{31.117}$$

   In particular, once inside one must continue to smaller values of r until one reaches either a singularity (as for Schwarzschild) or another horizon.

2. In the Reissner-Nordstrøm case there is indeed such a horizon, namely at $r = r_-$. The ingoing Eddington-Finkelstein coordinates are still valid at the inner horizon $r = r_-$ because the metric is regular there, the horizon sits (tautologically) at the finite value $r = r_-$ of the radial coordinate, while on that surface the coordinate

v (and the angular coordinates) still label the ingoing lightrays crossing the inner horizon. Indeed, the coordinates continue to be valid all the way up to $r = 0$.

Note that for this we do not need to know if the tortoise coordinate r^* is well-behaved at $r = r_-$. As a matter of fact, it is not, but this does not affect the Eddington-Finkelstein coordinates. It will, however, affect the Kruskal-Szekeres coordinates which will break down at $r = r_-$ (but have a larger region of validity across the past white hole horizon of the original Reissner-Nordstrøm patch).

3. At $r = r_-$, the function $f(r)$ again changes sign and outgoing lightrays are indeed outgoing, i.e. moving to larger values of r. Once these outgoing lightrays in the region $r < r_-$ reach the new (white hole) inner horizon at $r = r_-$, the original set of Eddington-Finkelstein coordinates (v, r) finally break down since v can only label the ingoing lightrays and the new r_- sits at advanced time $v = \infty$.

4. However, in this patch $r < r_-$ (whose metric is identical to that in the outside patch $r > r_+$), one can also construct and use outgoing Finkelstein coordinates (u, r), with u labelling the outgoing lightrays. These coordinates will not only cover the region $0 < r < r_-$, but they will extend across the new white hole inner and outer horizons $r_\mp$ into a new asymptotically flat Reissner-Nordstrøm patch.

5. From that region one can in principle continue into a new black hole region across another r_+, but now it is evidently the outgoing Eddington-Finkelstein coordinates that break down and one returns to step 1 and again constructs ingoing Eddington-Finkelstein coordinates to describe this.

6. It is now evident that, proceeding in this way, one can pave / tessellate the entire infinitely periodic fully extended non-extremal Reissner-Nordstrøm solution with ingoing and outgoing Eddington-Finkelstein coordinates (whose domains of validity overlap in the regions $r < r_-$ and $r > r_+$ where $f(r) > 0$). This is indicated in the Penrose diagram in Figure 35.

In the extremal case, when there is a double zero of $f(r)$, the story is similar, the only difference being that the region between r_- and r_+ is absent. The metric has the form

$$ds^2 = -\left(1 - \frac{m}{r}\right)^2 dv^2 + 2dv\, dr + r^2 d\Omega^2 \quad . \tag{31.118}$$

In particular, $r = m$ is a null surface and a horizon. In the patch covered by the above ingoing Eddington-Finkelstein coordinates one can only cross it along future-directed curves. The diffference is that now $f(r)$ is positive on both sides of the horizon, so that outgoing lightrays are really outgoing on both sides of the horizon. Ingoing Eddington Finkelstein coordinates can still also cover the patch $r < m$, but they cannot describe the new outgoing region beyond the new white whole inner horizon $r = m$. For this one needs to introduce outgoing Eddington-Finkelstein coordinates etc. Again one can pave

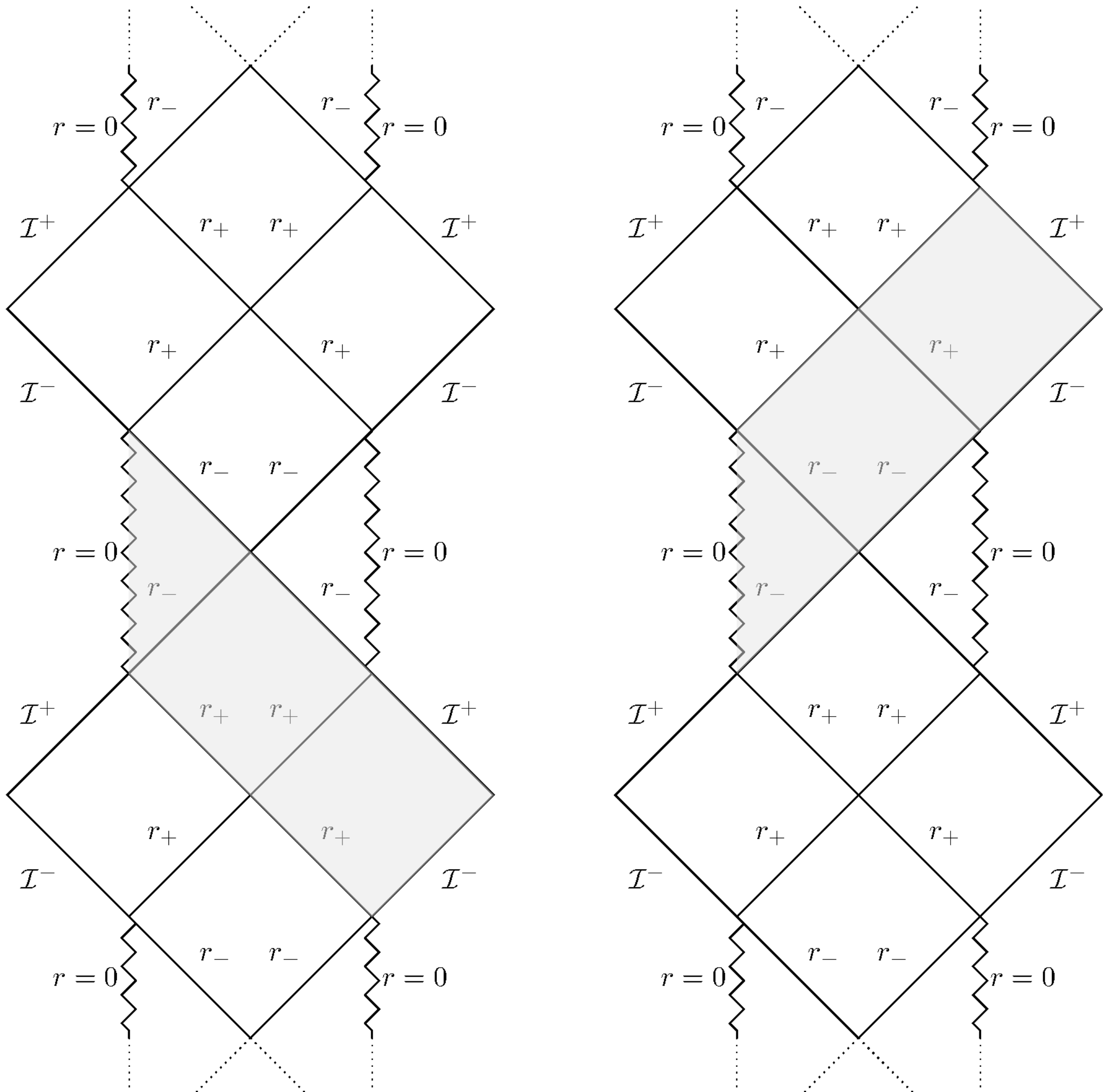


Figure 35: Penrose Diagram of the Reissner-Nordstrøm metric: regimes of validity of one set of ingoing (left panel) and outgoing (right panel) Eddington-Finkelstein coordinates are indicated by the shaded areas. These overlap in the regions $r < r_-$ or $r > r_+$ and thus one can cover the entire maximal extension with such coordinates and their mirror counterparts.

the infinitely periodic extremal Reissner-Nordstrøm geometry in this way, as indicated in Figure 36.

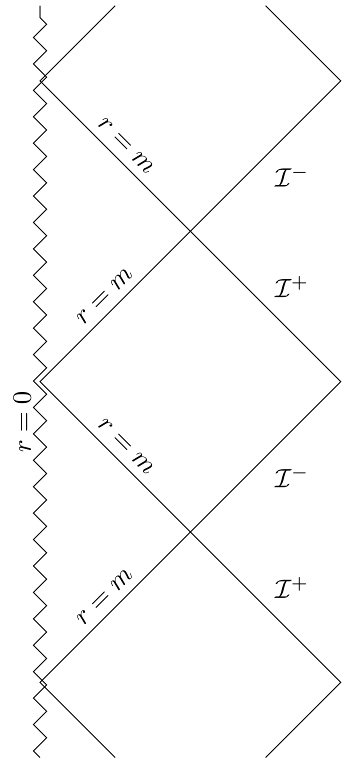


Figure 36: Penrose Diagram of the extremal Reissner-Nordstrøm metric.

31.9 Kruskal-Szekeres Coordinates: General Considerations

One can also introduce Kruskal-Szekeres coordinates via the same chain of transformations

$$(t,r) \to (u,v) \to (U,V) \to (T,X) \tag{31.119}$$

as in the Schwarzschild case. We again start by setting up the problem in the general context of metrics of the form (31.111),

$$ds^2 = f(r)[-dt^2 + (dr^*)^2] + r(r^*)^2 d\Omega^2 = -f(r)du\, dv + r(u,v)^2 d\Omega^2 \ , \tag{31.120}$$

and we will now be more specific and assume that $f(r)$ has a simple zero at $r = r_h$.

Since the issue is the elimination of the coordinate singularity at r_h, we can focus on the behaviour of $f(r)$ near $r = r_h$,

$$f(r) \approx (r - r_h) f'(r_h) + \ldots \tag{31.121}$$

(here is where the treatment for a double zero, say, would of course be different). Thus the tortoise coordinate can be approximated by

$$dr^* = f(r)^{-1} dr \approx \frac{dr}{(r - r_h) f'(r_h)} \quad \Rightarrow \quad r^* \approx \frac{1}{f'(r_h)} \log |r - r_h| \ . \tag{31.122}$$

Here $f'(r_h)$ has the same dual interpretation (27.177) as in section 27.10, namely on the one hand as the *inaffinity*, measuring the failure of the coordinate v to provide an affine parametrisation of the horizon generators $\xi = \partial_v$, and on the other hand as the *surface gravity*, providing a measure of the strength of the gravitational field at the horizon,

$$\kappa_h = \tfrac{1}{2}f'(r)|_{r=r_h} = \begin{cases} \text{inaffinity:} & \lim_{r\to r_h} \nabla_\xi \xi = \kappa_h \xi \\ \text{surface gravity:} & \kappa_h := \lim_{r\to r_h} f(r)^{1/2}\mathsf{a}(r) \end{cases} \tag{31.123}$$

To see this note that the derivation of the inaffinity given in (27.173) and (27.174) goes through verbatim in general,

$$\begin{aligned} (\nabla_\xi\xi)^\alpha\partial_\alpha &= \Gamma^\alpha_{vv}\partial_\alpha = (f'/2)(f\partial_r + \partial_v) \\ \lim_{r\to r_h} \nabla_\xi\xi &= \left(\tfrac{1}{2}f'(r)|_{r=r_h}\right)\xi = \kappa_h\xi \quad . \end{aligned} \tag{31.124}$$

and that the generalisation of the calculation of the acceleration for a static observer in section 26.2 gives

$$u^\alpha = (f(r)^{-1/2}, 0, 0, 0) \quad \Rightarrow \quad a^r = \tfrac{1}{2}f'(r) \quad . \tag{31.125}$$

In terms of κ_h, the approximate expression for the tortoise coordinate is

$$r^* \approx \frac{1}{2\kappa_h}\log|r - r_h| \tag{31.126}$$

(the suppressed subleading terms being regular as $r \to r_h$). Then near $r = r_h$ the function $f(r)$ can be approximated by

$$f(r) \approx f'(r_h)(r - r_h) \approx 2\kappa_h \mathrm{e}^{2\kappa_h r^*} = 2\kappa_h \mathrm{e}^{\kappa_h(r^* - t) + \kappa_h(r^* + t)} = 2\kappa_h \mathrm{e}^{-\kappa_h u}\mathrm{e}^{+\kappa_h v} \tag{31.127}$$

and for the (t, r)-part of the metric one has

$$ds^2 = -f(r)du\,dv \approx -2\kappa_h \mathrm{e}^{-\kappa_h u}du\,\mathrm{e}^{\kappa_h v}dv = -2\kappa_h^{-1}\,d\left(-\mathrm{e}^{-\kappa_h u}\right)\,d\left(+\mathrm{e}^{+\kappa_h v}\right) \quad . \tag{31.128}$$

We thus introduce the Kruskal coordinates (u_K, v_K) (with the normalisation as in (27.156)) by

$$u_K = -(\kappa_h)^{-1}\mathrm{e}^{-\kappa_h u} \quad , \quad v_K = +(\kappa_h)^{-1}\mathrm{e}^{+\kappa_h v} \tag{31.129}$$

and (if one doesn't like null coordinates) one can also introduce new time- and space-coordinates (t_K, x_K) via

$$u_K = t_K - x_K \quad , \quad v_K = t_K + x_K \quad , \tag{31.130}$$

and we have just seen that in terms of these Kruskal coordinates the metric near the horizon at $r = r_h$ takes the manifestly non-singular form

$$ds^2 = -f(r)du\,dv + r(u,v)^2 d\Omega^2 \approx -C du_K\,dv_K + r_h^2 d\Omega^2 \quad . \tag{31.131}$$

The precise value of the coefficient $C > 0$ will depend on the non-singular terms in r^* (which we have suppressed here), evaluated at $r = r_h$, and on the choice of integration constants, and is therefore arbitrary, and also irrelevant.

Thus these Kruskal coordinates provide us with a good system of coordinates not just in the original patch where the coordinates (t, r) were valid, but also across the future horizon at $u_K = 0$ and the past horizon at $v_K = 0$. In the Schwarzschild case, more than that was true, namely the Kruskal coordinates provided us with a coordinate system for the complete maximal extension of the Schwarzschild geometry. This is, however, not guaranteed by the above general construction and need not, and will not, be true in general.

31.10 Kruskal-Szekeres Coordinates: the Reissner-Nordstrøm Metric

Indeed, already for the Reissner-Nordstrøm metric we will be able to see this explicitly. To that end we will need the explicit expressions for the tortoise coordinate. In the non-extremal case we have

$$f(r) = \frac{(r - r_+)(r - r_-)}{r^2} \quad . \tag{31.132}$$

The surface gravities at the two horizons $r_\pm$ are

$$\kappa_\pm = \tfrac{1}{2} f'(r_\pm) = \pm \frac{r_+ - r_-}{2r_\pm^2} \quad . \tag{31.133}$$

In terms of these we can write $f(r)^{-1}$ as

$$f(r)^{-1} = 1 + \frac{1}{2\kappa_+} \frac{1}{r - r_+} + \frac{1}{2\kappa_-} \frac{1}{r - r_-} \quad , \tag{31.134}$$

and it is elementary to determine r^* in this case,

$$dr^* = \frac{r^2}{(r - r_+)(r - r_-)} dr \quad \Rightarrow \quad r^* = r + \frac{1}{2\kappa_+} \log |r - r_+| + \frac{1}{2\kappa_-} \log |r - r_-| \quad . \tag{31.135}$$

Thus $r_* \to -\infty$ as $r \to r_+$ and (recalling that $\kappa_- < 0$) $r^* \to +\infty$ as $r \to r_-$. Around each horizon one can introduce Kruskal coordinates, say

$$u_{K+} = -(\kappa_+)^{-1} \mathrm{e}^{-\kappa_+ u} \quad , \quad v_{K+} = +(\kappa_+)^{-1} \mathrm{e}^{+\kappa_+ v} \tag{31.136}$$

to cover the region around r_+. This coordinate system does not just cover the future black hole region (until $r = r_-$, see below) but (as for Schwarzschild) also a past white hole region (until $r = r_-$) and a mirror asymptotically flat region of the Reissner-Nordstrøm patch.

However, these coordinates break down as $r \to r_-$. This can be seen from the explicit expression of the metric in these coordinates, which display a (coordinate) singularity at $r = r_-$ (but we will forego this here). It can also be seen, more directly, and more to the point, from the fact that, as noted above, $r^* \to \infty$ for $r \to r_-$, so that

$$r \to r_- \quad \Rightarrow \quad r^* \to \infty \quad \Rightarrow \quad v - u \to \infty \quad \to \quad u_{K+} v_{K+} \to \infty \ , \tag{31.137}$$

so that the outer Kruskal coordinates u_{K+}, v_{k+} are singular at, and can therefore not be extended beyond, the inner horizon, as shown in Figure 37.

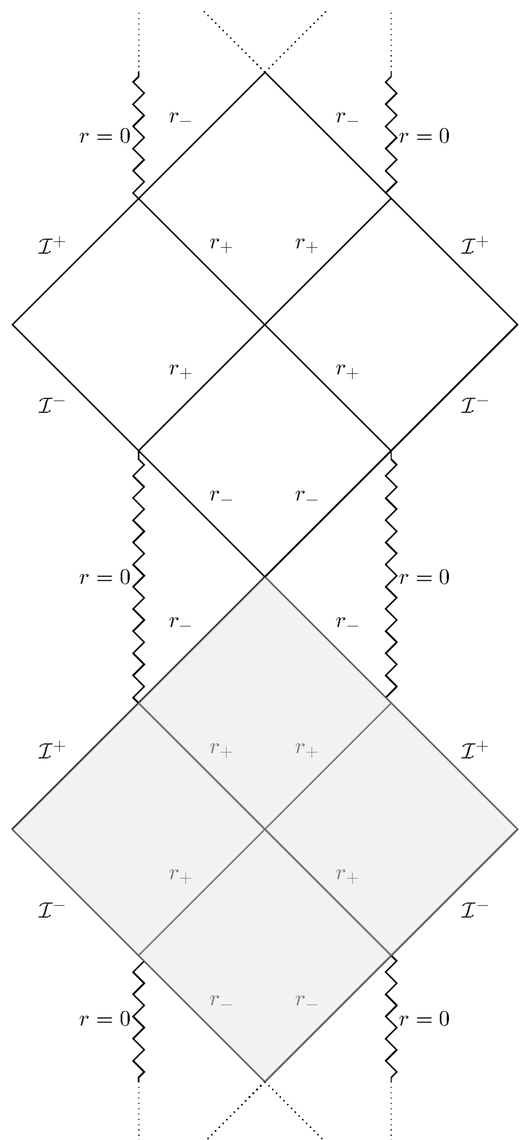


Figure 37: Penrose Diagram of the Reissner-Nordstrøm metric: regime of validity of one set of Kruskal-Szekeres coordinates.

In the region between the horizons one can then introduce the inner Kruskal coordinates (u_{K-}, v_{K-}) which extend beyond r_- (but become singular at r_+ instead). These two types of Kruskal coordinate systems can then be used alternatingly, each an infinite number of times, to pave the entire space-time.

Since thus the space-time cannot be covered by a single Kruskal coordinate patch, typically not a whole lot is gained by using Kruskal coordinates and for most purposes the simpler Eddington-Finkelstein coordinates are actually more convenient. However, as a matter of principle it is useful to know (and a remarkable fact in its own right) that there is a generalisation (due to Klösch and Strobl) of the Israel coordinates for the Schwarzschild metric discussed in section 27.11 that provides a global covering of the complete (infinitely periodically extended) Reissner-Nordstrøm space-time - see the reference in footnote 90 of that section.

Nevertheless, all in all, as in the case of the Kruskal diagram for the eternal fully extended Schwarzschild space-time, one should perhaps be somewhat skeptical of this intriguing and entertaining white hole - black hole structure and narrative for the extended Reissner-Nordstrøm space-time and take it with a substantial grain of salt:

1. First of all, for a collapsing star settling down to the non-extremal Reissner-Nordstrøm solution, the exotic regions beyond (i.e. before) the past white hole horizon r_+ are eliminated (as for Schwarzschild), as is the mirror region - but at first the infinite chain of white and black holes in the future remains intact.

2. Moreover, the outgoing inner horizon is known to be unstable to small perturbations via a phenomenon known as mass inflation[118], loosely speaking an infinite blue shift of infalling radiation experienced by an observer attempting to cross the horizon (anticipated by Penrose, 1968), and the inner horizon is expected to become singular in realistic situation.

3. One can repeat the story for the extremal case. In this case, the surface gravity is zero, but one can still construct Eddington-Finkelstein coordinates (as we have seen above), so that one can describe the region behind the horizon in this case. The instability of the inner horizon of a non-extremal black hole may however limit the validity of this picture and, in fact, seems to suggest that in the extremal case the outer = inner horizon may become singular.[119]

[118] E. Poisson, W. Israel, *Internal Structure of Black Holes*, Phys. Rev. D41 (1990) 1796. See also section 5.7.3 of E. Poisson, *A Relativist't Toolkit*.

[119] See D. Marolf, *The dangers of extremes*, `arXiv:1005.2999 [gr-qc]` and D. Garfinkle, *How extreme are extreme black holes?*, `arXiv:1105.2574` for a discussion of this.

32 Black Holes VI: Horizons

32.1 Introduction

So far, we have studied concretely 2 classes of exact solutions of the Einstein equations that can describe what we have called black holes, namely the Schwarzschild metric and the Reissner-Nordstrøm metric. However, this 2-parameter family of solutions to the Einstein(-Maxwell) equations is obviously very special, as the solutions are both static and spherically symmetric.

The aim of this section is to study properties of black holes in more generality, and therefore the first issue to address is what one actually means by a black hole. From the examples that we have studied, we know that the characteristic features arise from what is happening at the Schwarzschild horizon of the Schwarzschild black hole and at the outer horizon of the Reissner-Nordstrøm black hole. These examples suggest that black holes in general should be characterised and defined not in terms of what happens inside a black hole, but in terms of the properties of its "boundary" or horizon.

While the examples that we are already familiar with give us some idea, as we will recall in section 32.2, the Schwarzschild and Reissner-Nordstrøm horizons share a number of different properties and can thus also be characterised in many different ways. Therefore this does not automatically provide us with a unique candidate definition of a black hole boundary or horizon in a more general context.

Some disambiguation is provided by looking at the *Kerr metric* describing a rotating black hole (section 32.3). This metric is neither static nor spherically symmetric, but is still stationary as well as axially symmetric. We will revisit the different characterisation of the Schwarzschild horizon in this case, and at that stage essentially 3 a priori distinct characterisations of a (stationary) black hole survive:

1. Event Horizon

   This is the traditional global notion of a black hole that is meant to capture the idea that the black hole is a region of space-time that is invisible to an outside or asymptotic observer. Informally speaking, an event horizon is then the boundary of this black hole region.

   Until further notice, we will use the term "event horizon" in this way. A slightly more formal, gobal and causal, definition of the event horizon will then be given in section 32.4 (without, however, attempting to make this mathematically rigorous).

2. Killing Horizon

   This notion of a horizon relies on the existence of an asymptotically timelike Killing vector. It turns out to be a very convenient (local, geometric) characterisation of the (global, causal) event horizon of a stationary black hole.

Various aspects of the Killing horizon, and its relation with the event horizon, will be briefly discussed in sections 32.5 (rigidity theorems), 32.6 (surface gravity) and 32.7 (properties of the generating null congruences).

3. (Marginally) Trapped Surfaces

   Trapped Surfaces are closed spacelike 2-surfaces which are such that their area decreases locally along any future direction, in particular even along would-be "outgoing" lightrays. Thus this notion of trapped surfaces, due to Penrose (1965), captures the idea that in a sufficiently strong gravitational field, as in gravitational collapse, even outgoing lightrays are bent inwards.

   For stationary black holes, the event horizon (and thus the Killing horizon) can be characterised equivalently as the boundary of the region containing such trapped surfaces. In dynamical situations, however, trapped surfaces are in general dissociated from the event horizon and thus acquire an independent significance, as a means of (quasi-)locally detecting or defining black hole-like regions of space-time. These trapped surfaces and associated (quasi-local and geometric) notions of horizons and black hole boundaries will be briefly discussed in the context of a specific dynamical black hole metric, the *Vaidya metric*, in sections 32.8 and 32.9.

These trapped surfaces also play a central role in the singularity theorems of general relativity and are the prime indicators that a singularity will develop. From this perspective, it is perhaps the trapped surfaces that are fundamental, and the event horizon is only a considerate afterthought woven by a benign "cosmic censor" to hide the resulting singularity from the outside. It is therefore of interest to investigate the relation between the event horizon and various notions of black hole boundaries based on trapped surfaces, and this will be the subject of the last part of this section.

This section is unavoidably technically somewhat more advanced than other sections in this part of the notes. In particular, we will make extensive use of

- the properties of null geodesic congruences studied in sections 12.4 and 12.5;
- the properties of null hypersurfaces studied in section 17;
- Penrose diagrams and the associated notions of (conformal) infinity, in particular future null infinity $\mathcal{I}^+$, introduced in section 28.

32.2 The many Facets of the Schwarzschild Radius

Let us start by reconsidering the various features of the future horizon of the Schwarzschild metric, which we write either in the standard Schwarzschild coordinates or in advanced Eddington-Finkelstein coordinates (which extend across the future horizon) as

$$\begin{aligned} ds^2 &= -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \\ &= -f(r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \ \ , \end{aligned} \tag{32.1}$$

with

$$f(r) = 1 - \frac{2m}{r} = 1 - \frac{r_s}{r} \ \ . \tag{32.2}$$

In this metric, there are several apparently quite different things happening at $r = r_s$, and therefore different ways of characterising $r = r_s$. Everything that is said below is also valid for the outer horizon r_+ of a Reissner-Nordstrøm black hole with

$$f(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2} = \frac{(r - r_+)(r - r_-)}{r^2} \ \ . \tag{32.3}$$

We begin with the description in terms of the coordinate components of the metric:

1. At $r = r_s$ we have

$$g_{tt}(r_s) = 0 \quad \Leftrightarrow \quad g_{vv}(r_s) = 0 \ \ . \tag{32.4}$$

2. Likewise, we have (in both coordinate systems)

$$g^{rr}(r_s) = 0 \ \ . \tag{32.5}$$

In order to make the 1st characterisation somewhat less coordinate dependent, we can assign some additional physical significance to the coordinate t or the vector field $\xi = \partial_t$. For example, we can focus on the fact that static observers, i.e. those that stay at fixed spatial Schwarzschild coordinates (r, θ, ϕ), have worldlines with 4-velocity $u^\alpha \sim \xi^\alpha$.

3. We can then interpret the fact that $g_{tt}(r_s) = 0$ as the statement that such static observers can only exist for $r > r_s$. In this sense

$$\mathcal{S} \equiv \{\xi = \partial_t \text{ null}\} = \{r = r_s\} \quad \text{is a } \textit{static limit surface} \ \ . \tag{32.6}$$

4. Related to this is the fact that for a static observer

$$r = r_s \quad \text{is an } \textit{infinite redshift surface} \ \ . \tag{32.7}$$

Alternatively, we can focus on the fact that $\xi = \partial_t = \partial_v$ is an asymptotically timelike Killing vector of the metric:

5. The asymptotically timelike Killing vector ξ becomes lightlike at $r = r_s$,

$$r = r_s \quad \Rightarrow \quad g_{\alpha\beta}\xi^\alpha\xi^\beta = 0 \ \ , \tag{32.8}$$

and spacelike for $r < r_s$,

$$r < r_s \quad \Rightarrow \quad g_{\alpha\beta}\xi^\alpha\xi^\beta > 0 \ \ . \tag{32.9}$$

6. Moreover, we have that this locus is itself actually a null surface,

$$\mathcal{K} \equiv \{x : (g_{\alpha\beta}\xi^\alpha\xi^\beta)(x) = 0\} = \{r = r_s\} \quad \text{is a null hypersurface} \quad , \tag{32.10}$$

   Because the length of a Killing vector does not change in the direction of a Killing vector (see (9.61) or (9.62)), ξ is tangent to $\mathcal{K}$ (and therefore also normal to $\mathcal{K}$, cf. the discussion of null hypersurfaces in section 17.1).

We now turn to the 2nd characterisation $g^{rr}(r_s) = 0$:

7. We can reformulate this somewhat more invariantly as the statement that the normal vector to the hypersurfaces of constant r, $N_\alpha \sim \partial_\alpha r$, which is asymptotically spacelike, becomes null at $r = r_s$,

$$r = r_s \quad \Rightarrow \quad g^{\alpha\beta}\partial_\alpha r \partial_\beta r = 0 \tag{32.11}$$

   and actually becomes timelike for $r < r_s$,

$$r < r_s \quad \Rightarrow \quad g^{\alpha\beta}\partial_\alpha r \partial_\beta r < 0 \tag{32.12}$$

8. In particular, the hypersurface $r = r_s$ is null, and for $r < r_s$ one can only move through the spacelike hypersurfaces of constant r in the direction of decreasing r,

$$r < r_s \quad \Rightarrow \quad \dot{r} < 0 \quad \text{along future-directed timelike paths} \quad . \tag{32.13}$$

As we have seen, a crucial role in our analysis of the Schwarzschild black hole was played by analysing and understanding the behaviour of radial lightrays and lightcones, i.e. the causal structure of the space-time. Let us reconsider $r = r_s$ from this point of view:

9. In Eddington-Finkelstein coordinates, by construction lines of constant v (and constant (θ, ϕ)) describe ingoing radial lightrays. These are the trivial solutions of

$$-f(r)dv^2 + 2dvdr = 0 \quad . \tag{32.14}$$

   The other solution is described by

$$2(dr/dv) = f(r) \quad (\text{“outgoing” lightrays}) \quad , \tag{32.15}$$

   and therefore for $r < r_s$ these would-be outgoing lightrays actually also move to smaller values of r,

$$r < r_s \quad \Rightarrow \quad dr/dv < 0 \quad (\text{“outgoing” lightrays are not outgoing}) \quad . \tag{32.16}$$

   Thus $r = r_s$ is where the lightcones “tilt over”, and can only be crossed in the direction from $r > r_s$ to $r < r_s$.

We had already noted in section 27.5 that this can be phrased in a somewhat more geometric and invariant way in terms of the *expansions* θ_n and θ_ℓ of the in- and outgoing radiall null congruences defined by the null vector fields (27.115)

$$n = -\partial_r \quad , \quad \ell = \partial_v + \tfrac{1}{2} f(r) \partial_r \quad . \tag{32.17}$$

For these expansions we found

$$\theta_n = -\frac{2}{r} < 0 \quad \forall\, r \quad , \tag{32.18}$$

and

$$\theta_\ell = \frac{r-2m}{r^2} \quad \Rightarrow \quad \begin{cases} \theta_\ell > 0 & r > r_s \\ \theta_\ell = 0 & r = r_s \\ \theta_\ell < 0 & r < r_s \end{cases} \tag{32.19}$$

One says that for $r < r_s$ (resp. $r = r_s$), the spheres $S_{r,v}$ of constant r and v, with $\theta_n < 0$ and $\theta_\ell < 0$ (resp. $\theta_\ell = 0$) are *trapped* (resp. *marginally trapped*).

10. Thus, another way of characterising $r = r_s$ is as the statement that

$$r = r_s \quad \Rightarrow \quad \theta_n < 0 \quad , \quad \theta_\ell = 0 \quad \Rightarrow \quad S_{r_s,v} \ \textit{marginally trapped} \tag{32.20}$$

11. We can also rephrase this as the statement that the null surface $r = r_s$ is foliated by such marginally trapped spheres,

$$\mathcal{T} \equiv \cup_v S_{r_s,v} = \{r = r_s\} \quad . \tag{32.21}$$

Finally we can also turn to the "global, causal" characterisation of a black hole in terms of a (future) *event horizon*, defined here (for the time being) informally as the boundary of the region from which signals (lightrays) can be sent to an asymptotic observer (the more formal definition of the event horizon will be discussed in section 32.4):

12. Due to the time-independence of the Schwarzschild metric, this global property follows from the local behaviour of the lightcones established above, and therefore we have

$$\{r = r_s\} = \mathcal{H}^+ \quad \text{is a } \textit{(future) event horizon} \quad . \tag{32.22}$$

For the static and spherically symmetric Schwarzschild metric, all these 12 characterisations of $r = r_s$ (and perhaps some others I have overlooked or deliberately ignored) are equivalent. Some of these appear to be more closely related than others to what one might mean by a "horizon" or a "black hole", but clearly the Schwarzschild black hole alone is not enough to decide which of these criteria are pertinent or equivalent in more generality. As soon as one moves away either from the static situation or from the spherically symmetric situation, one finds that these different characterisation do no longer necesarily coincide (or are not even applicable) and even when applicable may capture different phenomena. Thus in order to decide which of these properties are the most useful or appropriate to capture at least some aspect of the "black hole-ness" of an object, we will now look at another example.

32.3 Kerr Metric: Ergosphere vs Killing Horizon and Event Horizon

This example is the *Kerr metric*, already briefly mentioned in section 30.1. In contrast to the Schwarzschild metric, it is neither static nor spherically symmetric, but it is still stationary and axially symmetric.

The Kerr metric is perhaps the single most important exact solution of the Einstein equations for astrophysical purposes, and there are a lot of things that should be said about the Kerr metric (and this is done in most respectable textbooks of general relativity), but here I will focus on those that are relevant for the (horizon) issue at hand.

In Boyer-Lindquist coordinates (t, r, θ, ϕ), the Kerr metric is (30.3),(30.5)

$$\begin{aligned} ds^2 &= -\left(1 - \frac{2mr}{\rho^2}\right) dt^2 - \frac{4mra\sin^2\theta}{\rho^2} dt\, d\phi + \frac{\Sigma}{\rho^2}\sin^2\theta d\phi^2 + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 \\ &= -\frac{\rho^2\Delta}{\Sigma}dt^2 + \frac{\Sigma}{\rho^2}\sin^2\theta(d\phi - \omega dt)^2 + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 \quad , \end{aligned} \tag{32.23}$$

where $\Delta, \rho, \Sigma, \omega$ are the (unfortunately somewhat complicated) functions

$$\begin{aligned} \Delta(r) &= r^2 - 2mr + a^2 \\ \rho(r,\theta)^2 &= r^2 + a^2\cos^2\theta \\ \Sigma(r,\theta) &= (r^2 + a^2)^2 - \Delta(r)a^2\sin^2\theta \\ \omega(r,\theta) &= -g_{t\phi}/g_{\phi\phi} = 2mar/\Sigma(r,\theta) \quad . \end{aligned} \tag{32.24}$$

For later use we note that the coefficients of the metric satisfy the simple (but rather unobvious) relation

$$g_{t\phi}^2 - g_{tt}g_{\phi\phi} = \Delta(r)\sin^2\theta \quad . \tag{32.25}$$

This also implies that the volume element $\sqrt{g}$ has the surprisingly simple Schwarzschild-like form

$$\sqrt{g} = \rho^2 \sin\theta \tag{32.26}$$

(but we will not make use of this result below).

As mentioned in section 30.1, this metric describes the gravitational field outside a rotating star or that of a rotating a black hole, with mass parameter m and angular momentum parameter a (and with the condition $|a| \leq m$, analogous to the condition $|q| \leq m$ for the Reissner-Nordstrøm metric excluding naked singularities - we will only look at the non-extremal case $m^2 > a^2$ in the following).

This metric is stationary, with time-translation Killing vector $\xi = \partial_t$, and axially symmetric, with the rotational Killing vector $\eta = \partial_\phi$, and the Boyer-Lindquist coordinates are evidently adapted to these two commuting symmetries. The most general Killing vector of the Kerr metric is thus of the form

$$K = a\xi + b\eta \tag{32.27}$$

with constant a, b. Because asymptotically the norm of η is proportional to r^2, while the norm of ξ is asymptotic to -1, the unique asymptotically timelike Killing vector of the Kerr metric is (up to a constant rescaling) the time-translation Killing vector $\xi = \partial_t$.

Since this metric is stationary, with Killing vector ξ, it makes sense to ask if or where this (asymptotically timelike) Killing vector becomes null. Likewise, because of the existence of a privileged (adapted) time-coordinate, there is a preferred class of observers, *static observers*, which remain at fixed values of the spatial coordinates (r, θ, ϕ), with 4-velocity

$$u^\alpha \sim \xi^\alpha \tag{32.28}$$

and it is legitimate ask if there is a static limit or infinite redshift surface for such observers. Both of these questions amount to determining the zeros of

$$g_{\alpha\beta}\xi^\alpha\xi^\beta = g_{tt} \quad , \tag{32.29}$$

and we will address this below.

Since the metric also has the axial Killing vevtor $\eta = \partial_\phi$, there is also a more general class of privileged observers, called *stationary observers*, who remain at fixed values of (r, θ) but rotate in the ϕ-direction with constant angular velocity Ω, so that

$$u^\alpha \sim \xi^\alpha_\Omega = \xi^\alpha + \Omega\eta^\alpha \quad , \tag{32.30}$$

and one can (and we will) also inquire about the existence of a corresponding stationary limit surface. Note that "constant angular velocity" here means "constant for an observer at constant (r, θ)", i.e. $\Omega = \Omega(r, \theta)$.

We will first consider $\xi = \partial_t$ and static observers. The question is thus if or where g_{tt} is zero. From the explicit expression for the metric given above one finds that

$$g_{tt}(r, \theta) = 0 \quad \Leftrightarrow \quad \rho^2 - 2mr = r^2 + a^2\cos^2\theta - 2mr = 0 \quad , \tag{32.31}$$

with solution (the rationale for the notation $r_{sl}(\theta)$ will become apparent below)

$$r = r_{sl}(\theta) = m + \sqrt{m^2 - a^2\cos^2\theta} \quad . \tag{32.32}$$

Thus at the poles $\theta = 0, \pi$ (on the axis of rotation) one has $r_{sl} = m + \sqrt{m^2 - a^2}$, and on the equatorial plane $\theta = \pi/2$ one has $r_{sl} = 2m$.

This surface

$$\mathcal{S} = \{r = r_{sl}(\theta)\} \tag{32.33}$$

has the following properties:

- $r = r_{sl}(\theta)$ defines the static limit surface for static observers (hence the notation r_{sl}), i.e. no static observers can exist for $r < r_{sl}(\theta)$.

- $r = r_{sl}$ also defines a surface of infinite redshift for static observers.
- The asymptotically timelike Killing vector $\xi = \partial_t$ becomes null at $r = r_{sl}(\theta)$.
- $r_{sl}(\theta)$ also defines what is commonly also called the *ergosphere*, and the region between the ergosphere and the event horizon, which we will pinpoint below, is then known as the *ergoregion*. This name arises from the fact (known as the *Penrose process*), that (some of) the rotational energy of a rotating black hole can be extracted from the ergoregion of a black hole (and ergon = work in ancient Greek).

For the Schwarzschild metric, $r_{sl}(\theta) \to 2m = r_s$ reduces to the Schwarzschild radius. Thus it is our first candidate for the horizon of the Kerr black hole. However, whatever other interesting properties the static limit surface or ergosphere $\mathcal{S}$ may have, such an interpretation is not warranted here:

1. For example, even though no static observers can exist for $r < r_{sl}(\theta)$, this does not by itself imply that one cannot escape from that region, and it is also not true. Indeed, while static observers cannot exist inside the ergosphere (static limit surface) $\mathcal{S}$, *stationary observers* with $u^\alpha \sim \xi^\alpha + \Omega\eta^\alpha$ can (for some range of $r < r_{sl}(\theta)$) provided that they are willing to rotate with, i.e. in the direction of, the black hole.

   More precisely, requiring that u^α or $\xi_\Omega = \xi + \Omega\eta$ be timelike,

   $$g_{tt} + 2\Omega g_{t\phi} + \Omega^2 g_{\phi\phi} < 0 \ , \tag{32.34}$$

   leads to the condition that

   $$\Omega_-(r,\theta) < \Omega(r,\theta) < \Omega_+(r,\theta) \tag{32.35}$$

   where

   $$\Omega_\pm = \frac{-g_{t\phi} \pm \sqrt{g_{t\phi}^2 - g_{tt}g_{\phi\phi}}}{g_{\phi\phi}} = \omega \pm \frac{\sqrt{\Delta}\rho^2}{\Sigma \sin\theta} \tag{32.36}$$

   are the two roots of the polynomial in (32.34). In the 2nd step I used the definition of ω in (32.24) and the identity (32.25).

   Now, on the ergosphere (static limit surface) $\mathcal{S}$ one has, by definition, $g_{tt} = 0$, so that $\Omega_- = 0$ there (note that $g_{t\phi}$ is negative), while Ω_- is negative (positive) outside (inside) the ergosphere,

   $$\Omega_- \quad \begin{cases} < 0 & \text{for} \quad r > r_{sl} \\ = 0 & \text{for} \quad r = r_{sl} \\ > 0 & \text{for} \quad r < r_{sl} \end{cases} \tag{32.37}$$

This is to be interpreted as the statement that outside the ergosphere stationary observers can exist that can rotate either with or against the sense of rotation of the black hole, while on and inside the ergosphere a stationary observer has no choice but to rotate with (i.e. to be dragged along by) the black hole. This condition $\Omega_- > 0$ continues to hold inside the ergosphere even when one adds momentum in the r (and/or θ) direction, of either sign, and such observers can then leave the ergopshere again. Thus the ergopshere is not like a horizon or 1-way membrane.

From the above explicit expression for the $\Omega_\pm$ we see that something special happens not only when $g_{tt} = 0$ (this we just discussed) but also when or where $\Delta(r) = 0$. We will come back to this below.

2. Another way of stating that the ergosphere (static limit surface) $\mathcal{S}$ is not very horizon-like is as the fact that $\mathcal{S}$ is a timelike surface, i.e. it has a spacelike normal (away from the axis of rotation). This can be seen from the fact that a (non-normalised) vector normal to

$$S(r,\theta) = r - r_{sl}(\theta) = 0 \tag{32.38}$$

will be

$$N_\alpha = \partial_\alpha S : \quad N_\alpha = (0, 1, -dr_{sl}/d\theta, 0) \ \ , \tag{32.39}$$

with norm

$$N_\alpha N^\alpha = g^{rr} + g^{\theta\theta}(dr_{sl}(\theta)/d\theta)^2 \ \ . \tag{32.40}$$

With

$$g^{rr} = \frac{\Delta}{\rho^2} \quad , \quad g^{\theta\theta} = \frac{1}{\rho^2} \tag{32.41}$$

this evaluates on $r = r_{sl}(\theta)$ to

$$N_\alpha N^\alpha = \frac{1}{2mr_{sl}} \frac{m^2 a^2 \sin^2\theta}{m^2 - a^2\cos^2\theta} \geq 0 \ \ , \tag{32.42}$$

with $N^\alpha N_\alpha = 0$ only at the poles. Such a timelike surface can never act as a horizon or 1-way membrane, since one can cross a timelike surface or timelike worldline multiple times in both directions (otherwise it would be really hard to meet people more than once!).

Thus, even though the Killing vector ξ becomes null on the ergopshere, this does not imply all by itself that the surface on which ξ becomes null is itself a null surface, even though this is what happened in the Schwarzschild case (we will see in section 32.5 that in general in the static case the former implies the latter).

Looking back at the list in section 32.2, we see that the (more or less equivalent) properties (1) and (3)-(5), as applied to $\xi = \partial_t$, describe the ergopshere but not a horizon (but we keep an open mind regarding condition (6) because, as mentioned above, ξ does not satisfy this condition since the ergosphere is not a null hypersurface).

Moving down in the list, we next have the (again more or less equivalent) conditions (2) and (7)-(9). For the Kerr metric, one has

$$g^{rr} = \frac{\Delta}{\rho^2} \tag{32.43}$$

and therefore

$$g^{rr}(r) = 0 \quad \Leftrightarrow \quad \Delta(r) = r^2 - 2mr + a^2 = 0 \ . \tag{32.44}$$

This has the 2 roots

$$r_\pm = m \pm \sqrt{m^2 - a^2} \ , \tag{32.45}$$

and we focus on r_+, as this is the one one encounters first. Note that

$$r_+ = m + \sqrt{m^2 - a^2} \leq m + \sqrt{m^2 - a^2 \cos^2\theta} = r_{sl}(\theta) \tag{32.46}$$

with equality only at the poles $\theta = 0, \pi$, i.e. on the axis of rotation.

Thus the surface $r = r_+$ is null, at this point one (radial) leg of the lightcone is aligned with this surface, and therefore this surface can (locally) only be crossed in one direction, in the case at hand from $r > r_+$ to $r < r_+$. Thus r_+ is our candidate for a black hole horizon.

It turns out that this also agrees with the event horizon. Instead of attempting to confirm this head-on, we first make the following observations regarding additional properties of the null surface $r = r_+$:

1. First of all we observe that at $\Delta(r) = 0$ one has $\Omega_+ = \Omega_- = \omega$. Thus $r = r_+$ is also the *stationary limit surface* beyond which stationary observers do not exist. Thus stationary observers approach the angular velocity

$$\Omega_h \equiv \omega(r_+) = \frac{a}{r_+^2 + a^2} \tag{32.47}$$

   at $r = r_+$. This is interpreted as (and called) the *angular velocity of the black hole*.

2. The calculation leading to (32.36) shows that at $r = r_+$ the Killing vector

$$\xi_h \equiv \xi_{\Omega_h} = \xi + \Omega_h \eta \tag{32.48}$$

   becomes null,

$$\left(g_{\alpha\beta}\xi_h^\alpha \xi_h^\beta\right)|_{r=r_+} = 0 \ . \tag{32.49}$$

   As noted at the beginning of this section, ξ_h is not asymptotically timelike. Nevertheless, a preferred normalisation for ξ (such as $\xi^\alpha\xi_\alpha \to -1$ asymptotically) also leads to a preferred normalisation for ξ_h.

3. By the same argument as for the Schwarzschild metric in section 32.2, since the length of a Killing vector does not change in the direction of the Killing vector, ξ_h is tangent to the null hypersurface $r = r_+$. Therefore this particular linear combination of the two Killing vectors actually satisfies property (6) of section 32.2,

$$\mathcal{K} = \{\xi_h \text{ null}\} = \{r = r_+\} \text{ is null} \quad , \quad \xi_h \text{ is tangent to } \mathcal{K} \ . \tag{32.50}$$

   We will formalise this property (a null hypersurface with a normal Killing vector) in terms of a *Killing horizon* below (section 32.5).

4. Because of the lack of spherical symmetry of the Kerr metric, the determination of the expansion of outgoing null congruences orthogonal to some 2-surface (of constant t and r, say) is somewhat more involved than for the Schwarzschild metric. Therefore, this is not the ideal way to check if our horizon candidate $r = r_+$ can also be described in terms of marginally trapped surfaces, as in the characterisations (10) and (11) of section 32.2.

   A better way to do this is to make use of the fact we just established that the Killing vector ξ_h is tangent to the null surface $r = r_+$. Since ξ_h is a Killing vector, the geometry cannot change along ξ_h. Moreover, because ξ_h is normal to $r = r_+$, it provides the null generators of $\mathcal{K}$ (cf. the discussion in section 17.2). Together, these two statements imply that the null geodesic congruence generated by ξ_h on $r = r_+$ must have zero expansion,

$$\theta_{\xi_h} = 0 \ , \tag{32.51}$$

   because if it had non-zero expansion, something would change along the congruence, e.g. the cross-sectional area. The formal argument for this will be given in section 32.5 below. This also shows that

$$\mathcal{T} = \{r = r_+\} \tag{32.52}$$

   is foliated by marginally trapped surfaces, and we see that for the Kerr metric $r = r_+$ also satisfies the properties (10) and (11) of section 32.2.

Finally we can turn to property (12), i.e. we return to the question if $r = r_+$ is actually the event horizon, as informally defined so far. We have seen that "outgoing" lightrays can only be truly outgoing for $r > r_+$. In general, the future behaviour of such (momentarily outgoing) lightrays depends on the future evolution of the geometry. In the case at hand, however, because the metric is stationary, we can extrapolate this statement all the way to the future to conclude that indeed $r = r_+$ is also the (future, outer) event horizon of the Kerr metric. (In addition, as for the Reissner-Nordstrøm metric, there are inner and/or past horizons at $r = r_-$, but we are not interested in these here).

Thus, from the Kerr metric we learn that there are essentially 3 a priori logically distinct ways of characterising the event horizon of a stationary black hole, namely in terms of

1. an event horizon
2. a Killing horizon
3. (marginally) trapped surfaces

(cf. the introduction to this section), and we will now formalise these notions in turn, starting with the event horizon.

32.4 Event Horizon

As mentioned in section 32.1, the characterisation of a black hole in terms of an event horizon is meant to capture the idea that the black hole is a region of space-time that is invisible to an asymptotic observer. Thus the boundary of this black hole region is such that it causally seals off part of the space-time from the outside, while the sealed-off region itself is then regarded as the interior (black hole) region.

Since this refers to asymptotic observers and the causal structure, it is useful to recall the Penrose diagrams for the Schwarzschild and Reissner-Nordstrøm solutions. In particular, as we already know that past horizons and white holes are an artefact of considering eternal black holes, we will focus on the external asymptotically flat region and the region around the future horizon, as represented e.g. by the Penrose diagrams for the collapse of a null shell (Figure 31) or of a star (Figure 33). The essential features of these diagrams are reproduced in Figure 38 below.

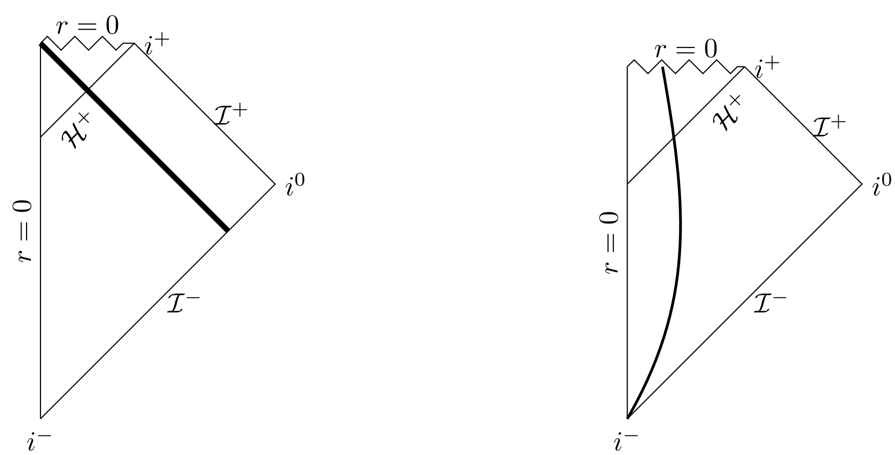


Figure 38: Penrose Diagram of the essential part of a (Schwarzschild) black hole: collapse of a null-shell on the left, collapse of a star on the right.

From these diagrams we can read off that what characterises the black hole in an asymptotically flat space-time is that it is the region of space-time from which one cannot send signals to future null infinity $\mathcal{I}^+$. Equivalently, the event horizon is the boundary of the region that can send signals to $\mathcal{I}^+$.

This is now also precisely captured by, and made precise in, the "official" definition of a future event horizon $\mathcal{H}^+$, as the boundary of the past of future null infinity $\mathcal{I}^+$,

$$\text{Future Event Horizon:} \quad \mathcal{H}^+ = \text{Future Boundary of the Causal Past of } \mathcal{I}^+ \tag{32.53}$$

This becomes a rigorous definition once all the terms appearing in it have been properly defined, but we will not attempt this here.[120] This definition is illustrated in Figure 39 which, none too surprisingly, does not differ significantly from the diagrams in Figure 38 (it does, however, deliberately remain agnostic about what happens inside the black hole, e.g. whether or not there is a singularity inside; or an inner horizon; or dragons; the definition does not address this).

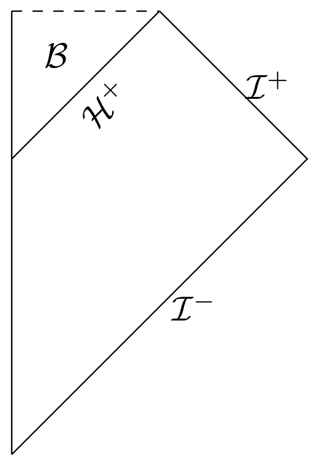


Figure 39: Definition of the event horizon and the black hole region: the future event horizon $\mathcal{H}^+$ is the (future) boundary of the past of future null infinity $\mathcal{I}^+$. The complement of the past of $\mathcal{I}^+$ is the black hole region $\mathcal{B}$, the region from which no signals can be sent to $\mathcal{I}^+$.

Here are some of the key features of this definition:

1. By its definition as a causal boundary, the event horizon $\mathcal{H}^+$ is a null hypersurface.
2. The definition relies on the existence of conformal infinity, in particular future null infinity $\mathcal{I}^+$. As such, this definition can be used in (suitably defined) asymptotically flat space-times, as well as for certain other asymptotics (e.g. asymptotically anti-de Sitter space-times). However, it cannot be used in spatially compact space-times.

[120] For a detailed treatment, see e.g. S. Hawking, G. Ellis, *The large scale structure of space-time* and sections 11 and 12 of R. Wald, *General Relativity*.

3. This definition is what is usually called *teleological* in the literature, i.e. given that the event horizon is defined as the "future boundary of the past of future null infinity", in order to define a black hole (or even in order to decide if there is a black hole at all somewhere right now) one needs to know the entire future evolution of the space-time (and then trace back lightrays from the infinite future to "today").

This definition of a black hole in terms of an event horizon has been tremendously useful, and has led to numerous and valuable insights into the nature of black holes. However, this time-honoured definition of a black hole is not completely unproblematic (the following, and other, points have all been made repeatedly in the past, in particular recently in the literature developing and advocating alternative quasi-local definitions of black hole boundaries; see the references in footnote 131 in section 32.9 below):

1. This definition of a black hole is so non-local in space that it rules out black holes in spatially compact universes.
2. It is also so non-local in time that it does not even allow astrophysicists to speak now about a supermassive black hole at the center of our galaxy.
3. Moreover, people working in numerical relativity will rightly consider an instruction to wait an infinite amount of time in order to then determine retroactively whether or not they had encountered a black hole in their simulations to be a somewhat non-constructive and non-productive procedure.

Perhaps black holes are indeed intrinsically so non-local objects that one cannot do better. However, in many ways black holes appear to be behave like reasonably local objects. It should also be kept in mind that, strictly speaking, the definition of asymptotically flat space-times and the associated construction of $\mathcal{I}^+$ and conformal infinity, were always meant to be idealisations of sufficiently distant observers in realistic space-times, say. Such idealisations are of course very common in phyics (spherical cows), but they are only useful if they actually simplify the analysis. If such idealisations give rise to their own technical problems (and there are indeed such problems[121]), then perhaps other idealised descriptions should be sought.

This suggests that the definition of a black hole in terms of an event horizon is perhaps not for all intents and purposes the best definition. For all these reasons, but also motivated by considerations involving the mechanics and thermodynamics of black hole horizons, in recent years a lot of work has gone into finding suitable definitions and

[121]See e.g. sections 1.4 and 1.5 of P. Chrusciel, *Black Holes*, `arXiv:gr-qc/0201053` for an incisive mathematical critique, and an analysis of the deficiencies of the corresponding "$\mathcal{I}^+$-based definition of a black hole".

quasi-local geometric characterisations of horizons and studying their properties. Most (if not all) of these rely in one way or another on marginally trapped surfaces and related concepts, and we will briefly return to this later on, after having discussed the Vaidya metric (sections 32.8 and 32.9).

32.5 Killing Horizons as Event Horizons of Stationary Black Holes

Looking back at the list in section 32.2, one of the characterisations of the Schwarzschild radius that a priori appears to have little to do with the most familiar properties or intuitive notions of a black hole, or with the event horizon, is property (6), that the horizon is a null surface with normal vector a Killing vector ($\xi = \partial_t$ in that case). Nevertheless, we saw that the event horizon of the Kerr black hole also has this property, albeit with respect to a different Killing vector $\xi_h = \xi + \Omega_h \eta$.

The fact that in both these examples the event horizon turned out to have this property is no coincidence. Indeed, there are so-called "rigidity theorems" which relate the global causal notion of an event horizon to the (a priori unrelated and independent) local, purely geometrical notion of a *Killing horizon*, which we can define eqivalently as

- a null hypersurface is a Killing horizon $\mathcal{K}$ of a Killing vector field ξ if $\xi|_{\mathcal{K}}$ is normal to $\mathcal{K}$;
- a Killing horizon is a null hypersurface $\mathcal{K}$ whose null generators are the integral curves of the restiction of a Killing vector field ξ to $\mathcal{K}$.

These above-mentioned rigidity theorems then state that under rather general conditions, and in a variety of circumstances, the event horizon of a *stationary* black hole must be a Killing horizon.[122]

One particular result along these lines is that in the static case the event horizon is a Killing horizon for the asymptotically timelike (and hypersurface-orthogonal) Killing vector ξ. In particular, validity of this statement requires that the hypersurface on which ξ becomes null, i.e. the static limit surface or infinite redshift surface, is itself a null surface, and thus a Killing horizon (something that, as we have seen, is not true for the Kerr metric).

[122]See R. Wald, *The Thermodynamics of Black Holes*, Living Rev. Relativity 4, (2001), 6; `http://www.livingreviews.org/lrr-2001-6`, `arXiv:gr-qc/9912119` for an overview and general discussion, S. Hawking and G. Ellis, *The large scale structure of space-time* for a detailed account of the classical results, and section 3.3.1 of P. Chrusciel, J. Lopes Costa, M. Heusler, *Stationary Black Holes: Uniqueness and Beyond*, Living Rev. Relativity 15 (2012) 7, `http://www.livingreviews.org/lrr-2012-7`, `arXiv:1205.6112 [gr-qc]` for a critical assessment of the current state of the art.

Subject to one simplifying technical assumption, this latter assertion is easy to prove. This assumption is that the static limit surface $\mathcal{S}$ is indeed a hypersurface that can be defined by $\xi^2 = 0$ (or at least as a connected component of this set). In other words, we assume that ξ^2 does not also vanish in some neighbourhood of a hypersurface. In that case, as in our general discussion of hypersurfaces in sections 15 - 17, we can characterise $\mathcal{S}$ in terms of its defining function

$$S(x) = -\xi^\alpha(x)\xi_\alpha(x) \tag{32.54}$$

as

$$\mathcal{S} = \{x :\ \xi^\alpha(x)\xi_\alpha(x) = 0\} = \{x :\ S(x) = 0\} \ . \tag{32.55}$$

We can then also choose $\partial_\alpha S$ as a non-vanishing normal to the surface.

Because the norm of a Killing vector does not change along the orbits of a Killing vector (see (9.61) or (9.62)), ξ is necessarily tangent to $\mathcal{S}$, and since ξ is null on $\mathcal{S}$, $\mathcal{S}$ cannot be a spacelike surface and therefore can be either a timelike surface (as for the Kerr metric) or a null surface (as for the Schwarzschild metric). What we want to show is that for ξ hypersurface-orthogonal the hypersurface $\mathcal{S}$ is null,

$$\xi \quad \text{hypersurface-orthogonal} \quad \Rightarrow \quad \mathcal{S} \quad \text{null} \ , \tag{32.56}$$

equivalently that the hypersurface-orthogonal ξ is actually orthogonal to $\mathcal{S}$,

$$\xi \quad \text{hypersurface-orthogonal} \quad \Rightarrow \quad \xi \ \text{normal to} \ \mathcal{S} \ . \tag{32.57}$$

Then one has a null surface with a Killing normal and therefore a Killing horizon.

In order to establish this, we start with the hypersurface-orthogonality condition, i.e. the Frobenius integrability condition (15.55)

$$\xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0 \ . \tag{32.58}$$

Since ξ is a Killing vector, $\nabla_\alpha\xi_\beta = -\nabla_\beta\xi_\alpha$ is anti-symmetric, and therefore this condition can also be written as

$$\xi_\alpha\nabla_\beta\xi_\gamma + \xi_\beta\nabla_\gamma\xi_\alpha = \xi_\gamma\nabla_\beta\xi_\alpha \ . \tag{32.59}$$

Contracting this with ξ^α, we see that on the static limit surface $\mathcal{S}$ we have

$$\xi_\beta\nabla_\gamma S = \xi_\gamma\nabla_\beta S \quad (\text{on } \mathcal{S}) \ . \tag{32.60}$$

Thus by the elementary linear algebra statement (17.31)

$$V_\alpha W_\beta = V_\beta W_\alpha \quad \Rightarrow \quad W_\alpha \sim V_\alpha \tag{32.61}$$

(provided that neither V nor W is identically zero), we can conclude that, since by our assumption $\partial_\alpha S \neq 0$ on $\mathcal{S}$, we have

$$\partial_\alpha S \neq 0 \quad \Rightarrow \quad \partial_\alpha S \sim \xi_\alpha \quad (\text{on } \mathcal{S}) \ . \tag{32.62}$$

Since ξ is null on $\mathcal{S}$, this shows that the normal vector to the surface is a null vector, and therefore the static limit surface $\mathcal{S}$ is a null surface with a normal Killing vector and therefore is a Killing horizon,

$$\mathcal{S} \text{ null} \quad \Rightarrow \quad \mathcal{S} = \mathcal{K} \ , \tag{32.63}$$

as claimed.

As the geometrically defined Killing horizon is much easier to work with than the globally defined event horizon, even in the stationary non-static case, it is common practice to base investigations of stationary black holes on the Killing horizon. In the following we will explore some of the more elementary properties of such Killing horizons.

Thus we assume that we are given a Killing vector ξ with a Killing horizon $\mathcal{K}$. Since $\mathcal{K}$ is a null surface, $\mathcal{K}$ will have all the properties of a general null hypersurface $\mathcal{N}$ described in section 17. In particular, the integral curves of ξ are the null geodesic generators of the surface, and there is a function $\kappa_\xi(x)$ on $\mathcal{K}$ (the inaffinity) such that

$$\xi^\beta \nabla_\beta \xi^\alpha = \kappa_\xi \xi^\alpha \ . \tag{32.64}$$

Special features of Killing horizons arise from the fact that these null geodesics generators are Killing vectors or orbits of the isometry group. Some properties of the inaffinity (or *surface gravity*) κ_ξ will be discussed in section 32.6 below, while the properties of the generating null congruence of a Killing horizon (and the comparison with those of a general event horizon) will be the subject of section 32.7.

All in all, Killing horizons turn out to provide a fairly satisfactory characterisation and description of stationary black holes. In particular it provides the basis of the laws of black hole machanics and black hole thermodynamics.[123] Nevertheless, this definition has some shortcomings:

1. First of all, Killing horizons are not necessarily associated with black holes. For example, the horizon $x = t$ of a Rindler observer ($\xi^1 = \xi^0$ in the notation of section 1.3, but here we use ξ to denote the Killing vector, not inertial coordinates) is a Killing horizon of the boost Killing vector (1.76)

$$\xi = x\partial_t + t\partial_x \tag{32.65}$$

Indeed,

$$\xi^\alpha \xi_\alpha = t^2 - x^2 \tag{32.66}$$

[123] See e.g. section 12.5 of R. Wald, *General Relativity*, or section 6 of P. Townsend, *Black Holes*, `arXiv:gr-qc/9707012v1`, G. Compère, *An introduction to the mechanics of black holes*, `arXiv:gr-qc/0611129`, or section 5.5 of E. Poisson, *A Relativist's Toolkit* for introductions to this subject.

is certainly null on the plane $t = x$, a normal vector to the surface $x - t = c$ is

$$N^a = \eta^{\alpha\beta}\partial_\beta(x-t) \quad \Rightarrow \quad N = \partial_x + \partial_t \ , \tag{32.67}$$

which is clearly null,

$$N^\alpha N_\alpha = 0 \tag{32.68}$$

and such that

$$\xi|_{x=t} = xN|_{x=t} \ . \tag{32.69}$$

2. This (counter-)example can be eliminated by the requirement that the spatial cross-sections of the Killing horizon be compact (so that the Killing horizon is what is known as *compactly generated*). Further conditions can be imposed to rule out other counterexamples.

3. Also, the definition of a Killing horizon requires the existence of a global asymptotically timelike Killing vector, and is thus not applicable in situations where either a Schwarzschild black hole forms from gravitational collapse, say, or where locally a black hole can be considered to be in equilibrium with its immediate surroundings but where there is some dynamics far away from the black hole.

4. Nevertheless, if black holes are not intrinsically and unavoidably very non-local objects one would expect some version of the laws of black hole mechanics to apply also to the (stationary portions) of the horizons of such objects. This was one of the motivations for developing the framework of *Isolated Horizons*.[124] These isolated horizons can be considered to be a special (null) case of definitions of horizons based on marginally trapped surfaces which I will briefly discuss later on.

32.6 Killing Horizons and Surface Gravity

As shown in section 17.2, and recalled in the previous section, the null normal vector field ξ of any null hypersurface $\mathcal{N}$ generates a null geodesic congruence; in particular one has

$$\xi^\beta \nabla_\beta \xi^\alpha = \kappa_\xi \xi^\alpha \tag{32.70}$$

for some function $\kappa_\xi(x)$ called the inaffinity. However, as also discussed in section 17.2, for a general null hypersurface $\mathcal{N}$ the function $\kappa_\xi(x)$ has no particular significance, since it can be changed (and even made to vanish) by replacing the normal vector ξ^α by $f\xi^\alpha$ for some non-vanishing function f on $\mathcal{N}$. In particular, if one chooses f such that $\ell^\alpha = f\xi^\alpha$ is affinely parametrised one has $\kappa_\ell(x) = 0$.

[124] See e.g. A. Ashtekar, B. Krishnan, *Isolated and dynamical horizons and their applications*, Living Rev. Relativity 7, (2004) 10. `http://www.livingreviews.org/lrr-2004-10`, `arXiv:gr-qc/0407042` and J. Engle, T. Liko, *Isolated horizons in classical and quantum gravity*, `arXiv:1112.4412 [gr-qc]`.

For a Killing vector field ξ, however, i.e. for $\mathcal{N} = \mathcal{K}$ a Killing horizon, we only have the freedom to rescale ξ by a constant, and if we have a preferred normalisation for ξ (as for the Schwarzschild and Kerr metrics), then κ_ξ is uniquely determined and is also known as the *surface gravity* of the Killing horizon or of the corresponding black hole. In this section, we will look at some elementary properties of the surface gravity κ_ξ of a black hole.

While we defined κ_ξ as the inaffinity (32.70) of the null geodesics generated by the Killing vector ξ on its Killing horizon $\mathcal{K}$, there are two commonly used alternative ways of defining and/or determining κ_ξ, and we start by introducing these.

1. For the 1st alternative definition, let us again assume, as in the proof of the statement (32.56) in the previous section, that the condition $\xi^\alpha \xi_\alpha = 0$ actually defines $\mathcal{K}$, i.e. that ξ is null locally only on $\mathcal{K}$ and not also in some neighbourhood of $\mathcal{K}$. Thus we can characterise $\mathcal{K}$ in terms of its defining function

$$S(x) = -\xi^\alpha(x)\xi_\alpha(x) \tag{32.71}$$

as

$$\mathcal{K} = \{x : \; S(x) = 0\} \quad , \tag{32.72}$$

and $\partial_\alpha S$ is normal to $\mathcal{K}$ and thus necessarily proportional to ξ_α,

$$\partial_\alpha S \sim \xi_\alpha \quad . \tag{32.73}$$

To determine the proportionality factor, we calculate

$$\partial_\alpha S = \partial_\alpha(-\xi^\beta \xi_\beta) = -2\xi^\beta \nabla_\alpha \xi_\beta = +2\xi^\beta \nabla_\beta \xi_\alpha = 2\kappa_\xi \xi_\alpha \quad . \tag{32.74}$$

Therefore we can alternatively and equivalently define κ_ξ by

$$\nabla_\alpha(-\xi^\beta \xi_\beta) = 2\kappa_\xi \xi_\alpha \quad . \tag{32.75}$$

This characterisation of κ_ξ is computationally convenient and also allows us to make contact with previous appearances of surface gravity in the context of the Schwarzschild or Reissner-Nordstrøm black hole.

For instance, for the Schwarzschild or Reissner-Nordstrøm metric with $\xi = \partial_v$ in Eddington-Finkelstein coordinates (focussing on the outer horizon r_+ for the latter), one has

$$S = -g_{vv} = f(r) \quad \Rightarrow \quad \partial_\alpha S = f'(r)\partial_\alpha r \quad , \tag{32.76}$$

Now

$$\xi_\alpha = g_{\alpha\beta}\xi^\beta = g_{\alpha v} \tag{32.77}$$

with $g_{vv} = -f$ and $g_{vr} = 1$, and therefore

$$\xi_\alpha|_{\mathcal{K}} = \partial_\alpha r \quad . \tag{32.78}$$

From this we can read off that

$$\kappa_\xi = \frac{1}{2}f'(r)|_{\mathcal{K}} = \frac{1}{2}f'(r)|_{r=r_s,r_+} \quad . \tag{32.79}$$

Thus this reproduces precisely the definition of surface gravity of the Schwarzschild metric first given in (27.177) in section 27.10 (and its analogue (31.123) for the Reissner-Nordstrøm metric).

2. For the 2nd alternative definition, we make use of the fact that, as the normal vector to $\mathcal{K}$, ξ is hypersurface-orthogonal and therefore satisfies the Frobenius integrability condition (15.55)

$$\xi_{[\alpha}\nabla_\beta\xi_{\gamma]} = 0 \quad . \tag{32.80}$$

   Since ξ is a Killing vector, this can also be written as

$$\xi_\gamma\nabla_\alpha\xi_\beta = -\xi_\alpha\nabla_\beta\xi_\gamma + \xi_\beta\nabla_\alpha\xi_\gamma \quad . \tag{32.81}$$

   Contracting this equation with $\nabla^\alpha\xi^\beta = -\nabla^\beta\xi^\alpha$, and using (32.70), one finds

$$\xi_\gamma\nabla_\alpha\xi_\beta\nabla^\alpha\xi^\beta = -\xi_\alpha\nabla^\alpha\xi^\beta\nabla_\beta\xi_\gamma - \xi_\beta\nabla^\beta\xi^\alpha\nabla_\alpha\xi_\gamma = -2(\kappa_\xi)^2\xi_\gamma \quad . \tag{32.82}$$

   Thus at points at which $\xi(x) \neq 0$, one can extract from this that κ_ξ can alternatively be defined as (or computed from)

$$(\kappa_\xi)^2 = -\tfrac{1}{2}(\nabla^\alpha\xi^\beta)(\nabla_\alpha\xi_\beta) \quad . \tag{32.83}$$

   By continuity this equation can then also be shown to hold at points at which $\xi(x) = 0$ (and at which then necessarily $(\nabla\xi)(x) \neq 0$ identically - cf. the argument in section 14.1).

Because κ_ξ is defined purely geometrically, one can (and should) expect $\kappa_\xi(x)$ to be constant along the isometry directions, i.e. along the null geodesic generators of $\mathcal{K}$,

$$\xi^\alpha(x)\partial_\alpha\kappa_\xi(x) = 0 \quad . \tag{32.84}$$

This is indeed true and not too difficult to prove, and we will do this below. Interestingly, in the situations where one has the rigidity theorems mentioned in section 32.5 at one's disposal one can prove a much stronger statement, namely that $\kappa_\xi(x)$ is not only constant along (the integral curves of) ξ, but actually constant all over the Killing horizon $\mathcal{K}$, but this requires more work. I will briefly come back to this at the end of this section.

I will give 3 proofs of (32.84),

1. using the characterisation (32.70) and the Lie derivative along ξ

2. using the characterisation (32.70) and the covariant derivative along ξ

3. using the characterisation (32.83) and the covariant derivative along ξ

1. The first proof is essentially a 1-line argument, and uses Lie derivatives. It relies on the fact that for a Killing vector ξ and any two vector fields X, Y one has

$$L_\xi(\nabla_X Y) = \nabla_{L_\xi X} Y + \nabla_X(L_\xi Y) \tag{32.85}$$

   (while for a non-Killing vector field there would be another term arising from the Lie derivative of the Christoffel symbols, which one could write symbolically as $(L_\xi \nabla)_X Y$). Since $L_\xi \xi = [\xi, \xi] = 0$, one has

$$0 = L_\xi(\nabla_\xi \xi) = L_\xi(\kappa_\xi \xi) = (L_\xi \kappa_\xi)\xi \quad . \tag{32.86}$$

   Thus $L_\xi \kappa = \xi^\alpha \partial_\alpha \kappa_\xi = 0$ at points where $\xi(x) \neq 0$, i.e. when one has a non-trivial orbit (and when $\xi(x) = 0$ at x, then evidently also $\xi^\alpha \partial_\alpha \kappa_\xi(x) = 0$ at that point, but this says nothing about $\kappa_\xi(x)$).

2. An alternative argument uses covariant instead of Lie derivatives, and the identity (13.3) of section 13.1 for the 2nd covariant derivative of Killing vectors, namely

$$\nabla_\gamma \nabla_\beta \xi^\alpha = R^\alpha_{\beta\gamma\delta}\xi^\delta \quad . \tag{32.87}$$

   Armed with this, we act with $\xi^\gamma \nabla_\gamma$ on the defining relation (32.70). Acting on the left-hand side we find

$$\begin{aligned} \xi^\gamma \nabla_\gamma(\xi^\beta \nabla_\beta \xi^\alpha) &= (\xi^\gamma \nabla_\gamma \xi^\beta)\nabla_\beta \xi^\alpha + \xi^\gamma \xi^\beta \nabla_\gamma \nabla_\beta \xi^\alpha \\ &= \kappa_\xi \xi^\beta \nabla_\beta \xi^\alpha + \xi^\gamma \xi^\beta R^\alpha_{\beta\gamma\delta}\xi^\delta \\ &= (\kappa_\xi)^2 \xi^\alpha \quad , \end{aligned} \tag{32.88}$$

   since the curvature term vanishes because of the anti-symmetry of the Riemann tensor. Acting on the right-hand side, we have

$$\begin{aligned} \xi^\gamma \nabla_\gamma(\kappa_\xi \xi^\alpha) &= (\xi^\gamma \partial_\gamma \kappa_\xi)\xi^\alpha + \kappa_\xi \xi^\gamma \nabla_\gamma \xi^\alpha \\ &= (\xi^\gamma \partial_\gamma \kappa_\xi)\xi^\alpha + (\kappa_\xi)^2 \xi^\alpha \quad . \end{aligned} \tag{32.89}$$

   Comparing the two, we deduce $\xi^\gamma \partial_\gamma \kappa_\xi(x) = 0$, as claimed.

3. The expression (32.83) for κ_ξ also provides one with a quick alternative proof along these lines of the constancy of κ_ξ along the orbits of ξ. As a consequence of (32.87), one has

$$\xi^\gamma \nabla_\gamma (\kappa_\xi)^2 = -(\nabla^\alpha \xi^\beta)\xi^\gamma(\nabla_\gamma \nabla_\alpha \xi_\beta) = -(\nabla^\alpha \xi^\beta) R_{\beta\alpha\gamma\delta}\xi^\gamma \xi^\delta = 0 \quad . \tag{32.90}$$

As mentioned before, it is also possible to show, with some additional hypotheses (most importantly the so-called *dominant energy condition*, cf. section 22.1), that κ_ξ is also constant along the other (spatial) directions of the horizon. In terms of the adapted coordinates of section 17.3 this is the statement that

$$E_k^\gamma \nabla_\gamma (\kappa_\xi)^2 = -(\nabla^\alpha \xi^\beta) R_{\beta\alpha\gamma\delta} E_k^\gamma \xi^\delta = 0 \quad . \tag{32.91}$$

However, the standard proofs of this fact all require some non-trivial or at least non-obvious gymnastics.[125]

32.7 Killing Horizons, Event Horizons and their Null Congruences

We now return to the generating null geodesics of a Killing horizon or of an event horizon, introduced for a general null hypersurface in section 17.2, and study them from the point of view of null geodesic congruences. In section 12.4 we had studied (affinely parametrised) null geodesic congruences ℓ^α and had, in particular, introduced the notion of expansion, shear and rotation of such a congruence, defined as the irreducible parts of the spatial projection $b_{\alpha\beta} = s_\alpha^\gamma s_\beta^\delta B_{\gamma\delta}$ (12.96) of $B_{\alpha\beta} = \nabla_\beta \ell_\alpha$ (12.99),

$$\begin{aligned} b_{\alpha\beta} &= \tfrac{1}{2}\theta_\ell s_{\alpha\beta} + \tfrac{1}{2}(b_{\alpha\beta} + b_{\beta\alpha} - \theta_\ell s_{\alpha\beta}) + \tfrac{1}{2}(b_{\alpha\beta} - b_{\beta\alpha}) \\ &= \tfrac{1}{2}\theta_\ell s_{\alpha\beta} + \sigma_{\alpha\beta} + \omega_{\alpha\beta} \quad . \end{aligned} \tag{32.92}$$

In particular, this led to the null Raychaudhuri equation (12.107),

$$\frac{d}{d\tau}\theta_\ell = -R_{\alpha\beta}\ell^\alpha \ell^\beta - \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta} \quad , \tag{32.93}$$

describing the evolution of the expansion θ_ℓ along the congruence generated by ℓ. In section 12.5, we had then subsequently extended this to non-affinely parametrised null congruences, with the result that there is just one additional term involving the inaffinity of the congruence (12.129),

$$L_\ell \theta_\ell = \frac{d}{d\tau}\theta_\ell = \kappa_\ell \theta_\ell - R_{\alpha\beta}\ell^\alpha \ell^\beta - \tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} + \omega^{\alpha\beta}\omega_{\alpha\beta} \quad . \tag{32.94}$$

We will now see that these results simplify drastically when restricted and specialised to a Killing horizon and the null congruence generating that Killing horizon:

- Because $\ell^\alpha = \xi^\alpha$ is hypersurface-orthogonal, by (12.119) the rotation vanishes.

$$\xi^\alpha \quad \text{hypersurface-orthogonal} \quad \Rightarrow \quad \omega_{\alpha\beta} = 0 \quad \text{on } \mathcal{K} \quad . \tag{32.95}$$

[125] See e.g. section 12.5 of R. Wald, *General Relativity*, or section 6 of P. Townsend, *Black Holes*, `arXiv:gr-qc/9707012v1`; for a discussion and a different argument see also I. Racz, R. Wald, *Global Extensions of Spacetimes Describing Asymptotic Final States of Black Holes*, `arXiv:gr-qc/9507055`.

- Because ξ^α is a Killing vector,

$$\nabla_\alpha \xi_\beta + \nabla_\beta \xi_\alpha = 0 \quad , \tag{32.96}$$

the symmetric part of $B_{\alpha\beta} = \nabla_\beta \xi_\alpha$ vanishes, and therefore also its spatial projection is zero, implying that the shear and expansion of this null congruence are zero,

$$\xi^\alpha \quad \text{Killing vector field} \quad \Rightarrow \quad \sigma_{\alpha\beta} = 0 \quad , \quad \theta_\xi = 0 \quad \text{on} \ \mathcal{K} \ . \tag{32.97}$$

- Thus the null congruence generating a Killing horizon is irrotational, shear-free and has zero expansion, and at this stage the Raychaudhuri equation collapses to

$$L_\xi \theta_\xi = -R_{\alpha\beta}\xi^\alpha \xi^\beta \tag{32.98}$$

- Since the expansion is zero on $\mathcal{K}$, $\theta_\xi = 0$, it does not vary along $\mathcal{K}$, and therefore also

$$L_\xi \theta_\xi = 0 \quad \text{on} \ \mathcal{K} \ . \tag{32.99}$$

Therefore we have

$$R_{\alpha\beta}\xi^\alpha \xi^\beta = 0 \quad \text{on} \ \mathcal{K} \ . \tag{32.100}$$

- If the Einstein equations are satisfied, this can be rephrased as the statement that

$$T_{\alpha\beta}\xi^\alpha \xi^\beta = 0 \quad \text{on} \ \mathcal{K} \ . \tag{32.101}$$

This can be interpreted as the statement that there is no flow of matter across the Killing = event horizon, evidently a necessary condition for a stationary black hole.

It is useful to contrast this with the corresponding equation for a general event horizon $\mathcal{H}^+$. This is still a null surface, and therefore has a null normal ℓ^α and the corresponding hypersurface-orthogonal generators. As a consequence, the generating null congruence of $\mathcal{H}^+$ satisfies

$$\frac{d}{d\tau}\theta_\ell = -\tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} - R_{\alpha\beta}\ell^\alpha \ell^\beta \quad . \tag{32.102}$$

Here we have chosen ℓ^α to generate affinely parametrised geodesics, as we are free to in this more general context where ℓ^α is not restricted by the Killing vector condition. Using the Einstein equations, we can also write this as

$$\frac{d}{d\tau}\theta_\ell = -\tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} - 8\pi G_N T_{\alpha\beta}\ell^\alpha \ell^\beta \quad . \tag{32.103}$$

Here the first 2 terms on the right-hand side are manifestly non-positive, and the last term will also be non-positive provided that the so-called *null energy condition* (cf. section 22.1)

$$k^\alpha k_\alpha = 0 \quad \Rightarrow \quad T_{\alpha\beta}k^\alpha k^\beta \geq 0 \tag{32.104}$$

is satisfied. Thus in that case one has

$$T_{\alpha\beta}\ell^\alpha\ell^\beta \geq 0 \quad \Rightarrow \quad \frac{d}{d\tau}\theta_\ell \leq 0 \ , \tag{32.105}$$

i.e. θ_ℓ cannot increase. As shown in section 12.4, this implies that if $\theta_\ell(\tau) < 0$ for some value of τ, then $\theta_\ell \to -\infty$ within a finite τ-interval. This is about as far as possible from the value $\theta_\ell = 0$ of the event horizon of a stationary black hole, and therefore for any event horizon that asymptotically becomes stationary one must have $\theta_\ell \geq 0$. This is a special case of a much more general result due to Penrose that the generators of an event horizon (whose definition requires the existence of a well-defined $\mathcal{I}^+$ etc.) have no future endpoints (and can therefore in particular not develop caustics with $\theta_\ell \to -\infty$),

$$\mathcal{H}^+ : \quad \theta_\ell \geq 0 \ . \tag{32.106}$$

Because θ_ℓ measures the change in the cross-sectional area of the null congruence,

$$\frac{d}{d\tau}\sqrt{s} = \theta_\ell\sqrt{s} \ , \tag{32.107}$$

we deduce that the cross-sectional area of the generating null congruence of an asymptotically stationary event horizon cannot decrease,

$$\frac{d}{d\tau}\sqrt{s} \geq 0 \ . \tag{32.108}$$

This is one of the key ingredients in Hawking's celebrated more general *Area Theorem* stating that the area of a black hole cannot decrease if the null energy condition is satisfied.

As shown in section 12.4, we can also write (32.103) as an equation (12.109) for the change in the expansion rate of the cross-sectional area $\sqrt{s}$ of the congruence, i.e. of the horizon in the case at hand, namely (using the Einstein equations and setting the rotation to zero)

$$\frac{d^2}{d\tau^2}\sqrt{s} = \left(+\tfrac{1}{2}\theta_\ell^2 - \sigma^{\alpha\beta}\sigma_{\alpha\beta} - 8\pi G_N T_{\alpha\beta}\ell^\alpha\ell^\beta\right)\sqrt{s} \ . \tag{32.109}$$

and this equation provides some insight into the behaviour of the event horizon.[126] In particular, one sees that even though (as shown above) θ_ℓ cannot increase, the rate of expansion of the horizon can increase, and will actually increase whenever the 1st term dominates over the other terms.

One seemingly counterintuitive consequence of this is that the growth rate of the horizon is largest when there is no matter and that it actually decreases when matter arrives to cross the event horizon into the black hole. In some sense this reflects, and is commonly attributed to, the global definition of an event horizon which requires one to know the

[126] See I. Booth, *Black Hole Boundaries*, `arXiv:gr-qc/0508107`, for an illuminating discussion and more details.

entire evolution of the black hole in the future in order to determine the location of the event horizon at some earlier time.

However, the above conclusion is true not just for an event horizon but more generally for the generating (and thus hypersurface-orthogonal) congruence of any null hypersurface (perhaps subject to the condition $\theta_\ell \geq 0$). It also becomes somewhat less counterintuitive when one compares it with the behaviour of radial null congruences in Minkowski space. This example was discussed in Remark 4 of section 12.4, where we observed that

- radially outgoing lightrays $\ell = \partial_v$, $v = t + r$, have expansion

$$\theta_\ell = \frac{2}{r} > 0 \quad , \tag{32.110}$$

- that θ_ℓ (of course) satisfies the flat space Raychaudhuri equation

$$\frac{d}{d\tau}\theta_\ell = \ell^\alpha \nabla_\alpha \theta_\ell = \partial_r(+2/r) = -2/r^2 = -\tfrac{1}{2}\theta_\ell^2 \quad , \tag{32.111}$$

  or

$$L_\ell \theta_\ell = -\tfrac{1}{2}\theta_\ell^2 \quad , \tag{32.112}$$

- that with cross-sectional area $\sqrt{s} \sim r^2$ one also has

$$\frac{1}{\sqrt{s}}\frac{d^2}{d\tau^2}\sqrt{s} = \frac{1}{r^2}(\partial_r)^2 r^2 = \frac{2}{r^2} = +\tfrac{1}{2}\theta_\ell^2 \quad , \tag{32.113}$$

Thus, while $\theta_\ell \sim r^{-1} \to 0$ as $r \to \infty$, indicating that the cross-sectional spheres become flatter and flatter for large r, the cross-sectional area grows like $\sqrt{s} \sim r^2$, leading to an acceleration of its growth.

A deviation of this behaviour thus signals the presence of a non-trivial curved spacetime and matter, and matter obeying the null energy condition will have an attractive focussing effect on lightrays and will therefore decrease the expansion rate of the congruence, in the case at hand that of the horizon, just as we saw above.

Thus, once one has an event horizon, its evolution behaves in a causal and predictable way (namely according to the Raychaudhuri equation). Nevertheless, the very fact that an event horizon can start forming in empty space (as e.g. in the collapse of a null shell), long before any matter has arrived, does reflect the "teleological" character of the event horizon.

32.8 Vaidya Metrics: (Marginally) Trapped Surfaces

So far, we have only explicitly considered stationary black holes. New features arise when one considers truly time-dependent dynamical black hole solutions. In general this is complicated, of course, but a tractable class of examples is provided by the so-called Vaidya metrics, already briefly mentioned in section 30.2. We will consider the ingoing Vaidya metrics, generalisations of the Schwarzschild metric in ingoing (advanced) Eddington-Finkelstein coordinates with a mass parameter $m = m(v)$ that is now allowed to depend on the retarded time coordinate v,

$$ds^2 = -f(v,r)dv^2 + 2dvdr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m(v)}{r} \quad . \tag{32.114}$$

We will make use of the following properties of these metrics (for a more detailed discussion of Vaidya metrics see sections 40 - 42):

1. These metrics are spherically symmetric, and they are written in coordinates that are adapted to this spherical symmetry and to ingoing radial lightrays, i.e. the lines of constant v (and constant angular coordinate) are ingoing lightrays, with r an affine parameter along these lightrays.

2. These metrics are solutions to the Einstein equations with a null energy-momentum tensor of the form (40.9)

$$T_{\alpha\beta} = \frac{m'(v)}{4\pi G_N r^2}\delta^v_\alpha \delta^v_\beta \quad . \tag{32.115}$$

   The null energy condition requires

$$m'(v) \geq 0 \quad , \tag{32.116}$$

   so that the mass $m(v)$ cannot decrease. These solutions can describe null dust (or "incoherent" radiation) either entering or forming a black hole.

3. A particular (but singular) example of this was the collapsing null shell of section 29.1 with mass function (29.5),

$$m(v) = m_f \Theta(v - v_0) \quad . \tag{32.117}$$

4. Another common class of examples is provided by a mass function that continuously interpolates between Minkowski space at early times, say, and the constant mass Schwarzschild metric with final mass m_f at late times,

$$m(v) = \begin{cases} 0 & v \leq v_0 \\ m_f & v \geq v_1 \end{cases} \tag{32.118}$$

   with $m'(v) > 0$ in the interval (v_0, v_1).

   (a) A concrete and analytically tractable example is provided by the linear mass function (choosing $v_0 = 0$ for notational convenience)

$$m(v) = \begin{cases} 0 & v \leq v_0 = 0 \\ (m_f/v_1)v & 0 \leq v \leq v_1 \\ m_f & v \geq v_1 \end{cases} \tag{32.119}$$

In this case, the energy momentm tensor $\sim m'(v)$ has jumps (discontinuities) at $v = v_0$ and $v = v_1$. We will look at this class of models (and variants thereof) in detail in section 42.

(b) One can of course also choose mass functions such that the metric and its 1st derivative with respect to v are continuous. A simple (and common) choice is

$$m(v) = m_f \frac{v^2}{v^2 + T^2} \quad \text{for } v \geq 0 \ , \tag{32.120}$$

with

$$m(0) = m'(0) = 0 \quad , \quad \lim_{v \to \infty} m(v) = m_f \quad , \quad m(T) = m_f/2 \ . \tag{32.121}$$

In the case of the Schwarzschild metric, the null hypersurface $r = 2m$ described the characteristic event horizon of a static black hole. It is evident that also for the general ingoing Vaidya metric something special happens at those points of space-time where

$$f(v, r) = 0 \quad \Leftrightarrow \quad r = 2m(v) \ . \tag{32.122}$$

We can also write this condition equivalently in the equally familiar form

$$r = 2m(v) \quad \Leftrightarrow \quad g^{rr} = 0 \ . \tag{32.123}$$

However, as we will discuss now, this is not the event horizon of the Vaidya black hole. Rather, depending on who one talks to this hypersurface is known

- either as the *apparent horizon* (or better apparent 3-horizon)
- or as a *marginally trapped tube* in more recent terminology,
- or as a *(future, outer) trapping horizon*
- or as an example of a *dynamical horizon* (when/where $m'(v) > 0$)

(these terms will be explained in section 32.9 below), and it is distinct from the event horizon unless $m(v) = m_0$ is constant.

In the present context it is first of all again the locus where the lightcones “tilt over”, i.e. the boundary between between the region where the so-called outgoing future-oriented lightrays are really locally outgoing in the sense that they move to larger values of r, $dr/d\tau > 0$, and the region where also the supposedly “outgoing” future-oriented lightrays move to smaller values of r, $dr/d\tau < 0$. This can be seen directly e.g. from the condition

$$-f(v, r)dv + 2dr = 0 \quad \Leftrightarrow \quad 2\frac{dr}{dv} = f(v, r) \ . \tag{32.124}$$

for outgoing (v not constant) lightrays in ingoing coordinates. Then one sees that

$$\frac{dr}{dv} = \tfrac{1}{2} f(v, r) \begin{cases} > 0 & \text{for } r > 2m(v) : \quad \text{truly outgoing} \\ < 0 & \text{for } r < 2m(v) : \quad \text{actually ingoing} \end{cases} \ . \tag{32.125}$$

As in our discussion of the analogous phenomenon for the Schwarzschild metric in Eddington-Finkelstein coordinates in section 27.5, we can rephrase this more geometrically and invariantly in terms of the expansion of null congruences. To that end we introduce the radial null vector fields

$$n = -\partial_r \quad , \quad \ell = \partial_v + \tfrac{1}{2} f(v,r)\partial_r \ , \tag{32.126}$$

with

$$n^2 = \ell^2 = 0 \quad , \quad n.\ell = -1 \ , \tag{32.127}$$

which are the obvious Vaidya counterparts of the null vector fields introduced in (27.115) for the Schwarzschild metric. Also the expansions turn out to be exactly like their Schwarzschild counterparts. In order to determine the expansions, we can consider the 2-spheres $S = S_{v,r}$ of constant r and v. The intrinsic geometry is characterised by the induced metric, in particular by the induced volume element

$$\sqrt{s} = r^2 \sin\theta \ . \tag{32.128}$$

Because of the spherical symmetry the extrinsic geometry of the 2-sphere can be completely characterised by the fractional change of the area element along ℓ and n, i.e. by the expansions

$$\theta_\ell = \frac{1}{\sqrt{s}} L_\ell \sqrt{s} \quad , \quad \theta_n = \frac{1}{\sqrt{s}} L_n \sqrt{s} \ . \tag{32.129}$$

Concretetly, using (32.128), one finds for the expansions (cf. (12.151) and (12.152))

$$\begin{aligned} \theta_\ell &= \frac{\ell^\alpha \partial_\alpha \sqrt{s}}{\sqrt{s}} = \frac{2}{r}\ell^\alpha \partial_\alpha r = \frac{2}{r}\ell^r \\ \theta_n &= \frac{n^\alpha \partial_\alpha \sqrt{s}}{\sqrt{s}} = \frac{2}{r} n^\alpha \partial_\alpha r = \frac{2}{r} n^r \ . \end{aligned} \tag{32.130}$$

These expansions are therefore a measure of the change of r (and hence the induced area) along the null directions n and ℓ. Since n is ingoing, one expects $\theta_n < 0$, and this expectation is indeed borne out in the ingoing Vaidya metric, for which one has, from (32.126), $n^r = -1 < 0$, and thus

$$\theta_n = -\frac{2}{r} < 0 \ . \tag{32.131}$$

This is indepedent of the mass function $m(v)$ and therefore, in particular, identical to the inward expansion (perhaps better: contraction) of a sphere of constant t and r in Minkowski space along an ingoing radial congruence of lightrays.

The non-triviality of the extrinsic geometry of the surface is then characterised by θ_ℓ, and in the case of the ingoing Vaidya metric we have, again from (32.126),

$$\ell = \partial_v + \tfrac{1}{2} f(v,r)\partial_r \quad \Rightarrow \quad \theta_\ell = \frac{r - 2m(v)}{r^2} \ . \tag{32.132}$$

Thus

$$\theta_\ell \quad \begin{cases} > 0 & \text{for } \; r > 2m(v) \\ = 0 & \text{for } \; r = 2m(v) \\ < 0 & \text{for } \; r < 2m(v) \end{cases} \tag{32.133}$$

Now in general for a 2-surface S with $\theta_n < 0$

$$S \;\text{ is called } \begin{cases} \text{untrapped} & \text{if } \; \theta_\ell > 0 \\ \text{marginally trapped} & \text{if } \; \theta_\ell = 0 \\ \text{trapped} & \text{if } \; \theta_\ell < 0 \end{cases} \tag{32.134}$$

and thus we can rephrase the above result as the statement that for the Vaidya metric

$$S_{v,r} \;\text{ is } \begin{cases} \text{untrapped} & \text{for } \; r > 2m(v) \\ \text{marginally trapped} & \text{for } \; r = 2m(v) \\ \text{trapped} & \text{for } \; r < 2m(v) \end{cases} \tag{32.135}$$

<u>Remarks:</u>

1. The null vector field $n = -\partial_r$ is future oriented and ingoing, and in terms of n the energy-momentum tensor of the Vaidya metric takes the characteristic ingoing form (see (40.57) and the general discussion in section 40.4)

$$T_{\alpha\beta} = \rho_{in} n_\alpha n_\beta \quad . \tag{32.136}$$

The null energy condition implies

$$T_{\alpha\beta}\ell^\alpha \ell^\beta = \rho_{in} \geq 0 \quad . \tag{32.137}$$

2. As mentioned above, r is an affine parameter along ingoing null geodesics, and thus n generates affinely parametrised geodesics,

$$n^\alpha \nabla_\alpha n^\beta = 0 \quad \Leftrightarrow \quad \kappa_n = 0 \quad . \tag{32.138}$$

For ℓ, on the other hand, one finds (see (41.8) in section 41.1, which contains a general discussion of Vaidya null geodesics)

$$\kappa_\ell = \frac{m(v)}{r^2} \quad , \tag{32.139}$$

which is again the obvious Vaidya counterpart of the Schwarzschild expression.

3. In particular, on $r = 2m(v)$ one finds

$$\kappa_\ell|_{r=2m(v)} = \frac{1}{4m(v)} \quad . \tag{32.140}$$

However, as we will see below, if (or where) $m(v)$ is not locally constant, ℓ is not tangent to $r = 2m(v)$. Thus it is not immediately obvious if this expression can have a useful interpretation as the surface gravity of the Vaidya metric.[127]

4. Introducing a time-coordinate $\tilde{t}$ by the relation

$$\tilde{t} = v - r \quad , \tag{32.141}$$

which is modelled on (and reduces to) the Kerr-Schild (or Eddington) time-coordinate $\tilde{t}$ defined by

$$v = t + r^* = \tilde{t} + r \tag{32.142}$$

for the Schwarzschild metric and introduced in (27.123),

- slices of constant $\tilde{t}$ give a foliation of the space-time by spacelike hypersurfaces
- the 2-spheres $S_{v,r}$ at constant values of (v, r) can equivalently be viewed as spheres $S_{\tilde{t},r}$ at constant values of $(\tilde{t}, r)$ and thus lying in the hypersurfaces of constant $\tilde{t}$.

It is good to keep in mind, however, that, while there is evidently a unique solution of $r = 2m(v)$ for a given v (i.e. on a slice of constant v), the solution need not be unique on a slice of constant $\tilde{t}$. For example, if $m(v) \sim v^k$, say, then substituting $v = \tilde{t} + r$ in the condition $r = 2m(v)$, for a fixed $\tilde{t}$ one obtains a polynomial equation of degree k for r. Moreover, the number of real and positive solutions to this equation may also jump as one varies $\tilde{t}$, leading to a perhaps unexpected behaviour and evolution of (marginally) trapped surfaces when viewed in a foliation by spacelike hypersurfaces.

5. It is straightforward to extend this analysis to a general spherically symmetric metric. In the *Bondi gauge*, or in so-called *radiative coordinates* (cf. the discussion in section 40.3), such a metric can be written as (40.30)

$$ds^2 = -\mathrm{e}^{2h(v,r)} f(v,r) dv^2 + 2\,\mathrm{e}^{h(v,r)} dv\, dr + r^2 d\Omega^2 \quad , \tag{32.143}$$

with

$$f(v,r) = 1 - \frac{2m(v,r)}{r} \quad , \tag{32.144}$$

where the mass function $m(v, r)$ can be invariantly characterised as the Misner-Sharp mass (24.82)

$$M_{MS}(z) \equiv m(z) = \frac{r(z)}{2} \left(1 - g^{ab}(z)\partial_a r(z)\partial_b r(z)\right) \quad . \tag{32.145}$$

[127] For some attempts to define surface gravity for non-Killing horizons, see e.g. A. Nielsen, M. Visser, *Production and decay of evolving horizons*, `arXiv:gr-qc/0510083`, A. Nielsen, J. Yoon, *Dynamical surface gravity*, `arXiv:0711.1445 [gr-qc]`, M. Pielahn, G. Kunstatter, A. Nielsen, *Dynamical Surface Gravity in Spherically Symmetric Black Hole Formation*, `arXiv:1103.0750 [gr-qc]`, B. Cropp, S. Liberati, M. Visser, *Surface gravities for non-Killing horizons*, `arXiv:1302.2383 [gr-qc]`.

Also in this case the spheres $S_{v,r}$ with

$$g^{rr} = f(v,r) = 0 \quad \Leftrightarrow \quad r = 2m(v,r) \tag{32.146}$$

are marginally trapped.

Thus in suitable coordinates spherically symmetric (marginally) trapped surfaces of spherically symmetric metrics are easy to find and identify.

32.9 Vaidya (and beyond): Marginally Trapped Tubes and Horizons

As mentioned in section 32.1, the existence of trapped surfaces is a characteristic feature of the region of the Schwarzschild black hole inside the future horizon (causal evolution is necessarily towards decreasing values of r), and of strong gravitational fields in general. Moreover, the event horizon of the Schwarzschild black hole can be equivalently characterised as the (null) hypersurface consisting of (and foliated by) the marginally trapped spheres with $r = 2m$

$$\text{Schwarzschild:} \quad \mathcal{T} = \cup_v S_{r_s,v} \quad = \quad \text{Event Horizon } \mathcal{H}^+ \tag{32.147}$$

(see (32.21) (characterisation (11)) in section 32.2).

In the present case we are thus also led to consider the union of all the marginally trapped spheres (as v varies),

$$\text{Vaidya:} \quad \mathcal{T} = \cup_v S_{r=2m(v),v} \quad . \tag{32.148}$$

So what is $\mathcal{T}$ and (how) is it related to the event horizon?

First of all, let us obtain some more information about $\mathcal{T}$ and, in passing, introduce some (actually quite a bit of) terminology:[128]

- In modern parlance a priori $\mathcal{T}$ defines what is known as a *marginally trapped tube* (MTT). This is simply any 3-surface foliated by marginally trapped surfaces (MTSs), i.e. closed surfaces with $\theta_n < 0$ and $\theta_\ell = 0$. The notation $\mathcal{T}$ was chosen to reflect this fact.
- Closed surfaces with just $\theta_\ell = 0$ and no condition on θ_n are known as marginally outer trapped surfaces (MOTSs), and correspondingly a marginally outer trapped tube (MOTT) is a 3-surface foliated by MOTSs.

[128] I. Booth, *Black Hole Boundaries*, `arXiv:gr-qc/0508107`, provides a quick overview of, and guided tour through, the zoo of quasi-local geometric definitions of horizons. I have found Table 1 of this article to be particularly helpful to familiarise myself with the terminology and acronyms, and as a reference.

- In the present case, $\mathcal{T}$ has the additional property that as one moves inwards, i.e. along n, the expansion θ_ℓ decreases, i.e. becomes negative "inside" of the MTT $\mathcal{T}$. Specifically, we have

$$(L_n\theta_\ell)|_{r=2m(v)} = -\frac{\partial}{\partial r}\frac{r-2m(v)}{r^2}|_{r=2m(v)} = -\frac{1}{r^2} < 0 \quad . \tag{32.149}$$

  This means that just inside $\mathcal{T}$ there are genuinely trapped surfaces with $\theta_\ell < 0$.

  An MTT with this property is called a *future outer trapping horizon* (FOTH) in the terminology introduced by Hayward in influential early work on trapped surfaces and associated notions of horizons.[129]

- This terminology is modelled on the future/past outer/inner horizons of the Reissner-Nordstrøm geometry. Thus a future innner trapping horizon (FITH) would have $L_n\theta_\ell > 0$ (lightrays can expand again after having crossed the innner horizon at $r = r_-$ from $r > r_-$ to $r < r_-$), and past trapping horizons are defined similarly.

- The induced metric on $\mathcal{T}$ is

$$ds^2|_{f(v,r)=0} = 4m'(v)dv^2 + (2m(v))^2 d\Omega^2 \quad . \tag{32.150}$$

  Thus, if the null energy condition $m'(v) \geq 0$ is satisfied, $\mathcal{T}$ is spacelike except when (or in regions where) $m'(v) = 0$, where it is null. This already shows that in a dynamical situation, $m'(v) \neq 0$, $\mathcal{T}$ cannot possibly be the event horizon (which is by definition a null hypersurface).

- It is also useful to explicitly construct the tangent vector field to $\mathcal{T}$ that connects the different MTSs (specifically, that connects the points with the same values of θ and ϕ as v varies), i.e. the evolution vector field of the MTSs.[130] This is the purely radial linear combination

$$\mathcal{E} = A\ell + Bn \tag{32.151}$$

  that leaves the condition $r = 2m(v)$ invariant,

$$\mathcal{E}(r-2m(v))|_{\mathcal{T}} = (A\ell^\alpha\partial_\alpha + Bn^\alpha\partial_\alpha)(r-2m(v)))\,|_{r=2m(v)} \stackrel{!}{=} 0 \quad . \tag{32.152}$$

  As we have seen above, n is not tangent to $\mathcal{T}$ (as θ_ℓ decreases along n) so we have $A \neq 0$ and we may as well choose $A = 1$. Moreover, along $\mathcal{T}$ one has $\ell = \partial_v$, and therefore explicitly this condition is

$$(\partial_v - B\partial_r)(r-2m(v)))\,|_{r=2m(v)} = -2m'(v) - B \stackrel{!}{=} 0 \quad , \tag{32.153}$$

[129] S. Hayward, *General Laws of Black-Hole Dynamics*, `arXiv:gr-qc/9303006`; see also S. Hayward, *Dynamics of black holes*, `arXiv:0810.0923 [gr-qc]` for a review of subsequent developments.

[130] See e.g. I. Booth, J. Martin, *On the proximity of black hole horizons: lessons from Vaidya*, `arXiv:1007.1642 [gr-qc]`.

so that $B = -2m'(v) < 0$ and

$$\mathcal{E} = \ell - 2m'(v)n = \partial_v + 2m'(v)\partial_r \quad . \tag{32.154}$$

The norm of this vector field is (using $\ell^2 = n^2 = 0, \ell.n = -1$ or $g_{rr} = 0, g_{vr} = 1$, and $g_{vv} = 0$ on $\mathcal{T}$)

$$\mathcal{E}^\alpha \mathcal{E}_\alpha = 4m'(v) \quad , \tag{32.155}$$

which confirms that $\mathcal{T}$ is spacelike (null) where $m'(v) > 0$ ($m'(v) = 0$).

- For a general FOTH with $\mathcal{E} = \ell + Bn$ one has

$$\mathcal{E}^\alpha \mathcal{E}_\alpha = -2B \tag{32.156}$$

  and therefore $\mathcal{T}$ is spacelike for $B < 0$, null for $B = 0$ and timelike for $B > 0$. In the null case, $\mathcal{E} = \ell$ is akin to the usual null tangent and normal of an event or Killing horizon.

- Given $\mathcal{E}$, it is of interest to look at the *expansion of* $\mathcal{T}$, i.e. at the change in the induced volume element $\sqrt{s}$ of the MTSs along $\mathcal{E}$,

$$\theta_{\mathcal{E}} = \frac{1}{\sqrt{s}} L_{\mathcal{E}} \sqrt{s}|_{\mathcal{T}} = (\theta_\ell - 2m'(v)\theta_n)|_{\mathcal{T}} \quad . \tag{32.157}$$

  Since by definition $\theta_\ell = 0$ on $\mathcal{T}$, and $\theta_n < 0$, the null energy condition $m'(v) \geq 0$ implies that

$$\theta_{\mathcal{E}} = -2m'(v)\theta_n \geq 0 \quad . \tag{32.158}$$

  Thus in this case $\mathcal{T}$ is non-contracting, i.e. either expanding or of constant area.

- More generally, one would find

$$\theta_{\mathcal{E}} = B\theta_n \quad , \tag{32.159}$$

  and thus with $\theta_n < 0$ the sign of the expansion of $\mathcal{T}$ is correlated with the signature of $\mathcal{T}$,

$$\begin{aligned} B < 0 &\quad \Rightarrow \quad \mathcal{T} \text{ spacelike and expanding} \\ B = 0 &\quad \Rightarrow \quad \mathcal{T} \text{ null and constant area} \\ B > 0 &\quad \Rightarrow \quad \mathcal{T} \text{ timelike and contracting} \quad . \end{aligned} \tag{32.160}$$

- A spacelike and expanding MTT is called a *Dynamical Horizon*, while *Isolated Horizons* (representing local equilibrium configurations of black holes) are modelled on null hypersurfaces with $\theta_\ell = 0$ (and topology $S^2 \times \mathbb{R}$). These and related quasi-local geometric notions of horizons have been intensely studied in recent years and are still an active area of research.[131]

- A timelike MTT $\mathcal{T}$ is somewhat peculiar because it is not very horizon-like, and is occasionally known as a *Timelike Membrane*. For the Vaidya metrics this cannot occur if the null energy condition $m'(v) \geq 0$ is satisfied. Surprisingly we will encounter such a timelike membrane when looking at and for trapped surfaces in the Oppenheimer-Snyder collapse geometry in section 32.12 below.

- Finally, it is or was quite common to use the term *apparent horizon* in this context, as a notion of a horizon associated with trapped surfaces (and a choice of foliation of the space-time by spacelike hypersurfaces). Because of various technical complications[132] and because of the difficulty in locating the apparent horizons even in situations where it is well-defined, the precise definition of an apparent horizon has been pretty much abandoned in favour of those given above (and will therefore not be given here).

  In practice, nowadays the term apparent (3-)horizon appears to be used as synonymous with, say, the outermost surface with $\theta_\ell = 0$ or the MTT consisting of such surfaces.

- With respect to a spherically symmetric foliation of the space-time, such as the one provided by the coordinate $\tilde{t}$ introduced above, the apparent (3-)horizon turns out to be given by the spherically symmetric MTT $\mathcal{T}$ identified above.

  For that reason, and because it is common practice, I will also frequently use the term "apparent horizon" in the following to refer to this particular MTT.

In summary, we have found that

1. the ingoing Vaidya space-time contains spherically symmetric marginally trapped spheres (MTSs) $S_{v,r}$ at $r = 2m(v)$;

2. these spherically symmetric MTSs foliate a spherically symmetric marginally trapped tube (MTT) $\mathcal{T}$;

3. this MTT $\mathcal{T}$ is

[131] See e.g. the following review articles (and references therein and thereto): A. Ashtekar, B. Krishnan, *Dynamical Horizons and their Properties*, `arXiv:gr-qc/0308033`; A. Ashtekar, B. Krishnan, *Isolated and dynamical horizons and their applications*, Living Rev. Relativity 7, (2004) 10. `http://www.livingreviews.org/lrr-2004-10`, `arXiv:gr-qc/0407042`; I. Booth, *Black Hole Boundaries*, `arXiv:gr-qc/0508107`; A. Nielsen, *Black holes and black hole thermodynamics without event horizons*, `arXiv:0809.3850 [hep-th]`; J. Engle, T. Liko, *Isolated horizons in classical and quantum gravity*, `arXiv:1112.4412 [gr-qc]`; B. Krishnan, *Quasi-local back hole horizons*, `arXiv:1303.4635 [gr-qc]`.

[132] See e.g. section 1.6 of P. Chrusciel, *Black Holes*, `arXiv:gr-qc/0201053`.

(a) null and and of constant area and consists of isolated horizon sections where $m'(v) = 0$,

(b) spacelike and expanding and consists of dynamical horizon sections where $m'(v) > 0$;

4. the MTT $\mathcal{T}$ is also a future outer trapping horizon (FOTH) in the sense of Hayward.

In this summary I have emphasised "spherical symmetry". Indeed, the MTSs and the MTT $\mathcal{T}$ identified above are not unique:

1. even in spherically symmetric space-times there can and will be non-spherically symmetric trapped surfaces and MTSs,

2. there are ways of stacking these into non-spherically symmetric MTTs,

3. and one can also study them from the point of view of non-spherically symmetric slicings of space-time into spacelike hypersurfaces.

We will return to the 1st and 2nd items in the discussion in section 32.13 below. For the 3rd, note that one simple (axially but) not spherically symmetric choice of slicing is provided by modifying (32.141) to

$$\tilde{t}_\alpha = v - r - \alpha r \cos\theta \quad , \tag{32.161}$$

where α is a constant indicating how far from spherical symmetry the constant $\tilde{t}_\alpha$ surfaces are.[133]

32.10 Vaidya Metrics: Apparent Horizon vs Event Horizon

In the previous subsection we have tentatively identified and defined various (quasi-)local geometric notions of black hole horizons or black hole boundaries based on trapped surfaces. These local geometric notions of a black hole horizon need to be distinguished from the global causal notion of a true *event horizon*, the boundary of the past of future null infinity $\mathcal{I}^+$ (see section 32.4), whose existence is usually taken to be the defining characteristic of a black hole. In the remainder of this section, we will look at various aspects of the relation between these two concepts of horizons, by way of examples and some general remarks.

Since, by its definition as a causal boundary, the event horizon is a null surface, it is already evident from the above examples (with $\mathcal{T}$ spacelike when/where $m'(v) > 0$),

[133] This slicing and corresponding MTSs and MTTs have been investigated in A. Nielsen, M. Jasiulek, B. Krishnan, E. Schnetter, *The slicing dependence of non-spherically symmetric quasi-local horizons in Vaidya Spacetimes*, `arXiv:1007.2990 [gr-qc]`.

that in general a (future, outer) trapping horizon (FOTH) or a marginally trapped tube (MTT) will not coincide with the event horizon (and by definition a dynamical horizon cannot coincide with the event horizon).

This is also easy to understand intuitively. The event horizon is much more of a global and subtle (teleological) object than, say, the apparent horizon, the spherically symmetric MTT $\mathcal{T}$, which we have been able to determine without any effort. In order to determine the event horizon, it is not enough to know if a lightray locally or instantaneously moves to larger values of r (this local information is completely captured by the expansions θ_n and θ_ℓ). In order to be able to assert that this lightray will reach infinity, i.e. $\mathcal{I}^+$, one needs to make sure that it continues to move to larger values of r in the future. As the future behaviour of the lightray depends on the future evolution of the geometry (e.g., in the present context, on the form of the mass function $m(v)$), it is clear that the location of the event horizon at a given time cannot be determined without knowing the (entire!) future evolution of that space-time.

More specifically, in the present context of the Vaidya metric, if one has an initially really outgoing lightray at some time v_i, i.e. at some $r_i > 2m(v_i)$, it is not guaranteed that this lightray will remain at $r > 2m(v)$ for all v. If it crosses the apparent horizon $r = 2m(v)$ at some later "time" v, it reaches a local maximum of r there,

$$r = 2m(v) \quad \Rightarrow \quad dr/dv = 0 \ , \tag{32.162}$$

and then (at least at first) returns to smaller values of r. In particular, what may have appeared initially to be a safe radial distance (where one can send lightrays outwards locally) can become unsafe in the future if the mass increases ($m'(v) > 0$, as we are assuming).

Evidently, then, in the time-independent Schwarzschild case with constant mass m_0 the notions of event horizon and apparent horizon agree. In particular, if one has an initially outgoing lightray at $r > 2m_0$, then it will continue to satisfy $r > 2m_0$ in the future because r is increasing. In general, however, the event horizon will not agree with the apparent horizon (and will typically lie outside the apparent horizon for a "growing" black hole).

In order to determine the future event horizon rather than the apparent horizon, one thus needs to determine the "last" outgoing lightray that can escape to infinity. The event horizon is spanned / generated by this S^2-family of lightrays (and this description again makes it manifest that the event horizon is null). Concretely, if the mass function is such that $m(v)$ tends to a finite limit $m(v) \to m_f$ as $v \to \infty$, then the event horizon is determined by solving the equation for outgoing null geodesics for $r(v)$ with future boundary condition $r(\infty) = 2m_f$,

$$\text{Event Horizon } \mathcal{H}^+ : \quad \frac{dr}{dv} = \frac{1}{2}\left(1 - \frac{2m(v)}{r}\right) \quad , \quad r(\infty) = 2m_f \ . \tag{32.163}$$

Depending on $m(v)$, this can be done either analytically (for a linear mass function see section 42) or numerically.

A typical Penrose diagram for the Vaidya space-time, here for the (prototypical) class of examples (32.118),

$$m(v) = \begin{cases} 0 & v \leq v_0 \\ m_f & v \geq v_1 \end{cases} \qquad (32.164)$$

is given in Figure 40.

As an aside: to really end up with a black hole space-time as displayed in the Figure, i.e. in order to avoid the formation of a naked singularity in this collapse, one has to impose the peculiar condition

$$\lim_{v\to 0+} (m(v)/v) > 1/16 \quad . \qquad (32.165)$$

See section 42 for a derivation of this for the linear mass dependence with $m(v) = \mu v$, leading to the requirement $16\mu > 1$ in this case, and compare with the Penrose diagrams in Figures 59 - 61.

Returning to Figure 40, note that, in particular, and as we already saw in the Penrose diagrams describing the thin null shell or Oppenheimer-Snyder collapse, an event horizon can exist even in flat regions of space-time, and starts growing from $r = 0$ in anticipation of matter falling in to form a black hole at a later time. Spherically symmetric trapped surfaces and the spherically symmetric MTT $\mathcal{T}$, on the other hand, exist only in the region $v > v_0$.

Because $\mathcal{T}$ is described by the equation $r = 2m(v)$, the above choice of mass function implies that $\mathcal{T}$ starts off at $r = 0$ at $v = v_0$, grows to $r = 2m_f$ at $v = v_1$ and agress with the Schwarzschild event horizon at $r = 2m_f$ in the Schwarzschild region $v > v_1$.

32.11 Example: Horizons in the Collapsing Thin Light Shell Geometry

It is particularly easy to describe this evolution and growth of the event horizon in the case of the collapsing spherical shell of null matter in Minkowski space discussed in section 29.1, with metric (29.3)

$$ds^2 = -f(v,r)dv^2 + 2dv\, dr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m_f}{r}\Theta(v) \quad . \qquad (32.166)$$

1. For $v > 0$, i.e. outside the shell, the metric is the Schwarzschild metric and the event horizon is simply the Schwarzschild event horizon at $r = r_s = 2m_f$.

2. Moreover, since outside the shell the geometry is the static Schwarzschild geometry, trapped surfaces exist everywhere in the region $r < 2m_f$, the hypersurface $r = 2m_f$ is foliated by spherically symmetric MTSs and thus outside the shell

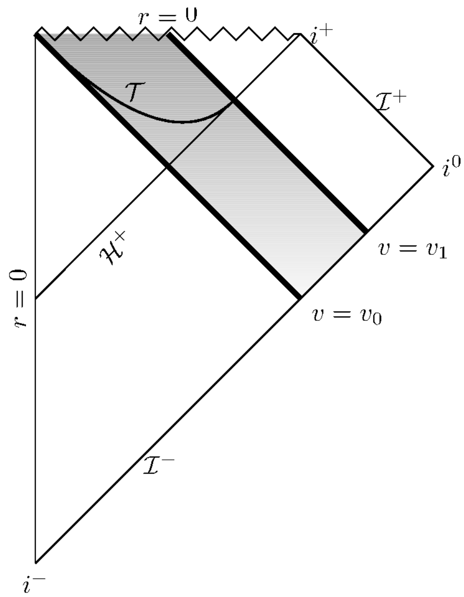


Figure 40: Event Horizon vs Apparent Horizon (Marginally Trapped Tube) $\mathcal{T}$ for the Vaidya metric, with infalling null matter in the interval $[v_0, v_1]$. For $v < v_0$, the geometry is that of Minkowski space, for $v \in [v_0, v_1]$ the geometry is described by the Vaidya metric, and for $v > v_1$ one has the Schwarzschild geometry with final mass $m = m_f$. The Event Horizon starts growing from $r = 0$ in the flat region and is described by $r = 2m_f$ in the Schwarzschild region. The Apparent Horizon (MTT) $\mathcal{T}$ is described by $r = 2m(v)$. Thus it starts off at $r = 0$ at $v = v_0$ and reaches $r = 2m_f$ at $v = v_1$, after which it agrees with the Event Horizon. In the interval $[v_0, v_1]$, $\mathcal{T}$ is spacelike whenever $m'(v) > 0$.

there is an MTT $\mathcal{T}$ (the apparent horizon with respect to a spherically symmetric foliation of space-time), and $\mathcal{T}$ agrees with the event horizon there.

3. To determine the event horizon in the interior of the shell, i.e. for $v < 0$, one needs to determine the S^2-family of outgoing radial lightrays in Minkowski space which reaches $r_s = 2m_f$ at $v = 0$, and thus connects to the exterior event horizon at the locus $v = 0$ of the shell.

   Outgoing lightrays in ingoing Minkowski coordinates (v, r) are described by

$$-dv + 2dr = 0 \quad \Leftrightarrow \quad u = t - r = v - 2r = c \text{ (constant)} \tag{32.167}$$

   i.e. by

$$r(v) = v/2 - c/2 \ . \tag{32.168}$$

   At $v = 0$ one has

$$r(v = 0) = -c/2 \overset{!}{=} r_s = 2m_f \quad \Rightarrow \quad c = -2r_s \ . \tag{32.169}$$

Therefore the event horizon is described parametrically by

$$r(v) = v/2 + r_s \ , \tag{32.170}$$

which starts growing from $r = 0$ at the time $v = -2r_s$, before the shell has arrived or crossed its Schwarzschild radius.

4. By contrast, there are no spherically symmetric (we will come back to this qualifier below) MTTs for $v \leq 0$, and thus the corresponding marginally trapped tube (apparent horizon) $\mathcal{T}$ is absent for $v < 0$.

These results are summarised in Figure 41.

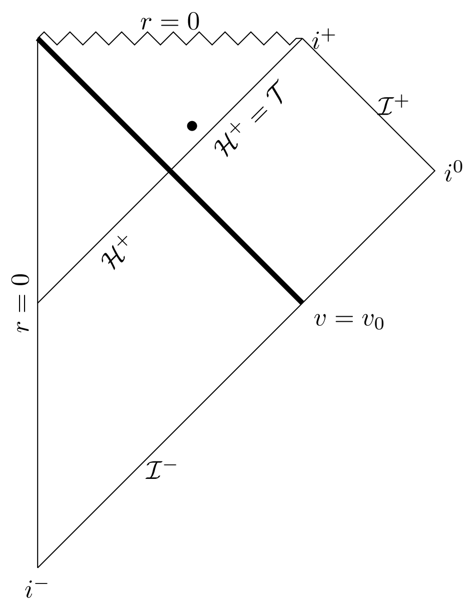


Figure 41: Event Horizon vs Apparent Horizon (Marginally Trapped Tube) $\mathcal{T}$ in the collapse of a thin null shell to a black hole. The worldline of the shell is given by the line $v = v_0$. In the region $v < v_0$ inside the shell the geometry is that of Minkowski space; the geometry outside the shell is Schwarzschild. Formation of the black hole occurs when the shell crosses the event horizon $\mathcal{H}^+$. The event horizon starts growing from $r = 0$ in the flat Minkowski region and is situated at $r = 2m_f$ outside the shell; the Apparent Horizon exists only outside the shell, and agrees with the Event Horizon there. The point indicated by a bullet represents a spherically-symmetric trapped sphere, and there are such trapped spheres for all points in the region $v > v_0, 0 < r < 2m_f$.

32.12 Example: Horizons in Oppenheimer-Snyder Collapse

The space-time geometry of a collapsing star, as described by the Oppenheimer-Snyder solution, provides another insightful illustration of the difference between apparent horizons (indicating locally the presence of trapped surfaces) and event horizons (describing globally the boundary of the region of space-time that is causally connected to infinity) in a time-dependent geometry.

The exterior geometry is given by the Schwarzschild metric and the interior geometry by a solution of the Friedmann equations describing a collapsing sphere of dust. In Painlevé-Gullstrand(-like) coordinates, the metric can be written as (see section 29.6)

$$ds^2 = \begin{cases} -d\tau^2 + (dr + \sqrt{2m/r}d\tau)^2 + r^2 d\Omega^2 & r > R(\tau) \\ -d\tau^2 + (dr - rH(\tau)d\tau)^2 + r^2 d\Omega^2 & r < R(\tau) \end{cases} \tag{32.171}$$

where $H(\tau)$ is the Hubble parameter for a dust-filled contracting universe,

$$H(\tau) = -\tfrac{2}{3}(-\tau)^{-1} < 0 \quad , \tag{32.172}$$

and the surface of the star is described by

$$r = R(\tau) = (9m/2)^{1/3}(-\tau)^{2/3} \equiv C(-\tau)^{2/3} \quad . \tag{32.173}$$

This solution describes a collapsing dust star for $\tau < 0$, collapsing to zero radius at time $\tau = 0$.

As the exterior geometry is just the Schwarzschild geometry, in the exterior region the event = apparent horizon is the null surface $r = 2m$, coming into existence at the time $\tau = \tau_f$ when the star crosses its Schwarzschild radius, i.e. at the time τ_f given by

$$R(\tau_f) = 2m \quad \Leftrightarrow \quad (9m/2)^{1/3}(-\tau_f)^{2/3} = 2m \quad \Leftrightarrow \quad \tau_f = -4m/3 \quad . \tag{32.174}$$

The interest is therefore in the formation and evolution of horizons in the interior of the star. In order to explore the causal structure of this (spherically symmetric) solution, we look at radial null rays, characterised by

$$d\tau^2 = (dr - rH(\tau)d\tau)^2 \quad . \tag{32.175}$$

This has the two branches of solutions

$$dr = (-1 + rH)d\tau \quad \text{or} \quad dr = (+1 + rH)d\tau \quad . \tag{32.176}$$

Since $H < 0$, the former describe ingoing radial null geodesics because

$$\frac{dr}{d\tau} = (-1 + rH) < 0 \qquad \text{(ingoing)} \quad , \tag{32.177}$$

while the latter, satisfying

$$\frac{dr}{d\tau} = (+1 + rH) \qquad \text{(“outgoing”)} \quad , \tag{32.178}$$

describe truly outgoing radial null geodesics only for $r < -1/H$, while these geodesics are also ingoing for $r > -1/H$. Thus there are marginally trapped spheres inside the star, centered at $r = 0$ and with radius $r = -1/H$. These define an *apparent horizon* or a *marginally trapped tube* $\mathcal{T}$ inside the star, at (recall that $\tau < 0$)

$$\text{Apparent Horion } \mathcal{T}: \quad r_{ah}(\tau) = -1/H(\tau) = -3\tau/2 \ , \tag{32.179}$$

and spheres $S_{r,\tau}$ with $r > r_{ah}(\tau)$ are trapped.

The metric induced on the apparent horizon is

$$ds^2|_{Hr=-1} = -\tfrac{3}{4}d\tau^2 + \tfrac{9}{4}\tau^2 d\Omega^2 \ , \tag{32.180}$$

which shows that the apparent horizon is a timelike hypersurface in this case. It is not clear if such an object deserves to be called a horizon at all, and the terminology *timelike membrane* has been proposed for a timelike hypersurface foliated by marginally trapped surfaces.

In order to determine the *event horizon*, we need to determine the interior outgoing lightrays that reach the surface of the star just as the surface of the star passes through its Schwarzschild radius, i.e. at the time $\tau = \tau_f$ determined in (32.174).

An outgoing lightray satisfies

$$\dot{r} = 1 + rH = 1 + 2r/3\tau \ . \tag{32.181}$$

The general solution of the homomgeneous equation is

$$\dot{r} = 2r/3\tau \quad \Rightarrow \quad r(\tau) = c_0(-\tau)^{2/3} \tag{32.182}$$

and a special solution of the inhomogeneous equation is

$$r(\tau) = 3\tau \quad \Rightarrow \quad \dot{r} = 1 + 2r/3\tau \ . \tag{32.183}$$

Thus the general solution for outgoing lightrays is

$$r(\tau) = 3\tau + c_0(-\tau)^{2/3} \ . \tag{32.184}$$

The integration constant is determined by selecting the outgoing lightray with $r(\tau_f) = 2m$,

$$r(\tau_f) = 2m \quad \Rightarrow \quad c_0 = 3C \ , \tag{32.185}$$

with C defined in (32.173), leading to the parametric equation

$$r_{eh}(\tau) = 3\left[\tau + R(\tau)\right] \tag{32.186}$$

for the event horizon.

Collecting our intermediate results, we see that the surface of the star, the apparent horizon and the event horizon are described by

$$\begin{array}{lll} \text{Surface of the Star:} & R(\tau) & = (9m/2)^{1/3}(-\tau)^{2/3} \\ \text{Apparent Horizon:} & r_{ah}(\tau) & = -3\tau/2 \\ \text{Event Horizon:} & r_{eh}(\tau) & = 3\tau + 3R(\tau) \end{array} \tag{32.187}$$

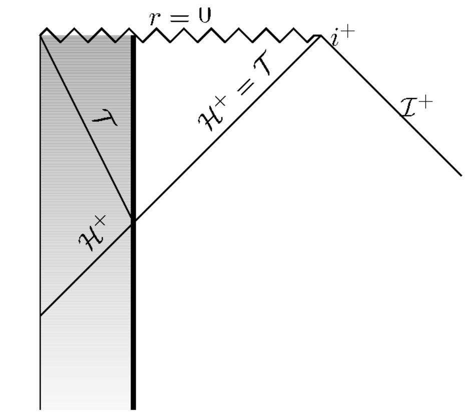


Figure 42: Event Horizon vs Apparent Horizon (Marginally Trapped Tube) $\mathcal{T}$ for the Oppenheimer-Snyder collapse geometry. The shaded region is the interior of the star, the surface of the star follows $r = R(\tau)$. Outside the star, one has the Schwarzschild geometry with event horizon = apparent horizon the null hypersurface at $r = 2m$ and the region $r < 2m$ contains spherically symmetric trapped surfaces. The Event Horizon starts growing from $r = 0$ in the interior of the star. Inside the star, there is also a timelike marginally trapped tube, a Timelike Membrane, at $r = r_{ah}(\tau)$ that starts at $r = 2m$ as the star crosses its Schwarzschild radius, and then shrinks to $r = 0$ at the time $\tau = 0$ of complete collapse. At time τ spheres inside the star centered at $r = 0$ and with radius $r > r_{ah}(\tau)$ are trapped.

This agrees with the results reported in the reference in footnote 99 in section 29.6.

In order to understand and visualise these results, we note the following:

1. At time $\tau = \tau_f = -4m/3$, for the event horizon we have

$$r_{eh}(\tau = \tau_f) = 2m \tag{32.188}$$

by construction, and also for the apparent horizon we have

$$r_{ah}(\tau = \tau_f) = -(3/2)(-4m/3) = 2m \quad , \tag{32.189}$$

as it should be.

2. The event horizon starts growing from the non-singular center of the star $r = 0$ at the time $\tau = \tau_i < 0$ determined by

$$r_{eh}(\tau_i) = 0 \quad \Rightarrow \quad \tau_i = -9m/2 < \tau_f < 0 \quad . \tag{32.190}$$

Its growth rate is

$$\dot{r}_{eh}(\tau) = 3 - 2(9m/2)^{1/3}(-\tau)^{-1/3} \geq 0 \quad \text{for} \quad \tau_i \leq \tau \leq \tau_f \tag{32.191}$$

with

$$\dot{r}_{eh}(\tau_i) = 1 \quad , \quad \dot{r}_{eh}(\tau_f) = 0 \tag{32.192}$$

so that, in particular, this matches differentiably onto the (exterior) constant radius event horizon $r_{eh} = 2m$ for $\tau > \tau_f$.

Note that, even though the event horizon grows between τ_i and τ_f, its growth rate (expansion velocity) decreases, in anticipation of the fact that no more matter will fall in after $\tau = \tau_f$. This is a vivid illustration of the "precognition" the event horizon appears to have because it is only defined once the entire evolution of the space-time is known.

3. The growth (rather "shrink") rate of the apparent horizon is

$$\dot{r}_{ah}(\tau) = -3/2 \ , \tag{32.193}$$

so the apparent horizon starts forming as the star crosses its Schwarzschild radius at $\tau = \tau_f$. Indeed for $\tau < \tau_f$ one would have $r_{ah}(\tau) > R(\tau)$, but this would be outside the star (and we determined the apparent horizon by studying the interior lightrays). It then shrinks from $r = 2m$ at $\tau = \tau_f$ to $r = 0$ at $\tau = 0$.

4. Thus for $\tau_f < \tau < 0$ the apparent horizon has two "branches", one inside the star and one outside. Inside the star, trapped spheres only occur *outside* the apparent horizon, i.e. in the region between the apparent horizon and the surface of the star. Outside the star, they of course occur in the Schwarzschild black hole region $r < r_{eh} = 2m$.

These results are summarised and schematically indicated in Figure 42.

A similar analysis can be performed for the $k = +1$ Oppenheimer-Snyder geometry, but this is conceptually identical and has no relevant new features (beyond somewhat more calculational complexity).

32.13 Concluding Comments: Trapped Regions vs Event Horizons

In the above examples, we have seen that in general dynamical situations marginally trapped tubes or trapping horizons will typically not coincide with the event horizon.

Moreover, from what we have seen so far, locally nothing particularly untoward or dangerous seems to be happpening in the region between the two, the danger apparently revealing itself only through the future evolution of the space-time.

However, in the above we have restricted attention to spherically symmetric MTSs and MTTs (the apparent horizon), and I had mentioned that even in spherical symmetry there can and will be non-spherically symmetric MTSs and MTTs. This non-uniqueness of MTTs, and the question what happens in the region between the apparent horizon and event horizon provide motivations to consider the entire region of space-time containing trapped surfaces, i.e. the *Trapped Region* $\mathbb{T}$ defined to be the set of space-time points which lie on at least one trapped surface, spherically symmetric or not. The boundary

$$\mathcal{B} = \partial\mathbb{T} \tag{32.194}$$

of this trapped region, the *Trapping Boundary*, is then a natural candidate for the black hole boundary, independent of any choices, and automatically inheriting all the symmetries of the space-time. It turns out to be surprisingly difficult and delicate, however, to determine this region precisely, even in simple examples such as the ones we have discussed here, e.g. the Vaidya metrics, and I will close this section with some comments on this subject.

1. First of all, instead of considering trapped surfaces one can also consider outer trapped surfaces ($\theta_\ell < 0$, no condition on θ_n), and thus the region $\mathbb{T}_o$ covered by outer trapped surfaces, and its boundary

$$\mathcal{B}_o = \partial\mathbb{T}_o \ \ . \tag{32.195}$$

   It was conjectured by Eardley that this 3-surface actually coincides with the event horizon, and this conjecture was established by Ben-Dov in the case of Vaiyda space-times with $m'(v) \geq 0$ and with finite total mass, i.e. with $m(v)$ bounded from above.[134]

2. Thus, if or when Eardley's conjecture holds, this seems to provide the desired almost local characterisation of the event horizon. However, this is somewhat misleading because the outer trapped spacelike surfaces that are required extend far into the future and in this way manage to feed back the information about the future evolution of the space-time into the location of the boundary 3-surface at an earlier time. This is also referred to as the *clairvoyant* property of (outer) trapped surfaces.[135] Thus in spite of these results there appears to be no good local in time characterisation of the event horizon.

3. Ben-Dov also showed that the restriction to genuinely trapped surfaces with $\theta_n < 0$ as well is not enough to fill out the space between the apparent and event horizons. This issue has been further analysed by Bengtsson and Senovilla who showed that genuinely trapped surfaces can in principle extend into parts of the flat region and investigated how far such genuinely trapped surfaces can extend into the

[134] D. Eardley, *Black Hole Boundary Conditions and Coordinate Conditions*, `arXiv:gr-qc/9703027`; I. Ben-Dov, *Outer Trapped Surfaces in Vaidya Spacetimes*, `arXiv:gr-qc/0611057`.
[135] See e.g. J. Senovilla, *Trapped Surfaces*, `arXiv:1107.1344 [gr-qc]`.

intermediate region for Vaidya space-times.[136] However, it appears that at present the location of $\mathcal{B}$ has been significantly constrained but has not yet been pinned down precisely.

4. A particularly illuminating illustration of Ben-Dov's proof of Eardley's conjecture and the work on trapped surfaces in Vaidya space-times is provided by the collapsing thin null shell geometry, Figure 43.

   As described before, point A lies on (and represents) a spherically symmetric trapped surface. Points B and C both lie in the flat Minkowski region of space-time, inside the event horizon. There can be no (closed) trapped surfaces lying entirely in the flat part of space-time ($v < v_0$). In particular, B, C do not lie on any spherically symmetric trapped surface. However,

   (a) there exist trapped surfaces ($\theta_\ell < 0, \theta_n < 0$) that can extend into the flat region, and B represents a point that lies on such a (necessarily not spherically symmetric) trapped surface,
   (b) any point in the region inside the event horizon can be shown to lie on some outer trapped surface ($\theta_\ell < 0$, no condition on θ_n),
   (c) for points sufficiently close to the event horizon and sufficiently far from the shell (point C) there are no truly trapped surfaces ($\theta_\ell < 0, \theta_n < 0$) through that point.

   It is intriguing (and again a reflection of the non-local character of event horizons and trapped surfaces) that therefore the flat region around B is quite different from the flat region around C, and that one cannot appeal to translation invariance in this flat region to rule this out.

5. A particularly interesting non-Vaidya example is provided by the Oppenheimer-Snyder collapse geometry.[137] Here an elementary observation is that due to the homogeneity (in particular the translation invariance in r) of the interior metric, a trapped sphere centered at $r = 0$ can be translated in r (as long as it stays inside the star) to give new trapped spheres, no longer centered at $r = 0$. Thus the region $\mathbb{T}$ containing trapped surfaces inside the star is definitely larger than the region enclosed by the apparent horizon $\mathcal{T}$ and the boundary of the star in Figure 42, but again its precise location has not yet been conclusively determined.

[136] I. Bengtsson, J. Senovilla, *A Note on trapped Surfaces in the Vaidya Solution*, `arXiv:0809.2213 [gr-qc]`, J. Senovilla, *On the boundary of the region containing trapped surfaces*, `arXiv:0812.2767 [gr-qc]`, and J. Aman, I. Bengtsson, J. Senovilla, *Where are the trapped surfaces?*, `arXiv:0912.3691v1 [gr-qc]`. See also I. Bengtsson, J. Senovilla, *The region with trapped surfaces in spherical symmetry, its core, and their boundaries*, `arXiv:1009.0225 [gr-qc]` for a much more general analysis of the same question, and J. Senovilla, *Trapped Surfaces*, `arXiv:1107.1344 [gr-qc]` for an updated account of what is known about the trapping boundary $\mathcal{B}$.

[137] I. Bengtsson, E. Jakobsson, J. Senovilla, *Trapped surfaces in Oppenheimer-Snyder black holes*, `arXiv:1306.6486 [gr-qc]`.

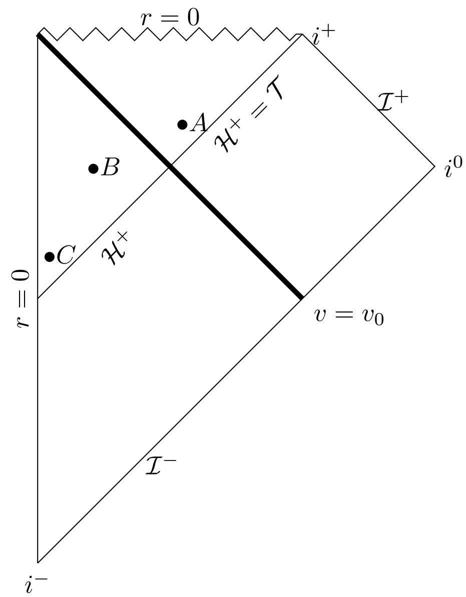


Figure 43: Trapped Surfaces in the collapse of a thin null shell to a black hole. See the body of the text for the description.

6. Apart from some very general properties, very little seems to be known at present about $\mathbb{T}$ and $\mathcal{B}=\partial\mathbb{T}$ in situations without spherical symmetry.

It seems appropriate to close this section with a quotation from the article just mentioned:

> We find it puzzling, and indeed intriguing, that the very simple questions we ask are so difficult to answer.[137]

F: Cosmology

33 Cosmology I: Basics

33.1 Preliminary Remarks

We now turn away from considering isolated systems (stars) to some (admittedly very idealised) description of the universe as a whole. This subject is known as Cosmology. It is certainly one of the most fascinating subjects of theoretical physics, dealing with such issues as the origin and ultimate fate and the large-scale structure of the universe.

Due to the difficulty of performing cosmological experiments and making precise measurements at large distances, many of the most basic questions about the universe are still unanswered today:

1. PAST:

   What actually happened at (or even before) what is usually called the Big Bang?

2. PRESENT:

   is our universe spatially finite or infinite?

3. FUTURE:

   Will our universe keep expanding forever or will it recollapse?

4. MATTER:

   What is Dark Matter?

5. ENERGY

   - Is Dark Energy, responsible for what appears to be a current phase of accelerated expansion of the universe, a cosmological constant?
   - If Dark Energy is indeed a cosmological constant, why is the cosmological constant so small and what determines its value?

While recent precision data, e.g. from supernovae surveys and detailed analysis of the cosmic microwave background radiation, suggest answers to at least some of these questions, these answers leave less wiggle-room for philosophical prejudices or esthetic preferences and actually just make the universe more mysterious than ever.

Of course, we cannot study any of these questions in detail, in particular because an important role in studying these questions is played by the interaction of cosmology with astronomy, astrophysics and elementary particle physics, each of these subjects deserving at least a course of its own.

Fortunately, however, many of the important features any realistic cosmological model should display are already present in some very simple models, the so-called *Friedmann-Lemaître-Robertson-Walker Models* (FLRW models) already studied in the 20's and

30's of the last century. They are based on the simplest possible ansatz for the metric compatible with the assumption that on large scales the universe is roughly homogeneous and isotropic (cf. the next section for a more detailed discussion of this *Cosmological Principle*) and have become the 'standard model' of cosmology.

We will see that they already display all the essential features such as

1. a Big Bang
2. an expanding universes (Hubble expansion) and a cosmological redshift.
3. different long-term behaviour (eternal expansion versus recollapse)

Our first aim will be to make maximal use of the symmetries that simple cosmological models should have to find a simple ansatz for the metric. Our guiding principle will be the *Cosmological Principle.*

33.2 Fundamental Assumption: The Cosmological Principle

At first, it may sound impossibly difficult to find solutions of the Einstein equations describing the universe as a whole. However: if one looks at the universe at large (very large) scales, in that process averaging over galaxies and even clusters of galaxies, then the situation simplifies a lot in several respects:

1. First of all, at those scales non-gravitational interactions can be completely ignored because they are either short-range (the nuclear forces) or compensate each other at large distances (electro-magnetism).
2. Furthermore we assume that the earth, and our solar system, or even our galaxy, have no privileged position in the universe (this is occasionally referred to as the *Copernican Principle*). This means that at large scales the universe should look the same from any point in the universe. Mathematically this means that there should be translational symmetries from any point of space to any other, in other words, space should be homogeneous.
3. Also, we assume that, at large scales, the universe looks the same in all directions. Thus there should be rotational symmetries and hence space should be isotropic.

Together, the second and third assumptions form the *Cosmological Principle*, which is the starting point for our discussion of cosmology and on which much of the work in cosmology is based. It is plausible (and true) that the assumption of isotropy (around us) can be tested experimentally / observationally, while testing the assumption of homogeneity is evidently going to be more tricky.[138]

Making the above assumptions, it follows from our discussion in section 14, that the n-dimensional space (of course $n = 3$ for us) has n translational and $n(n-1)/2$ rotational Killing vectors, i.e. that the spatial metric is *maximally symmetric*. For $n = 3$, we will thus have six Killing vectors, two more than for the Schwarzschild metric, and the ansatz for the metric will simplify accordingly.

Note that, since we know from observation that the universe expands, we do not require a priori a maximally symmetric space-*time* as this would imply that there is also a timelike Killing vector.

What simplifies life considerably is the fact that (cf. the discussion in section 14) there are only three species of maximally symmetric spaces (for any n), namely

- flat space $\mathbb{R}^n$ (with its standard Euclidean metric),
- the sphere S^n (with its standard "round" metric),
- and its negatively curved counterpart, the n-dimensional pseudosphere or hyperboloid we will call H^n.

Thus, for a space-time metric with maximally symmetric spacelike 'slices', the only unknown is the time-dependence of the overall size of the metric. More concretely, the metric can (now fixing the number of spatial dimensions to be $n = 3$) be chosen to be

$$ds^2 = -dt^2 + a^2(t)(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2) \quad , \tag{33.1}$$

where $k = 0, \pm 1$ corresponds to the three possibilities mentioned above. Thus the metric contains only one unknown function, the 'radius' or cosmic scale factor $a(t)$. This function will be determined by the Einstein equations via the matter content of the universe (we will of course be dealing with a non-vanishing energy-momentum tensor), modelled by a perfect fluid.

33.3 Fundamental Observations I: Olbers' Paradox

One paradox, popularised by Olbers (1826) but noticed before by others is the following. He asked the seemingly innocuous question "Why is the sky dark at night?". According to his calculation, reproduced below, the sky should instead be infinitely bright.

[138] The extent to which these assumptions, in particular homogeneity, can be observationally tested is discussed e.g. in C. Clarkson, R. Maartens, *Inhomogeneity and the foundations of concordance cosmology*, `arXiv:1005.2165 [astro-ph.CO]`, G. Ellis, *Inhomogeneity effects in Cosmology*, `arXiv:1103.2335 [astro-ph.CO]`, R. Maartens, *Is the Universe homogeneous?*, `arXiv:1104.1300 [astro-ph.CO]`.

The simplest assumption one could make in cosmology (prior to the discovery of the Hubble expansion) is that the universe is static, infinite and homogeneously filled with stars. In fact, this is probably the naive picture one has in mind when looking at the stars at night, and certainly for a long time astronomers had no reason to believe otherwise.

However, these simple assumptions immediately lead to a paradox, namely the conclusion that the night-sky should be infinitely bright (or at least very bright) whereas, as we know, the sky is actually quite dark at night. This is a nice example of how very simple observations can actually tell us something deep about nature (in this case, the nature of the universe). The argument runs as follows.

1. Assume that there is a star of brightness (luminosity) L at distance r. Then, since the star sends out light into all directions, the apparent luminosity A (neglecting absorption) will be
$$A(r) = L/4\pi r^2 \quad . \tag{33.2}$$

2. If the number density ν of stars is constant, then the number of stars at distances between r and $r + dr$ is
$$dN(r) = 4\pi\nu r^2 dr \quad . \tag{33.3}$$
Hence the total energy density due to the radiation of all the stars is
$$E = \int_0^\infty A(r)dN(r) = L\nu \int_0^\infty dr = \infty \quad . \tag{33.4}$$

3. Therefore the sky should be infinitely bright.

Now what is one to make of this? Clearly some of the assumptions in the above are much too naive. The way out suggested by Olbers is to take into account absorption effects and to postulate some absorbing interstellar medium, but this is also too naive because in an eternal universe we should now be in a stage of thermal equilibrium. Hence the postulated interstellar medium should emit as much energy as it absorbs, so this will not reduce the radiant energy density either.

Of course, the stars themselves are not transparent, so they could block out light completely from distant sources, but if this is to rescue the situation, one would need to postulate so many stars that every line of sight ends on a star, but then the night sky would be bright (though not infinitely bright) and not dark.

Modern cosmological models can resolve this problem in a variety of ways. For instance, the universe could be static but finite (there are such solutions, but this is nevertheless an unlikely scenario) or the universe is not eternal since there was a 'Big Bang' (and this is a more likely scenario).

We have already discussed one of the fundamental inputs of simple cosmological models, namely the cosmological principle. This led us to consider space-times with maximally-symmetric spacelike slices. One of the few other things that is definitely known about the universe, and that tells us something about the time-dependence of the universe, is that it expands or, at least that it appears to be expanding.

In fact, in the 1920's and 1930's, the astronomer Edwin Hubble made a remarkable discovery regarding the motion of galaxies. He found that light from distant galaxies is systematically redshifted (increased in wave-length λ), the increase being proportional to the distance d of the galaxy,

$$z := \frac{\Delta\lambda}{\lambda} \propto d \quad . \tag{33.5}$$

Hubble interpreted this redshift as due to a Doppler effect and therefore ascribed a *recessional velocity* $v = cz$ to the galaxy. While, as we will see, this pure Doppler shift explanation is not tenable or at least not always the most useful way of phrasing things, the terminology has stuck, and Hubble's law can be written in the form

$$v = Hd \quad , \tag{33.6}$$

where H is *Hubble's constant.* To set the historical record straight: credit for this fundamental discovery should perhaps (also) go to G. Lemaître.[139]

We will see later that in most cosmological models H is actually a function of time, so the H in the above equation should then be interpreted as the value H_0 of H today. It is one of the main goals of observational cosmology to determine H_0 and H as precisely as possible, and the main problem here is naturally a precise determination of the distances of distant galaxies. This is a complex and fascinating issue in its own right, but one that we will not go into here (safe for a brief mention of the *luminosity distance* at the end of section 34.9).[140] I will just conclude this section with one comment on the units usually employed to express galactic distances and the Hubble constant H_0.

Galactic distances are frequently measured in *mega-parsecs* (Mpc). A parsec is the distance from which a star subtends an angle of 2 arc-seconds at the two diametrically opposite ends of the earth's orbit. This unit arose because of the old trigonometric

[139] See e.g. M. Way, H. Nussbaumer, *The linear redshift-distance relationship: Lemaître beats Hubble by two years*, `arXiv:1104.3031v1 [physics.hist-ph]`; J.-P. Luminet, *Editorial note to "The beginning of the world from the point of view of quantum theory"*, `arXiv:1105.6271v1 [physics.hist-ph]`, and in particular also M. Livio, *The Expanding Universe: Lost (in Translation) and Found*, `http://hubblesite.org/pubinfo/pdf/2011/36/pdf.pdf` for an important addition to this debate.

[140] For a beautiful introduction to this subject of the *Cosmic Distance Ladder*, see the pdf-slides of a public talk by Fields Medalist (the mathematics counterpart of a Nobel Laureate!) Terry Tao at `http://terrytao.files.wordpress.com/2010/10/cosmic-distance-ladder.pdf` (you can also find a video of the talk on youtube).

method of measuring stellar distances (a triangle is determined by the length of one side and the two adjacent angles). 1 parsec is approximately 3×10^{18} cm, a little over 3 light-years. The Hubble constant is therefore often expressed in units of km s^{-1} (Mpc)$^{-1}$. The best currently available estimates point to a value of H_0 in the range (using a standard parametrisation)

$$\begin{aligned} H_0 &= 100h\text{km/s/Mpc} \ , \\ h &= 0.71 \pm 0.06 \ . \end{aligned} \tag{33.7}$$

We will usually prefer to express it just in terms of inverse units of time. The above result leads to an order of magnitude range of

$$H_0^{-1} \approx 10^{10} \text{ years} \tag{33.8}$$

(whereas Hubble's original estimate was more in the 10^9 year range).

34 Cosmology II: Geometry and Physics of Robertson-Walker Metrics

34.1 Mathematical Model: the Robertson-Walker Metric

Having determined that the metric of a maximally symmetric space is of the simple form (14.27), we can now deduce that a space-time metric satisfying the Cosmological Principle can be chosen to be of the form (33.1),

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right] \quad . \tag{34.1}$$

Here we have used the fact that (as in the ansatz for a spherically symmetric metric) non-trivial g_{tt} and g_{tr} can be removed by a coordinate transformation. This metric is known as the *Friedmann-Robertson-Walker metric* or just the Robertson-Walker metric, and spatial coordinates in which the metric takes this form are called *comoving coordinates*, for reasons that will become apparent below. The function $a(t)$, the "radius of the universe", is known as the *cosmic (or cosmological) scale factor*.

It is perhaps useful at this point to recall the symmetries of this metric (the isometries). See sections 14.4 and 39.1 for a more detailed and general discussion:

1. For $k = 0$, the spatial part of the metric is just (up to the overall factor $a(t)^2$), the Euclidean metric on $\mathbb{R}^3$. The isometry group of the Euclidean metric is the 6-parameter Euclidean group $\mathsf{E}(3)$, consisting of spatial translations and spatial rotations,

$$k = 0: \quad \text{Isometry Group} = \mathsf{E}(3) = \mathsf{SO}(3) \ltimes \mathbb{R}^3 \quad . \tag{34.2}$$

   These are also symmetries of the space-time metric (34.1) for $k = 0$. For generic choices of $a(t)$, there are no further isometries of the metric.

2. For $k = +1$, the spatial part of the metric is just (up to the overall factor $a(t)^2$), the standard metric on the 3-sphere S^3, the one induced on S^3 from the Euclidean metric

$$ds^2 = dx^2 + dy^2 + dz^2 + dw^2 \tag{34.3}$$

   on $\mathbb{R}^4$ via the embedding

$$S^3 = \{(x, y, z, w) \in \mathbb{R}^4 : \; x^2 + y^2 + z^2 + w^2 = 1\} \quad . \tag{34.4}$$

   The isometry group of this metric on S^3 is therefore the group $\mathsf{SO}(4)$ of 4-dimensional rotations leaving the 4-dimensional Euclidean metric (and norm) invariant,

$$k = +1: \quad \text{Isometry Group} = \mathsf{SO}(4) \quad . \tag{34.5}$$

This group is also 6-dimensional. These are also symmetries of the space-time metric (34.1) for $k = +1$. For generic choices of $a(t)$, there are no further isometries of the metric.

3. For $k = -1$, the spatial part of the metric is just (up to the overall factor $a(t)^2$), the standard metric on the 3-hyperboloid H^3, the one induced on H^3 from the Lorentzian metric

$$ds^2 = dx^2 + dy^2 + dz^2 - dw^2 \tag{34.6}$$

on $\mathbb{R}^4$ via the embedding

$$H^3 = \{(x, y, z, w) \in \mathbb{R}^4 : \; x^2 + y^2 + z^2 - w^2 = -1\} \; . \tag{34.7}$$

The isometry group of this metric on H^3 is therefore the group $\mathsf{SO}(3,1)$ of 4-dimensional rotations leaving the 4-dimensional Lorentzian metric (and norm) invariant,

$$k = -1 : \quad \text{Isometry Group} = \mathsf{SO}(3,1) \; . \tag{34.8}$$

This group is also 6-dimensional. These are also symmetries of the space-time metric (34.1) for $k = -1$. For generic choices of $a(t)$, there are no further isometries of the metric.

<u>Remarks:</u>

1. The isometry group for $k = -1$ happens to be isomorphic to the Lorentz group of 4-dimensional Minkowski space. However, here it is realised purely on the spatial part of the metric, while the time coordinate t is not transformed at all. We will later on write the $k = -1$ metric in such a way, via a suitable coordinate transformation, that the $\mathsf{SO}(3,1)$ symmetries are realised precisely by standard Lorentz transformations on the coordinates (cf. section 34.6).

2. As mentioned in the discussion around (14.29), a rescaling of k by a positive constant, $k \to k/L^2$, is equivalent to an overall constant rescaling of the spatial metric,

$$ds^2 = \frac{dr^2}{1 - kr^2/L^2} + r^2 d\Omega^2 = L^2 \left(\frac{d\tilde{r}^2}{1 - k\tilde{r}^2} + \tilde{r}^2 d\Omega^2 \right) \; , \tag{34.9}$$

where $\tilde{r} = r/L$. In the context of the Robertson-Walker metric (34.1), a scaling of k is thus equivalent to a scaling of the cosmic scale factor $a(t)$. This scaling freedom can be used

   - either to fix $k = 0, \pm 1$ once and for all,
   - or to normalise $a(t)$ in a convenient way, e.g. by the condition $a(t_0) = 1$ for t_0 the time "today".

In the latter case, the 3 different possibilities for the spatial geometry, are distinguished by $k < 0, k = 0, k > 0$. For the most part we work with the first option, but for certain questions (like "why is space so close to being flat today?") it is convenient to rephrase and express this in terms of k ("why is k so close to zero?"), which is evidently only meaningful if one does not restrict k to the 3 discrete values $k = 0, \pm 1$.

3. Another convenient way of representing this metric, that we will occasionally make use of below, is as

$$ds^2 = -dt^2 + a(t)^2(d\psi^2 + g_k(\psi)^2 d\Omega^2) \quad . \tag{34.10}$$

where (cf. (14.34) and (14.35))

$$g_k(\psi) = \begin{cases} \psi & k = 0 \\ \sin\psi & k = +1 \\ \sinh\psi & k = -1 \end{cases} \tag{34.11}$$

4. The metric of the three-space at constant t is

$$g_{ij} = a^2(t)\tilde{g}_{ij} \quad , \tag{34.12}$$

where $\tilde{g}_{ij}$ is the maximally symmetric spatial metric. Thus for $k = +1$, $a(t)$ directly gives the size (radius) of the universe. For $k = -1$, space is infinite, so no such interpretation is possible, but nevertheless $a(t)$ still sets the scale for the geometry of the universe, e.g. in the sense that the curvature scalar $R^{(3)}$ of the metric g_{ij} is related to the curvature scalar $\tilde{R}^{(3)}$ of $\tilde{g}_{ij}$ by

$$R^{(3)}(t) = \frac{1}{a^2(t)}\tilde{R}^{(3)} \quad . \tag{34.13}$$

Finally, for $k = 0$, three-space is flat and also infinite, but one could replace $\mathbb{R}^3$ by a three-torus T^3 (still flat but now compact) and then $a(t)$ would once again be related directly to the size of the universe at constant t.

5. Through the dependence of $a(t)$ on t, proper length scales and distances in the constant time surfaces depend on time. Thus $a(t)$ changes or sets the scale, i.e. $a(t)$ plays the role of a cosmological scale factor.

6. Note that the case $k = +1$ opened up for the very first time the possibility of considering, even conceiving, an unbounded but finite universe! These and other generalisations made possible by a general relativistic approach to cosmology are important as more naive (Newtonian) models of the universe immediately lead to paradoxes or contradictions (as we have seen e.g. in the discussion of Olbers' paradox in section 33.3).

34.2 Timelike Geodesics and Comoving Observers

We now look at timelike geodesics in the Robertson-Walker geometry. We will analyse the general case momentarily, but we can already identify one privileged class of geodesic observers by inspection. Indeed, as already noted (way back) in the discussion around (3.11), in a space-time of the Robertson-Walker form there is a particularly simple class of timelike geodesics, given by fixing once and for all the spatial coordinates of the observer.

Let us quickly rederive this result here. Note that, since $g_{tt} = -1$ is a constant and the off-diagonal time - space components of the metric are zero, $g_{tk} = 0$, one has

$$\Gamma_{\mu tt} = 0 \ . \tag{34.14}$$

Therefore the vector field ∂_t is geodesic, which can be expressed as the statement that

$$\nabla_t \partial_t := \Gamma^\mu_{\ tt} \partial_\mu = 0 \ . \tag{34.15}$$

In simpler terms this means that the curves $\vec{x} = const.$ ($\vec{x}$ referring to the spatial coordinates),

$$\tau \to (t(\tau), \vec{x}(\tau)) = (\tau, \vec{x}_0) \tag{34.16}$$

are geodesics.

Hence, in this coordinate system, observers remaining at fixed values of the spatial coordinates are in free fall. In other words, the coordinate system is falling with them or *comoving*, and the proper time τ along such geodesics coincides with the coordinate time or cosmic time t, $d\tau = dt$. It is these observers of constant $\vec{x}$ or constant (r, θ, ϕ) who all see the same isotropic universe at a given value of t.

Remarks:

1. This may sound a bit strange but a good way to visualise such a coordinate system is, as in Figure 44, as a mesh of coordinate lines drawn on a balloon that is being inflated or deflated (according to the behaviour of $a(t)$). Draw some dots on that balloon (that will eventually represent galaxies or clusters of galaxies). As the balloon is being inflated or deflated, the dots will move but the coordinate lines will move with them and the dots remain at fixed spatial coordinate values. Thus, as we now know, regardless of the behaviour of $a(t)$, these dots follow a geodesic, and we will thus think of galaxies in this description as being in free fall.

2. Recall that we had already encountered analogous comoving coordinates in our discussion of Lemaître coordinates for the Schwarzschild metric in section 27.3 and subsequently in equation (29.26) of section 29.3, when we had introduced proper time of the freely falling particles on the surface of the star to describe the metric induced on the surface of the star.

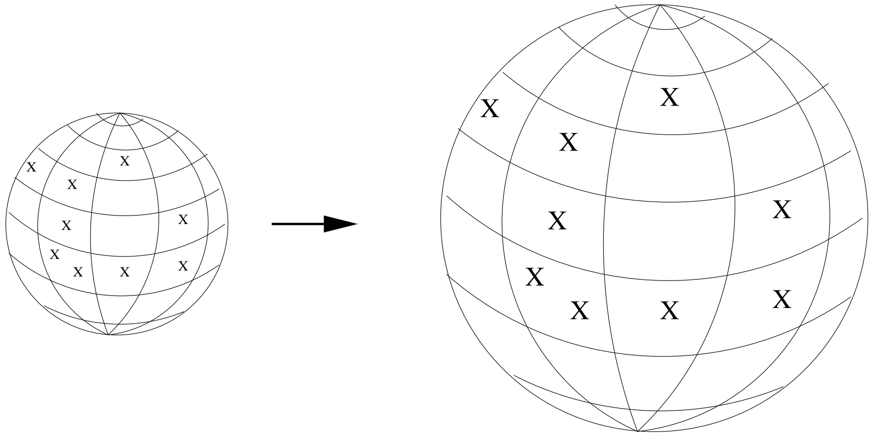


Figure 44: Illustration of a comoving coordinate system: Even though the sphere (universe) expands, the X's (galaxies) remain at the same spatial coordinates. These trajectories are geodesics and hence the X's (galaxies) can be considered to be in free fall. The figure also shows (cf. the discussion in section 6.8) that it is the number density per unit coordinate volume that is conserved, *not* the density per unit proper volume.

3. Another advantage of the comoving coordinate system is that the six-parameter family of isometries just acts on the spatial part of the metric. Indeed, let $K^i\partial_i$ be a Killing vector of the maximally symmetric spatial metric. Then $K^i\partial_i$ is also a Killing vector of the Robertson-Walker metric. This would not be the case if one had e.g. made an x-dependent coordinate transformation of t or a t-dependent coordinate transformation of the x^i. In those cases there would of course still be six Killing vectors, but they would have a more complicated form.

The worldlines of comoving observers discussed above are special timelike geodesics. To discuss the general case, it will be convenient to write the metric in the form (34.10)

$$ds^2 = -dt^2 + a(t)^2(d\psi^2 + g_k(\psi)^2 d\Omega^2) \ . \tag{34.17}$$

By spatial maximal symmetry and the associated conserved (angular) momenta, we can without loss of generality consider motion in the (t,ψ)-direction, so that $\dot{t}$ and $\dot{\psi}$ are related by

$$\dot{t}^2 - a(t)^2\dot{\psi}^2 = 1 \ , \tag{34.18}$$

where, as usual, an overdot refers to differentiation with respect to proper time.

Even though we do not have a timelike Killing vector (and its associated conserved energy) to further simplify this, in the case at hand we have plenty of spacelike Killing vectors $V^\alpha\partial_\alpha$ with $V^t = 0$. Among them there will be ψ-translational Killing vectors which have the form

$$V = f(\theta,\phi)\partial_\psi + \dots \tag{34.19}$$

(the 3-dimensional counterparts of the Killing vectors $V_{(1)}, V_{(2)}$ (9.55) of the 2-sphere, say). Associated to any such Killing vector and the timelike geodesic there is the conserved momentum

$$P = \dot{x}^\alpha V_\alpha = a(t)^2 \dot{\psi} f(\theta, \phi) \ . \tag{34.20}$$

Since θ and ϕ are constant along the lightray, we can absorb $f(\theta, \phi)$ into the definition of P, and thus we have

$$\dot{\psi} = P/a^2 \ . \tag{34.21}$$

Therefore (34.18) can be written as

$$\dot{t} = \sqrt{1 + P^2/a^2} \ . \tag{34.22}$$

In particular, we see that comoving observers are characterised by $P = 0$,

$$P = 0 \quad \Rightarrow \quad \dot{\psi} = 0 \quad \text{(comoving)} \ , \tag{34.23}$$

and that precisely for these observers the cosmic time t coincides with their proper time,

$$P = 0 \quad \Leftrightarrow \quad \dot{t} = 1 \quad \Leftrightarrow \quad dt = d\tau \ . \tag{34.24}$$

Nevertheless, even in the general case it is useful to combine the two previous equations to obtain an equation for ψ as a function of t, namely (assuming $P \neq 0$)

$$\frac{d\psi(t)}{dt} = \frac{\dot{\psi}}{\dot{t}} = \frac{P}{\sqrt{a(t)^4 + P^2 a(t)^2}} = \frac{1}{a(t)\sqrt{1 + a(t)^2/P^2}} \ . \tag{34.25}$$

In general, even for simple power-law behaviours for $a(t)$ (which we will typically find as solutions to the Einstein equations in the spatially flat case $k = 0$), this equation cannot be solved in closed form (but can be approximated by a tractable, even elementary, integral when $a(t) \gg |P|$ or when $a(t) \ll |P|$).

34.3 Velocity - Distance Relation, Recessional Velocities and the Hubble Sphere

The Robertson-Walker metric immediately, and in complete generality, implies a crude distance - velocity relation reminiscent of Hubble's law. Namely, let us ego- or geocentrically place ourselves at the origin $r = 0$ (remember that because of maximal symmetry this point is as good as any other and in no way privileged). Consider another galaxy following the comoving geodesic at the fixed value $r = r_1$. Its "instantaneous" proper distance $R_p(t)$ at time t can be calculated from

$$dR_p = a(t)\frac{dr}{(1 - kr^2)^{1/2}} \quad \Rightarrow \quad R_p(t) = a(t) f_k(r_1) \ , \tag{34.26}$$

where

$$f_k(r) = \begin{cases} r & k=0 \\ \arcsin r & k=+1 \\ (\sinh)^{-1} r & k=-1 \end{cases} \tag{34.27}$$

Note that the $f_k(r)$ are the inverses of the functions $g_k(\psi)$ defined in (34.11) (the precise form of $f_k(r)$ will however be irrelevant for this argument). If we use the coordinates (t,ψ) instead of (t,r), the instantaneous proper distance to a point with coordinate ψ_1 simply has the form

$$R_p(t) = a(t)\psi_1 \quad . \tag{34.28}$$

It follows that its proper velocity is

$$V_p(t) \equiv \frac{d}{dt} R_p(t) = \dot{a}(t)\psi_1 = H(t) R_p(t) \quad . \tag{34.29}$$

Here we have introduced the *Hubble parameter*

$$H(t) = \frac{\dot{a}(t)}{a(t)} \quad , \tag{34.30}$$

which plays a pivotal role in cosmology.

The relation (34.29) clearly expresses something like Hubble's law $v = Hd$ (33.6): all objects run away from each other with velocities proportional to their distance. And the above derivation shows that such an (at first sight perhaps surprising) relationship between (some kind of) velocity and (some kind of) distance can arise as a general purely geometric statement about the geometry of space-time. We will have much more to say about $H(t)$, and about the relation between distance and redshift z (which is what is actually observed), below.

Calculating the second time-derivative of $R_p(t)$, one finds

$$\frac{d^2}{dt^2} R_p(t) = (\frac{d}{dt} H(t)) R_p(t) + H(t) \frac{d}{dt} R_p(t) = \frac{\ddot{a}(t)}{a(t)} R_p(t) \tag{34.31}$$

since

$$\frac{d}{dt} H(t) = \frac{\ddot{a}(t)}{a(t)} - \frac{\dot{a}(t)^2}{a(t)^2} = \frac{\ddot{a}(t)}{a(t)} - H(t)^2 \quad . \tag{34.32}$$

Thus the cosmological expansion or contraction can be visualised as acting like a linear harmonic oscillator force on the separation of comoving objects, with (in general time-dependent) real or imaginary frequency $\omega(t)$,

$$\frac{d^2}{dt^2} R_p(t) + \omega(t)^2 R_p(t) = 0 \quad , \quad \omega(t)^2 = -\frac{\ddot{a}(t)}{a(t)} \quad . \tag{34.33}$$

Universes with an accelerating expansion thus lead to imaginary frequencies, and hence to an exponential-like rather than harmonic motion (over periods of time during which the time-dependence of the frequency can be neglected).

If the object is not comoving, i.e. is not sitting at a fixed value of ψ, say, one has

$$R_p(t) = a(t)\psi(t) \quad , \tag{34.34}$$

and then its proper velocity is

$$V_p(t) = \frac{d}{dt}R_p(t) = H(t)R_p(t) + a(t)\frac{d\psi(t)}{dt} \quad . \tag{34.35}$$

Here "proper" refers to the fact that it is the t-derivative of the instantaneous proper distance, but it should be kept in mind that t is not the proper time for non-comoving observers. This proper velocity can be decomposed into what are known as its *recessional velocity* (equal to the proper velocity for a comoving object) and its *peculiar velocity* (relative to the recessional velocity or *Hubble flow*),

$$V_p(t) = H(t)R_p(t) + a(t)\frac{d\psi(t)}{dt} \equiv V_{rec}(t) + V_{pec}(t) \quad . \tag{34.36}$$

As a consequence of (34.18), one has

$$V_{pec}(t)^2 = a(t)^2 \left(\frac{d\psi(t)}{dt}\right)^2 < 1 \quad . \tag{34.37}$$

Alternatively this follows from (34.25), which allows us to express the peculiar velocity as

$$V_{pec}(t) = a(t)\frac{d\psi(t)}{dt} = \frac{1}{\sqrt{1 + a(t)^2/P^2}} \quad . \tag{34.38}$$

Either way we see that the peculiar velocity is always "subluminal".

However, the recessional velocity is not restricted in this way and in most cosmological models superluminal recessional velocities will occur for objects which are sufficiently far away (in the sense of having a sufficiently large proper distance $R_p(t)$).

Remarks:

1. There is absolutely nothing illegal or pathological about this because V_{rec} measures the rate of an increase in distance between two objects, not any locally measurable velocity. For example, even in Special Relativity, if in your rest-frame you send off two objects in opposite directions at speeds $> c/2$ each, then the distance between them grows at a rate (measured with respect to your proper time) larger than c, but clearly you have not violated or disproven Special Relativity by doing this.

2. Moreover, in a curved space-time a comparison of vectors (say velocity vectors) at different points requires some care. Indeed, as explained in section 5.8, this requires some notion of parallel transport in order to have two vectors at the same point that one can meaningfully compare.[141]

[141] See e.g. A. Kaya, *Hubble's law and faster than light expansion speeds*, `arXiv:1107.5168 [physics.gen-ph]` for an analysis of this issue from the parallel transport point of view.

3. For a further illustration of the harmlessness of superluminal recessional velocities in the present context, it suffices to consider the special case $a(t) = t$, in which case one has

$$a(t) = t \quad \Rightarrow \quad H(t) = 1/t \quad \Rightarrow \quad V_{rec}(t) = R_p(t)/t \ . \tag{34.39}$$

   At any given time t, objects further away than t appear to have superluminal recession velocities. On the other hand, as shown in section 37.1, the space-time with $a(t) = t$ and $k = -1$ is just a part of Minkowski space. Thus even Minkowski space can be foliated in such a way (by hyperboloids in the future lightcone) that events appear to have superluminal recession velocities, but evidently there is nothing here that violates any of the postulates of Special Relativity.

4. Cosmologists frequently refer to the sphere beyond which the recessional velocity exceeds the speed of light as the *Hubble sphere*. Its radius $R_H(t)$ at time t, the *Hubble radius*, is (restoring for once and temporarily the speed of light c)

$$V_{rec}(t) = c \quad \Leftrightarrow \quad R_H(t) = c/H(t) \ . \tag{34.40}$$

   Misleadingly, this surface is also often referred to as the *Hubble horizon*, the reason for this apparently being the idea or belief that we can never observe objects outside the Hubble sphere, but this is in general not correct. In particular, it is not correct to say (perhaps based on a mistaken analogy with special relativistic reasoning) that objects with recessional velocities $V_{rec} > c$ are infinitely redshifted and therefore invisible to us. There is indeed a cosmological redshift, worked out in section 34.8, and there is also an ensuing Hubble-like redshift - distance relation in Robertson-Walker geometries, derived in section 34.9. However, this cannot be written as a standard special relativistic recessional velocity - redshift relation involving the recessional velocity.[142]

5. It is true that there are limits to how much of the universe one can observe at any given time, and it is also true that in certain situations the Hubble radius $R_H(t)$ provides one with an order of magnitude estimate of the size of the visible universe.

   However, it is certainly misleading (even though some people appear to be obsessed with this) to think of the visible universe (or the inside of the Hubble sphere) as somehow being like the inside of a Schwarzschild black hole or some such nonsense. In fact, with any standard definition of a black hole going beyond pop-sci culture wisdom this statement is so obviously wrong or misleading in so many respects that I don't even know where to start (so don't get me started). Nevertheless I will briefly revisit this claim in section 36.9. For the time you may enjoy poking holes into this statement yourself . . .

[142] See also T. Davis, C. Lineweaver, *Superluminal Recession Velocities*, `arXiv:astro-ph/0011070` for a nice discussion of this.

We will return to some of these issues in a (slightly) more quantitative way later on, in sections 36.7 – 36.9, with the Friedmann equations (i.e. the Einstein equations for the standard model of cosmology we are in the process of developing) at our disposal.

34.4 Painlevé-Gullstrand-like Coordinates for Comoving Observers

In section 27.2 we introduce coordinates for the Schwarzschild metric that are adapted to radial geodesic observers, known as Painlevé-Gullstrand coordinates. We can do something analogous for the comoving observers of the Robertson-Walker metrics.

For $k = 0$ the Robertson-Walker metric is

$$ds^2 = -dt^2 + a(t)^2(dr^2 + r^2 d\Omega^2) \quad . \tag{34.41}$$

To put this into PG-like form, we keep t (which is, after all, already the proper time of comoving observers) but introduce, instead of the comoving coordinate r the area radius

$$\tilde{r}(t, r) = a(t)r \quad . \tag{34.42}$$

In terms of this the Robertson-Walker metric takes the PG-like form

$$\begin{aligned} ds^2 &= -(1 - \tilde{r}^2 H(t)^2)dt^2 - 2\tilde{r}H(t)dtd\tilde{r} + (d\tilde{r}^2 + \tilde{r}^2 d\Omega^2) \\ &= -dt^2 + (d\tilde{r} - \tilde{r}H(t)dt)^2 + \tilde{r}^2 d\Omega^2 \quad , \end{aligned} \tag{34.43}$$

where, as above, $H(t) = \dot{a}(t)/a(t)$ is the Hubble parameter.

<u>Remarks:</u>

1. This is clearly analogous to the Schwarzschild metric in PG coordinates (27.12)

$$\begin{aligned} ds^2 &= -(1 - 2m/r)dT^2 + 2\sqrt{2m/r}dTdr + (dr^2 + r^2 d\Omega^2) \\ &= -dT^2 + \left(dr + \sqrt{2m/r}dT\right)^2 + r^2 d\Omega^2 \quad . \end{aligned} \tag{34.44}$$

2. Just as the Schwarzschild metric in PG coordinates is adapted to observers with $dr = -\sqrt{2m/r}dT$ (for which T is proper time, $\dot{r} = -\sqrt{2m/r}$ describing geodesic radial free fall with $E = 1$), the metric (34.43) is adapted to observers with

$$\frac{d}{dt}\tilde{r} = H(t)\tilde{r} \tag{34.45}$$

which are precisely the comoving observers $\tilde{r}(t) = a(t)r$ with r fixed obeying the Hubble law (34.29).

3. Performing the same coordinate transformation to the area radius $\tilde{r} = a(t)r$ for $k \neq 0$, one finds the metric

$$ds^2 = -dt^2 + \frac{1}{1 - k\tilde{r}^2/a^2}\left(d\tilde{r} - \tilde{r}H(t)dt\right)^2 + \tilde{r}^2 d\Omega^2 \quad . \tag{34.46}$$

4. While these PG-like coordinates are not widely used in the cosmological context, we make use of them in section 29.5, in the description of the interior geometry of a collapsing star (because this PG-like form of the metric makes it particularly easy to match the interior metric to the exterior Schwarzschild metric).

34.5 Conformal Time η

Writing the Robertson-Walker metric as

$$ds^2 = -dt^2 + a(t)^2 d\tilde{s}^2 = a(t)^2[-dt^2/a(t)^2 + d\tilde{s}^2] \quad , \tag{34.47}$$

where a tilde refers to the maximally symmetric spatial metric, we see that it is natural to introduce a new time-coordinate η through

$$d\eta = dt/a(t) \quad , \tag{34.48}$$

in terms of which the Robertson-Walker metric takes the simple form

$$d\eta = dt/a(t) \quad \Rightarrow \quad ds^2 = a^2(\eta)(-d\eta^2 + d\tilde{s}^2) \tag{34.49}$$

($a(\eta)$ is short (and sloppy) for $a(t(\eta))$). In terms of polar coordinates (34.10), this becomes

$$ds^2 = a^2(\eta)(-d\eta^2 + d\psi^2 + g_k(\psi)^2 d\Omega^2) \quad . \tag{34.50}$$

Remarks:

1. In particular, "radial" null lines are determined by $d\eta = \pm d\psi$, as in flat space, and η is also known as *conformal time*. This coordinate is very convenient for discussing the causal structure of the Friedmann-Lemaître-Robertson-Walker universes.

2. Since the η-dependence resides exclusively in the overall conformal factor of the metric, the vector ∂_η is a *conformal Killing vector* of the Robertson-Walker metric in the sense of (10.5),

$$C = \partial_\eta : \quad \nabla_\alpha C_\beta + \nabla_\beta C_\alpha = 2\frac{a'(\eta)}{a(\eta)}\, g_{\alpha\beta} = 2\dot{a}(t)\, g_{\alpha\beta} \quad . \tag{34.51}$$

3. We will use conformal time immediately below to establish the conformal flatness of the Robertson-Walker metrics, in section 37.3 to solve the cosmological Einstein equations in a particular case, and we will use the above conformal Killing vector and the associated conserved charge for null geodesics (10.9) in the discussion of the cosmological redshift in section 34.8.

In the spatially flat case $k = 0$, (34.50) shows that Robertson-Walker metrics are conformally flat, here written in radial polar coordinates with $\psi = r$,

$$k = 0 \quad \Rightarrow \quad ds^2 = a^2(\eta)(-d\eta^2 + dr^2 + r^2 d\Omega^2) = a(\eta)^2 \eta_{\alpha\beta} dx^\alpha dx^\beta \quad . \tag{34.52}$$

This is actually true for all Robertson-Walker metrics, i.e. also for $k \neq 0$, but the coordinate transformation required to exhibit this as explicitly as in the $k = 0$ case is somewhat more involved (and, as we will see below, the conformal factor does not depend just on the Minkowski time coordinate).

As a first step it will be convenient to introduce null coordinates

$$u = \eta - \psi \quad , \quad v = \eta + \psi \quad , \tag{34.53}$$

in terms of which the metric (34.50) takes the form

$$ds^2 = a^2(\eta)\left[-du\, dv + g_k((v-u)/2)^2\, d\Omega^2\right] \quad . \tag{34.54}$$

This metric is conformally flat if the line element in brackets, which does not depend on the cosmic scale factor $a(t)$ or $a(\eta)$, is conformally flat, so let us focus on

$$d\bar{s}^2 = ds^2/a^2(\eta) = -du\, dv + g_k((v-u)/2)^2\, d\Omega^2 \quad , \tag{34.55}$$

where, recall, $g_k(\psi) = \sin(\mathrm{h})\psi$ for $k = \pm 1$ respectively.

Given a metric of this form, it is natural to consider transformations of the form

$$U = U(u) \quad , \quad V = V(v) \tag{34.56}$$

because under such transformations $dudv$ and $dUdV$ are just related by an overall conformal factor. It is now straightforward to check, using some basic trigonometric identities (or their hyperbolic counterparts) that the specific coordinate transformation

$$U = 2\tan(\mathrm{h})u/2 \quad , \quad V = 2\tan(\mathrm{h})v/2 \tag{34.57}$$

(we will consider an alternative transformation for $k = -1$ below) is such that it transforms $d\bar{s}^2$ into

$$d\bar{s}^2 = \left(\cos(\mathrm{h})^2 u/2\right)\left(\cos(\mathrm{h})^2 v/2\right)\left[-dU\, dV + \tfrac{1}{4}(V-U)^2 d\Omega^2\right] \quad . \tag{34.58}$$

Now the metric in brackets is just the Minkowski metric, written in radial null coordinates, as can be seen by undoing the transformation to null coordinates through

$$U = T - R \quad , \quad V = T + R \quad , \tag{34.59}$$

which results in

$$-dU\, dV + \tfrac{1}{4}(V-U)^2 d\Omega^2 = -dT^2 + dR^2 + R^2 d\Omega^2 \quad , \tag{34.60}$$

as claimed. This establishes explicitly the conformal flatness of the Robertson-Walker metrics even for $k \neq 0$. In particular, therefore, the Weyl tensor of a Robertson-Walker metric is zero (section 11.4), and the Riemann tensor can be expressed in terms of its traces, the Ricci tensor and the Ricci scalar.

Note that for $k \neq 0$ the conformal factor relating the Robertson-Walker metric to the Minkowski metric is not just a function of the Minkowski time T alone, but a particular function of T and R,

$$k \neq 0: \qquad ds^2_{FRW} = \Omega^2(T,R)\left(-dT^2 + dR^2 + R^2 d\Omega^2\right) \quad . \tag{34.61}$$

This could hardly be otherwise because (recall the discussion in section 34.1) for $k \neq 0$ the isometry group is not the Euclidean group, the symmetry group of the Euclidean metric $dR^2 + R^2 d\Omega^2$, but rather it is $\mathsf{SO}(4)$ for $k = +1$ and $\mathsf{SO}(3,1)$ for $k = -1$.

In the latter case, we can make this symmetry manifest by slightly modifying the above procedure. To that end consider, instead of (34.57), the transformation

$$k = -1: \quad U = T - R = \mathrm{e}^{u} \quad , \quad V = T + R = \mathrm{e}^{v} \quad , \tag{34.62}$$

leading to

$$T = \mathrm{e}^{\eta}\cosh\psi \quad , \quad R = \mathrm{e}^{\eta}\sinh\psi \quad . \tag{34.63}$$

Then a straightforward calculation, using

$$R^2 = \mathrm{e}^{2\eta}\sinh^2\psi = UV\sinh^2\psi \quad , \tag{34.64}$$

shows that

$$\begin{aligned} ds^2 &= -dt^2 + a(t)^2(d\psi^2 + \sinh^2\psi \; d\Omega^2) \\ &= a(\eta)^2\left[-du\; dv + \sinh^2\psi\; d\Omega^2\right] \\ &= \frac{a(\eta)^2}{UV}\left[-dU\; dV + UV\sinh^2\psi\; d\Omega^2\right] \\ &= \mathrm{e}^{-2\eta}a(\eta)^2\left[-dT^2 + dR^2 + R^2 d\Omega^2\right] \quad . \end{aligned} \tag{34.65}$$

Here the term in brackets is just the 4-dimensional Minkowski metric,

$$-dT^2 + dR^2 + R^2 d\Omega^2 = -dT^2 + d\vec{X}^2 \quad , \tag{34.66}$$

and the conformal prefactor is only a function of η (or t) which, via

$$T^2 - R^2 = \mathrm{e}^{2\eta} \quad , \tag{34.67}$$

is itself only a function of the Lorentz invariant quantity

$$T^2 - R^2 = T^2 - \vec{X}^2 \quad . \tag{34.68}$$

Written in this way, the complete $\mathsf{SO}(3,1)$ symmetry of the $k = -1$ metrics is manifest. In the original comoving coordinates, it was realised as the spatial isometry group of the hyperboloid H^3, with t not transformed. In the new form of the metric, it is explicitly realised as the Lorentz transformations of (an auxiliary) 4-dimensional Minkowski space-time with coordinates $(T, \vec{X})$, but with the original time coordinate t or η transforming as a Lorentz scalar under these transformations, as it should.

34.7 Area Measurements and Number Counts

The aim of this and the subsequent sections is to learn as much as possible about the general properties of Robertson-Walker geometries (without using the Einstein equations) with the aim of looking for observational means of distinguishing e.g. among the models with $k = 0, \pm 1$.

To get a feeling for the geometry of the Schwarzschild metric, we studied the properties of areas and lengths in the Schwarzschild geometry. Spatial length measurements are rather obvious in the Robertson-Walker geometry, so here we focus on the properties of areas.

We write the spatial part of the Robertson-Walker metric in polar coordinates as (34.10)

$$ds^2 = a(t)^2[d\psi^2 + g_k^2(\psi)d\Omega^2] \ , \tag{34.69}$$

where $g_k(\psi) = \psi, \sin\psi, \sinh\psi$ for $k = 0, +1, -1$ (see (14.35)). Now the radius of a surface $\psi = \psi_0$ around the point $\psi = 0$ (or any other point, our space is isotropic and homogeneous) is given by

$$\rho = a\int_0^{\psi_0} d\psi = a\psi_0 \ . \tag{34.70}$$

On the other hand, the area of this surface is determined by the induced metric $a^2 g_k^2(\psi_0)d\Omega^2$ and is

$$A(\rho) = a^2 g_k^2(\psi_0)\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta = 4\pi a^2 g_k^2(\rho/a) \ . \tag{34.71}$$

For $k = 0$, this is just the standard behaviour

$$A(\rho) = 4\pi\rho^2 \ , \tag{34.72}$$

but for $k = \pm 1$ the geometry looks quite different. For $k = +1$, we have

$$A(\rho) = 4\pi a^2 \sin^2(\rho/a) \ . \tag{34.73}$$

Thus the area reaches a maximum for $\rho = \pi a/2$ (or $\psi = \pi/2$), then *decreases* again for larger values of ρ and goes to zero as $\rho \to \pi a$. Already the maximal area, $A_{max} = 4\pi a^2$ is much smaller than the area of a sphere of the same radius in Euclidean space, which would be $4\pi\rho^2 = \pi^3 a^2$.

This behaviour is best visualised by replacing the three-sphere by the two-sphere and looking at the circumference of circles as a function of their distance from the origin (see Figure 45).

For $k = -1$, we have

$$A(\rho) = 4\pi a^2 \sinh^2(\rho/a) \ , \tag{34.74}$$

so in this case the area grows much more rapidly with the radius than in flat space.

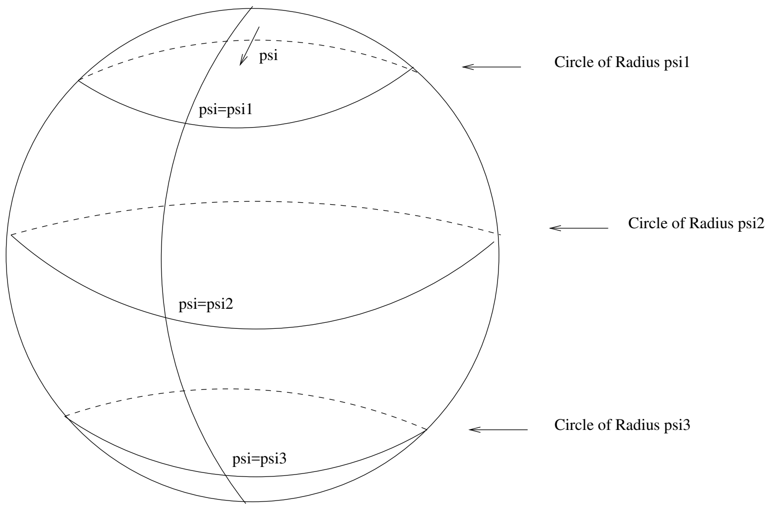


Figure 45: Visualisation of the $k = +1$ Robertson-Walker geometry via a two-sphere of unit radius: Circles of radius ψ, measured along the two-sphere, have an area which grows at first, reaches a maximum at $\psi = \pi/2$ and goes to zero when $\psi \to \pi$. E.g. the maximum value of the circumference, at $\psi = \pi/2$, namely 2π, is much smaller than the circumference of a circle with the same radius $\pi/2$ in a flat geometry, namely $2\pi \times \pi/2 = \pi^2$. Only for ψ very small does one approximately see a standard Euclidean geometry.

In principle, this distinct behaviour of areas in the models with $k = 0, \pm 1$ might allow for an empirical determination of k. For instance, one might make the assumption that there is a homogeneous distribution of the number and brightness of galaxies, and one could try to determine observationally the number of galaxies as a function of their apparent luminosity. As in the discussion of Olbers' paradox, the radiation flux would be proportional to $F \propto 1/\rho^2$. In Euclidean space ($k = 0$), one would expect the number $N(F)$ of galaxies with flux greater than F, i.e. distances less than ρ to behave like ρ^3, so that the expected Euclidean behaviour would be

$$N(F) \propto F^{-3/2} \quad . \tag{34.75}$$

Any empirical departure from this behaviour could thus be an indication of a universe with $k \neq 0$, but clearly, to decide this, many other factors (redshift, evolution of stars, etc.) would have to be taken into account. This illustrates as a matter of principle how the geometry of the spatial slices influences, and can be encoded in, observable quantities. In practice, however,

> [...] the statistical uncertainties, together with source evolution (which affects the detection probabilities), prevent this from being a useful test of k.[143]

[143] G. Ellis, R. Maartens, M. MacCallum, *Relativistic Cosmology*, section 7.6.

The most important information about the cosmic scale factor $a(t)$ comes from the observation of shifts in the frequency of light emitted by distant sources.

To calculate the expected shift in a Robertson-Walker geometry, let us again place ourselves at the origin $r = 0$. We consider a radially travelling electro-magnetic wave (a lightray) and consider the equation $d\tau^2 = 0$ or

$$dt^2 = a^2(t)\frac{dr^2}{1-kr^2} \quad . \tag{34.76}$$

Since the cosmological scale factor $a(t)$ sets the length scale, one may expect that wave lengths at different times are related by

$$\frac{\lambda(t_1)}{\lambda(t_0)} = \frac{a(t_1)}{a(t_0)} \quad , \tag{34.77}$$

leading to the relation

$$\frac{\omega(t_0)}{\omega(t_1)} = \frac{a(t_1)}{a(t_0)} \quad , \tag{34.78}$$

among the frequencies. This is indeed the correct result.

As in our discussion of the gravitational redshift in section 3.5, I will analyse this situation in two ways, in a geometric optics approach, where we trace the lightrays in the above geometry, and in a slightly more covariant language using the geodesic equation and the symmetries and associated conserved charges. I will also give a third, essentially one-line (but perhaps at first somewhat obscure looking), derivation based on the conformal Killing vector (34.51) and its associated conserved charge.

1. Let us assume that the wave leaves a galaxy located at $r = r_1$ at the time t_1. Then it will reach us at $r = 0$ at a time t_0 given by

$$f_k(r_1) = \int_{r_1}^{0} \frac{dr}{\sqrt{1-kr^2}} = \int_{t_1}^{t_0} \frac{dt}{a(t)} \quad . \tag{34.79}$$

   Note that there will only be a solution to this equation if the light from the galaxy at $r = r_1$ actually reaches us at a time t_0. In this sense galaxies whose light has not yet reached us (or may perhaps never reach us) are implicitly (and now, having said this, explicitly) excluded from the analysis - after all, such galaxies are not particularly useful for analysing redshifts.

   As typical galaxies will be comoving, i.e. have have constant spatial coordinates, $f_k(r_1)$ (34.27) is time-independent. If the next wave crest leaves the galaxy at r_1 at time $t_1 + \delta t_1$, it will arrive at a time $t_0 + \delta t_0$ determined by

$$f_k(r_1) = \int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{dt}{a(t)} \quad . \tag{34.80}$$

Subtracting these two equations and making the (eminently reasonable) assumption that the cosmic scale factor $a(t)$ does not vary significantly over the period δt given by the frequency of light, we obtain

$$\frac{\delta t_0}{a(t_0)} = \frac{\delta t_1}{a(t_1)} \quad . \tag{34.81}$$

Indeed, say that $b(t)$ is the integral of $1/a(t)$. Then we have

$$b(t_0 + \delta t_0) - b(t_1 + \delta t_1) = b(t_0) - b(t_1) \quad , \tag{34.82}$$

and Taylor expanding to first order, we obtain

$$b'(t_0)\delta t_0 = b'(t_1)\delta t_1 \quad , \tag{34.83}$$

which is the same as (34.81). Therefore the observed frequency ω_0 is related to the emitted frequency ω_1 by

$$\frac{\omega_0}{\omega_1} = \frac{a(t_1)}{a(t_0)} \quad , \tag{34.84}$$

precisely as anticipated in (34.78). I will comment further on this result below.

2. As in derivation 2 of section 3.5, we describe a lightray by the null wave vector $k^\mu = (\omega, \vec{k})$. The frequency measured by an observer with velocity u^μ is then $\omega = -u^\mu k_\mu$ (3.117).

   Adapting the discussion of timelike geodesics in the Robertson-Walker geometry in section 34.2 to null rays, we can choose the wave vector to be of the form

$$k^\mu = (\dot{t}, \dot{\psi}, 0, 0) \tag{34.85}$$

   with

$$-\dot{t}^2 + a(t)^2\dot{\psi}^2 = 0 \quad . \tag{34.86}$$

   In section 3.5 we used the timelike Killing vector of the static spherically symmetric metric, and its associated conserved energy, to relate the measured frequencies for static observers at different radial positions and to determine the gravitational redshift. Here we use one of the spatial Killing vectors to deduce, as in section 34.2, that

$$\dot{\psi} = P/a^2 \tag{34.87}$$

   for some constant P. Thus

$$\dot{t} = |P|/a \quad \Rightarrow \quad k^\alpha = (|P|/a)\,(1, \pm 1/a, 0, 0) \quad . \tag{34.88}$$

   The observers we are interested in are the comoving observers at fixed values of the spatial coordinates, i.e. with $u^\alpha = (1, 0, 0, 0)$. Thus these measure the frequency

$$\omega(t) = -k^\alpha u_\alpha = |P|/a(t) \quad . \tag{34.89}$$

In particular, for the ratio of frequencies at times t_0 and t_1 one has

$$\frac{\omega(t_0)}{\omega(t_1)} = \frac{a(t_1)}{a(t_0)} \quad , \tag{34.90}$$

in complete agreement with the result (34.78) found previously in (34.84).

3. Alternatively, and even more quickly, one can use the *conformal Killing vector* (34.51)

$$C = \partial_\eta = a(t)\partial_t \quad \text{or} \quad C^\alpha = a(t)u^\alpha \tag{34.91}$$

of the Robertson-Walker metric, and the associated conserved quantity $C_\alpha \dot{x}^\alpha$ (10.9) for null geodesics, to deduce

$$C_\alpha k^\alpha = -a(t)\omega(t) = \text{const.} \quad , \tag{34.92}$$

leading imediately to the same conclusion (34.90).

Astronomers like to express this result in terms of the redshift parameter (see the discussion of Hubble's law above)

$$z = \frac{\lambda_0 - \lambda_1}{\lambda_1} \quad , \tag{34.93}$$

which in view of the above result we can write as

$$z = \frac{a(t_0)}{a(t_1)} - 1 \quad . \tag{34.94}$$

Thus if the universe expands one has $z > 0$ and there is a redshift while in a contracting universe with $a(t_0) < a(t_1)$ the light of distant glaxies would be blueshifted.

<u>REMARKS:</u>

1. This cosmological redshift has nothing to do with the star's own gravitational field - that contribution to the redshift is completely negligible compared to the effect of the cosmological redshift.

2. Unlike the gravitational redshift we discussed before, this cosmological redshift is symmetric between receiver and emitter, i.e. light sent from the earth to the distant galaxy would likewise be redshifted if we observe a redshift of the distant galaxy.

3. However, like the gravitational redshift, the final result depends only on the position (time) of emission and arrival of the lightray, not on the intermediate gravitational field (cosmic scale factor). This illustrates that fundamentally the redshift is due to the different reference frames used by emitter and observer, not due to the fact that something happens to the lightray along the way. We will briefly return to this matter from a slightly different perspective in section 34.11.

4. While the previous remark suggests a purely Doppler-like explanation of the cosmological redshift it is best to think of the redshift as a combined effect of gravitational and Doppler redshifts. Without additional choices (like preferred families of intermediate observers) it is not very meaningful to separate this into the two and/or to interpret this only in terms of one of them.[144]

5. Nowadays, astronomers tend to express the distance of a galaxy not in terms of light-years or megaparsecs, but directly in terms of the observed redshift factor z, the conversion to distance then following from some version of Hubble's law. It is good to keep in mind that when cosmologists talk about small distances, i.e. small redshifts $z \approx 0.1$, this corresponds to a distance of approximately 1 billion light-years! The largest observed redshift of a galaxy is currently $z \approx 10$, corresponding to a distance of the order of 13 billion light-years.

6. The cosmic microwave background radiation (CMBR), which originated just a couple of 100.000 years after the Big Bang (≈ 370.000 years), has $z \gtrsim 1000$. This was the time when atoms were formed, and the CMBR photons were decoupled and emitted. This happened at a temperature of $T_{dec} \approx 3000K$. Comparing with the fact the temperature of the CMBR today is $T_{cmbr} \approx 3K$ and using $T \propto a^{-1}$ (this is essentially a reformulation of our above result for the redshift, since - up to conversion factors $\hbar$ and k - frequency = energy = temperature), one then finds the above-quoted estimate for z.[145]

34.9 Redshift - Distance Relation (Hubble's Law)

We have seen that there is a cosmological redshift in Robertson-Walker geometries. Our aim will now be to see if and how these geometries are capable of explaining Hubble's law that the redshift is approximately proportional to the distance and how the Hubble constant is related to the cosmic scale factor $a(t)$.

For a long time, reliable data for cosmological redshifts as well as for distance measurements were only available for small values of z, and thus it was common to consider the case where $t_0 - t_1$ and r_1 are small, i.e. small on cosmological scales. This allows one to find a redshift-distance relation which can e.g. be written as a power-series in z. Such a formula is not quite good enough for modern purposes, however, and I will come back to this below.

[144] For a lucid discussion of this issue, actually in the end arguing in favour of the pure Doppler interpretation, see E. Bunn, D. Hogg, *The kinematic origin of the cosmological redshift*, `arXiv:0808.1081`.

[145] For an introduction to the rich physics of the CMBR see e.g. the review articles D. Samtleben, S. Staggs, B. Winstein, *The Cosmic Microwave Background for Pedestrians: A Review for Particle and Nuclear Physicists*, `arXiv:0803.0834 [astro-ph]`, or A. Jones, A. Lasenby, *The Cosmic Microwave Background*, Living Rev. Relativity 1, (1998), 11; `http://www.livingreviews.org/lrr-1998-11`.

Assuming the validity of such an expansion, this allows us in particular to expand $a(t)$ in a Taylor series,

$$a(t) = a(t_0) + (t - t_0)\dot{a}(t_0) + \tfrac{1}{2}(t - t_0)^2\ddot{a}(t_0) + \ldots \tag{34.95}$$

Let us introduce the *Hubble parameter* $H(t)$ (which already made a brief appearance in (34.30) of section 34.3) and the *deceleration parameter* $q(t)$ by

$$\begin{aligned} H(t) &= \frac{\dot{a}(t)}{a(t)} \\ q(t) &= -\frac{a(t)\ddot{a}(t)}{\dot{a}(t)^2} \quad , \end{aligned} \tag{34.96}$$

and denote their present day values by a subscript zero, i.e. $H_0 = H(t_0)$ and $q_0 = q(t_0)$. $H(t)$ measures the expansion velocity as a function of time while $q(t)$ measures whether the expansion velocity is increasing or decreasing. We will also denote $a_0 = a(t_0)$ and $a(t_1) = a_1$. In terms of these parameters, the Taylor expansion can be written as

$$a(t) = a_0(1 + H_0(t - t_0) - \tfrac{1}{2}q_0 H_0^2(t - t_0)^2 + \ldots) \quad . \tag{34.97}$$

Higher order terms in this expansion are known as *jerk* (3rd derivative) and *snap* (4th derivative).[146]

We can use the expansion (34.97) to express z as a function of $(t_0 - t_1)$, and we can in principle use (34.79) to express r_1 as a function of $(t_0 - t_1)$. Combining the two results and eliminating $(t_0 - t_1)$, one therefore obtains the sought-for relation between the redshift z and the (coordinate-) distance r_1. The result is given in (34.104). The derivation is primarily an exercise in inverting series expansions and not *per se* particularly enlightning.

From (34.97) one finds that the redshift parameter z, as a power series in time, is

$$\frac{1}{1+z} = \frac{a_1}{a_0} = 1 + (t_1 - t_0)H_0 - \tfrac{1}{2}q_0 H_0^2(t_1 - t_0)^2 + \ldots \tag{34.98}$$

or

$$z = (t_0 - t_1)H_0 + (1 + \tfrac{1}{2}q_0)H_0^2(t_0 - t_1)^2 + \ldots \tag{34.99}$$

For small $H_0(t_0 - t_1)$ this can be inverted,

$$t_0 - t_1 = \frac{1}{H_0}[z - (1 + \tfrac{1}{2}q_0)z^2 + \ldots] \quad . \tag{34.100}$$

We can also use (34.79) to express $(t_0 - t_1)$ in terms of r_1. On the one hand we have

$$\int_{t_1}^{t_0} \frac{dt}{a(t)} = r_1 + \mathcal{O}(r_1^3) \quad , \tag{34.101}$$

[146] If you want to know why the 5th and 6th order derivative terms are known as *crackle* and *pop* respectively, look at M. Dunajski, G. Gibbons, *Cosmic Jerk, Snap and Beyond*, `arXiv:0807.0207 [gr-qc]`.

while expanding $a(t)$ in the denominator we get

$$
\begin{aligned}
\int_{t_1}^{t_0} \frac{dt}{a(t)} &= \frac{1}{a_0} \int_{t_1}^{t_0} \frac{dt}{(1+(t-t_0)H_0+\ldots)} \\
&= \frac{1}{a_0} \int_{t_1}^{t_0} dt\,[1+(t_0-t)H_0+\ldots] \\
&= \frac{1}{a_0}[(t_0-t_1)+t_0(t_0-t_1)H_0-\tfrac{1}{2}(t_0^2-t_1^2)H_0+\ldots] \\
&= \frac{1}{a_0}[(t_0-t_1)+\tfrac{1}{2}(t_0-t_1)^2H_0+\ldots] \ . \qquad (34.102)
\end{aligned}
$$

Therefore we get

$$
r_1 = \frac{1}{a_0}[(t_0-t_1)+\tfrac{1}{2}(t_0-t_1)^2H_0+\ldots] \ . \qquad (34.103)
$$

Using (34.100), we finally obtain

$$
a_0r_1 = \frac{1}{H_0}[z-\tfrac{1}{2}(1+q_0)z^2+\ldots] \ . \qquad (34.104)
$$

This clearly indicates to first order a linear dependence of the redshift on the distance of the galaxy and identifies H_0, the present day value of the Hubble parameter, as playing the role of the Hubble constant introduced in (33.6).

<u>Remarks:</u>

1. Note that the linear relation (34.29) between recessional velocity and distance of comoving objects is exact while (34.104) shows that the relation between redshift and distance is only approximately linear for small z.

2. Nowadays, cosmologists routinely deal with objects with redshifts $z > 1$. For such objects, the relation (34.104), a power-series exansion in z, is evidently not appropriate. In section 38.1 we will derive a "non-perturbative" formula for $H = H(z)$ (the value of the Hubble parameter at the time an object emitted the light that we now observe with redshift z), namely (38.8)

$$
H(z) = H_0(1+z)\left[1+(\Omega_M)_0 z+(\Omega_\Lambda)_0\left(\frac{1}{(1+z)^2}-1\right)\right]^{1/2} \ . \qquad (34.105)
$$

   Here Ω_M and Ω_Λ are the so-called *density parameters* associated to matter and a cosmological constant, the subscript 0 denoting their value today (so that $H(z)$ is expressed in terms of quantities that are in principle directly or indirectly observable).

3. Returning to the case of small z, even in that case (34.104) is not yet a very useful way of expressing Hubble's law even in that case. First of all, the distance a_0r_1 that appears in this expression is not the proper distance (unless $k=0$), but is at least equal to it in our approximation. Note that a_0r_1 is the present distance to the galaxy, not the distance at the time the light was emitted.

4. Even proper distance is not directly measurable or observable and thus, to compare this formula with experiment, one needs to relate r_1 to the measures of distance used by astronomers.

One practical way of doing this is based on the so-called *luminosity distance* d_L. If for some reasons one knows the absolute luminosity of a distant star (for instance because it shows a certain characteristic behaviour known from other stars nearby whose distances can be measured by direct means - such objects are known as *standard candles*), then one can compare this absolute luminosity L with the apparent luminosity A. Then one can define the luminosity distance d_L by (cf. (33.2))

$$d_L^2 = \frac{L}{4\pi A} \quad . \tag{34.106}$$

We thus need to relate d_L to the coordinate distance r_1. The key relation is

$$\frac{A}{L} = \frac{1}{4\pi a_0^2 r_1^2} \frac{1}{1+z} \frac{a_1}{a_0} = \frac{1}{4\pi a_0^2 r_1^2 (1+z)^2} \quad . \tag{34.107}$$

Here the first factor arises from dividing by the area of the sphere at distance $a_0 r_1$ and would be the only term in a flat geometry (see the discssion of Olbers' paradox). In a Robertson-Walker geometry, however, the photon flux will be diluted. The second factor is due to the fact that each individual photon is being redshifted. And the third factor (identical to the second) is due to the fact that as a consequence of the expansion of the universe, photons emitted a time δt apart will be measured a time $(1+z)\delta t$ apart. Hence the relation between r_1 and d_L is

$$d_L = (L/4\pi A)^{1/2} = r_1 a(t_0)(1+z) \quad . \tag{34.108}$$

Intuitively, the fact that for z positive d_L is larger than the actual (proper) distance of the galaxy can be understood by noting that the redshift makes an object look darker (further away) than it actually is.

This can be inserted into (34.104) to give an expression for the redshift in terms of d_L, *Hubble's law*

$$d_L = H_0^{-1}[z + \tfrac{1}{2}(1-q_0)z^2 + \ldots] \quad . \tag{34.109}$$

The program would then be to collect as much astronomical information as possible on the relation between d_L and z in order to determine the parameters q_0 and H_0.

34.10 Klein-Gordon Scalar Field in a Cosmological Background

The study of fields propagating in a cosmological background space-time plays an important role in cosmology, in particular when these fields arise as perturbations of the metric itself (cosmological perturbations). Here we will be content with simply deriving the equations of motion of a free massive Klein-Gordon scalar field Ψ in a spatially flat ($k=0$) Robertson-Walker background

$$ds^2 = -dt^2 + a(t)^2 d\vec{x}^2 \quad , \tag{34.110}$$

described by the standard action (6.11)

$$S[\Psi] = \int \sqrt{g} d^4x \left[-\tfrac{1}{2} g^{\alpha\beta} \partial_\alpha \Psi \partial_\beta \Psi - \tfrac{1}{2} m^2 \Psi^2 \right] \quad , \tag{34.111}$$

and very briefly qualitatively discussing some of the properties of these equations. This a continuation and variation of the theme begun in section 6.3 (general formalism for scalar fields in a gravitational field), section 7.8 (scalar field in Rindler coordinates), and section 26.8 (scalar field in the Schwarzschild space-time).

In the original coordinates $(t, \vec{x})$, the action explicitly takes the form

$$S[\Psi] = \tfrac{1}{2} \int dt\, d^3x\, a(t)^3 \left(\dot{\Psi}^2 - (\vec{\nabla}\Psi)^2 / a(t)^2 - m^2 \Psi^2 \right) \quad , \tag{34.112}$$

leading to the equation of motion

$$\ddot{\Psi} + 3(\dot{a}/a)\dot{\Psi} - (\Delta\Psi)/a^2 + m^2 \Psi = 0 \quad . \tag{34.113}$$

This exhibits a characteristic drag/friction term proportional to the Hubble parameter $H(t)$.

However, in order to make the setting as close as possible to that of a scalar field in Minkowski space (which is useful e.g. if one is intent on quantising the scalar field afterwards), it is useful to employ the conformal time coordinate η, already introduced in (34.48) and defined by

$$\begin{aligned} ds^2 &= -dt^2 + a(t)^2 d\vec{x}^2 = a(t)^2 (-dt^2/a(t)^2 + d\vec{x}^2) \\ &\equiv A(\eta)^2 (-d\eta^2 + d\vec{x}^2) \quad , \end{aligned} \tag{34.114}$$

with $A(\eta(t)) = a(t)$ and $d/d\eta = a(d/dt)$. Denoting an η-derivative by a prime, $d\Psi/d\eta = \Psi'$, the action can then more explicitly be written as

$$S[\Psi] = \tfrac{1}{2} \int d^3x\, d\eta\, A(\eta)^2 \left[(\Psi')^2 - (\vec{\nabla}\Psi)^2 - m^2 A(\eta)^2 \Psi^2 \right] \quad . \tag{34.115}$$

We see that the field Ψ has a non-canonical kinetic term $\sim A(\eta)^2 (\Psi')^2$, leading to Euler-Lagrange equations of motion containing “friction” terms,

$$\tfrac{d}{d\eta}(A^2 \Psi') = A^2 (\Psi'' + 2(A'/A)\Psi') \quad , \tag{34.116}$$

which are awkward (in the classical theory, but even more so for quantisation). Happily, these non-canonical terms can be eliminated by the field redefinition

$$\phi(\eta, \vec{x}) = A(\eta) \Psi(\eta, \vec{x}) \quad . \tag{34.117}$$

Indeed, up to a total derivative term one then finds

$$S[\Psi] \to S[\phi] = \tfrac{1}{2}\int d^3x\; d\eta\; \Big[(\phi')^2 - (\vec{\nabla}\phi)^2 - (m^2A^2 - A''/A)\phi^2\Big] \quad . \tag{34.118}$$

<u>Remarks:</u>

1. Observe that $S[\phi]$ is the standard action for a Klein-Gordon field in Minkowski space, its only mildly exotic feature being the time-dependent mass term with an effective mass
$$m^2_{\text{eff}}(\eta) = m^2A(\eta)^2 - A''(\eta)/A(\eta) \quad . \tag{34.119}$$
Thus the interaction of the scalar field ϕ with the gravitational background is entirely encoded in this time-dependent mass term. Note that this effective mass term is even present when the original field is massless, $m^2 = 0$.

2. This purely geometric contribution to the mass term can be interpreted as an induced non-minimal coupling to the scalar curvature R of the space-time. Indeed, we have
$$A''/A = (1/a)(a\frac{d}{dt})(a\frac{d}{dt})a = \frac{d}{dt}(a\dot{a}) = a\ddot{a} + \dot{a}^2 \quad . \tag{34.120}$$
Comparison with the result (35.7) for the Ricci scalar of the Robertson-Walker metric for $k = 0$ shows that
$$A''/A = Ra^2/6 = RA^2/6 \quad , \tag{34.121}$$
so that the effective mass can also be written as
$$m^2_{\text{eff}}(\eta) = A(\eta)^2(m^2 - R(\eta)/6) \quad . \tag{34.122}$$

3. This non-minimal coupling to the scalar curvature, and the factor $1/6$, are (or should be) reminiscent of the *conformal coupling* $\xi R\phi^2$ of a scalar field discussed in section 22.3. If instead of with the action (34.111) we start off with the non-minimally coupled action (22.102)
$$S_\xi[\Psi] = -\tfrac{1}{2}\int \sqrt{g}d^4x\; \Big(g^{\alpha\beta}\partial_\alpha\Psi\partial_\beta\Psi + \xi R\Psi^2\Big) \quad , \tag{34.123}$$
then instead of (34.118) we will find the action
$$S_\xi[\phi] = \tfrac{1}{2}\int d^3x\; d\eta\; \Big[(\phi')^2 - (\vec{\nabla}\phi)^2 + A(\eta)^2(\xi - 1/6)R(\eta)\phi^2\Big] \quad . \tag{34.124}$$
In particular, when ξ takes the value (22.103)
$$\xi = \frac{D-2}{4(D-1)} = +1/6 \quad , \tag{34.125}$$

for a conformal coupling, the action

$$S_{\xi=1/6}[\phi] = \tfrac{1}{2}\int d^3x\, d\eta\, \left[(\phi')^2 - (\vec{\nabla}\phi)^2\right] \tag{34.126}$$

is simply the action of a free massless scalar field in Minkowski space. This is as it should be: the action with $\xi = 1/6$ (and $m^2 = 0$) is conformally invariant, and the Robertson-Walker metric is conformally flat (this is manifest in (34.114)). Thus the action with this value of ξ must reduce to the free action in Minkowski space, the rescaling (34.117) of the scalar field reflecting the non-trivial conformal weight of a scalar field in $D = 4$.

When one introduces the non-mimimal ξ-coupling in the action together with a non-zero explicit mass term, then everything goes through as above, the only difference being that the effective mass is now ξ-dependent,

$$m^2_{\text{eff}}(\eta) = A(\eta)^2\left(m^2 + (\xi - 1/6)R(\eta)\right) \quad . \tag{34.127}$$

4. The equations of motion are

$$\phi''(\eta,\vec{x}) - \Delta\phi(\eta,\vec{x}) + m^2_{\text{eff}}(\eta)\phi(\eta,\vec{x}) = 0 \quad . \tag{34.128}$$

Spatial flatness $k = 0$ brings with it the simplifying feature that we can expand the spatial dependence of the fields in standard Fourier modes. Upon spatial Fourier expansion,

$$\phi(\eta,\vec{x}) \sim \int d^3k\; \phi_{\vec{k}}(\eta)\; \mathrm{e}^{i\vec{k}.\vec{x}} \quad , \tag{34.129}$$

one sees that each mode $\phi_{\vec{k}}$ satisfies the equation

$$\phi''_{\vec{k}}(\eta) + \omega_k^2(\eta)\phi_{\vec{k}}(\eta) = 0 \tag{34.130}$$

of a time-dependent harmonic oscillator, with time-dependent frequency

$$\omega_k^2(\eta) = m^2_{\text{eff}}(\eta) + k^2 \quad . \tag{34.131}$$

5. The crucial feature of this action and the mode equations are their explicit time-dependence which means that the energy of ϕ is not conserved. This in turn will lead to the important phenomena of particle or mode production in a cosmological background.

6. For example, one can consider the (evidently highly idealised) situation where the cosmic scale factor is asymptotically constant in the remote past and in the remote future. During these early and late periods the metric is essentially the Minkowski metric (possibly up to a rescaling of the coordinates), and one thus has a preferred notion of particles during those eras, uniquely determined by the asymptotic Poincaré symmetry. However, these definitions of particles need

not (and will amost invariably not) agree when there is an intermediate time-dependent phase. For instance, the early time vacuum Heisenberg state would not be interpreted or seen as a vacuum by the late time observer, and this disagreement about the particle content is then interpreted as a particle production due to the time-dependent gravitational field. See the references in footnote 87 (section 27.7) for a detailed discussion of these and other related fascinating issues.

34.11 Comments on Cosmic Expansion as "Expansion of Space"

As an aside, and as a conclusion to this section, let me make some comments on (and issue a *caveat* regarding) the seductive picture of an inflating balloon as the model for an expanding universe, as depicted e.g. in Figure 44 of section 34.2.

1. For many purposes, this picture certainly provides an instructive and illuminating analogy:

   - it illustrates how an expanding universe can look the same to all comoving observers;
   - in particular, it illustrates how it is possible for everything to move away from any given (comoving) observer without that observer actually being singled out as special;
   - it shows that a spatially homogeneous expansion naturally gives rise to a universal velocity-distance relation;
   - it illustrates (or is at least meant to illustrate) that expansion of the universe is something intrinsic and does not mean expansion in and into some space into which the universe has somehow been embedded.

2. The above picture of the inflating balloon is often used to describe the expansion of the universe as "an expansion of space itself", and it is then, in view of the success of this picture, tempting to *ascribe* the behaviour of light and test particles in the universe (which are mathematically described by the geodesic equations) to such an expansion of space.

   This manner of speaking and this imagery may provide some further useful intuitive understanding for some effects in the general relativistic description of cosmology. However, unless one defines precisely what one means by this, the notion of expanding space is not without its pitfalls and as with all analogies here one runs the risk of pushing this analogy too far. In particular, the danger hides in the above word "ascribe", i.e. in the risk of confusing cause and effect, or cause and effective description. Fundamentally, there is no (new?) force that (somehow) acts on space to (somehow) make it expand, and that can therefore be invoked to "explain" (in a Newtonian way) the behaviour of particles and light in such a space-time.[147]

3. The danger of thinking of the cosmological expansion in terms of the expansion of space and a corresponding agent responsible for this is well-illustrated by the cosmological redshift discussed in section 34.8, and derived there on the basis of the null geodesic equations.

   If (somehow) the expansion of space were fundamentally responsible for this effect, one would expect this effect to be cumulative and to depend on the gravitational field (or cosmic scale factor) $a(t)$ during the entire period of propagation. However, as we have seen, and as stressed at the end of section 34.8, the final result (34.94) for the redshift, $z = a(t_0)/a(t_1) - 1$, depends only on the values of the cosmological scale factor at the times t_0 and t_1, indicating that the redshift is *not* a cumulative effect due to the expansion of space while it was traversed by the lightray, but that it can (equally intuitively and perhaps more correctly) be ascribed to the fact that emitter and observer do not share the same inertial frame.

4. The issue also, and in particular, arises when it comes to frequently asked questions such as "which objects participate in the cosmic expansion?" (do you expand with the universe? does a hydrogen atom? does our solar system?) or, to use cosmologists' jargon, "which objects join the Hubble flow?", which have generated a lot of confusion over the decades. Here again the "expanding space" image may lead to a misleading intuition, in particular when space is then viewed as some kind of viscous fluid which will invariably drag other objects along with it when it expands.[148]

5. These kinds of questions have a long history in general relativity, dating back at least to an article by Einstein and Straus in 1945 entitled *The Influence of the Expansion of Space on the Gravitational Fields Surrounding the Individual Stars*. The Einstein-Straus solution, known as the *Einstein-Straus vacuole* is a space-time that is obtained by a cut-and-paste procedure from a suitable cosmological solution, removing a ball of mass M and replacing it by a Schwarzschild solution of the same mass. This is essentially an inside-out version of the Oppenheimer-Snyder collapse solution (removing a ball from Schwarzschild and replacing it

[147] I strongly recommend the lucid and engaging discussion of these issues in M. Francis, L. Barnes, J. Berian James, G. Lewis, *Expanding Space: the Root of all Evil?*, `arXiv:0707.0380 [astro-ph]`. See also A. Whiting, *The Expansion of Space: Free Particle Motion and the Cosmological Redshift*, `arXiv:astro-ph/0404095`, and J. Peacock, *A diatribe on expanding space*, `arXiv:0809.4573 [astro-ph]` for related discussions.

[148] For a clear discussion of these issues for geodesic but non-comoving particles, highlighting the necessity to define precisely and carefully what one means by "joining the Hubble flow", see L. Barnes, M. Francis, J. Berian James, G. Lewis, *Joining the Hubble Flow: Implications for Expanding Space*, `arXiv:astro-ph/0609271`.

by a contracting cosmological solution modelling the collapsing star) discussed in section 29. The Einstein-Straus procedure can also be applied multiple times around various "centers" and can then be used to model inhomogeneities in an otherwise homogeneous universe (and in this context the model is then known to cosmologists as the *Swiss cheese model*).

6. Since the work of Einstein and Straus, a lot of work has gone into finding exact solutions of the Einstein equations that describe gravitational objects like stars or black holes somehow "embedded" into cosmological backgrounds. On the basis of such exact solutions one can then (try to) answer the question if a given bound object takes part in the cosmic expansion or not, and try to to develop some intuition for this issue that complements intuition coming from more Newtonian considerations.

   There is a common folklore statement or rule of thumb to the effect that "gravitationally (or otherwise) bound systems do not expand with the universe", and while this statement undoubtedly has a certain validity it requires a more precise formulation to decide if or when such a statement is not only true but also has some non-trivial content.

7. The most prominent class of solutions among these hybrid star-cosmology solutions (apart from black holes in (anti-)de Sitter space, the Schwarzschild (anti-)de Sitter metrics (30.7)) is the so-called McVittie solution, found already in 1933. It consists of a crude superposition of a (in the simplest case $k=0$) Robertson-Walker metric,

$$ds^2 = -dt^2 + a(t)^2 d\vec{x}^2 \quad , \tag{34.132}$$

   with the Schwarzschild metric in isotropic coordinates (24.46)

$$ds^2 = -\frac{(1-\frac{m}{2|\vec{x}|})^2}{(1+\frac{m}{2|\vec{x}|})^2}dt^2 + (1+\frac{m}{2|\vec{x}|})^4\, d\vec{x}^2 \quad , \tag{34.133}$$

   and thus has the form

$$ds^2 = -\frac{(1-\frac{m}{2a(t)|\vec{x}|})^2}{(1+\frac{m}{2a(t)|\vec{x}|})^2}dt^2 + (1+\frac{m}{2a(t)|\vec{x}|})^4\, a(t)^2\, d\vec{x}^2 \quad . \tag{34.134}$$

   While this metric is easy to write down, it leads to a somewhat peculiar energy-momentum tensor, and therefore its physical interpretation and significance are somewhat obscure. These issues, as well as aspects of the global structure of the McVittie space-time, continue to be debated in the literature to this day.[149]

[149] For a pedagogical review of these topics and a discussion of the literature see e.g. M. Carrera, D. Giulini, *On the influence of the global cosmological expansion on the local dynamics in the Solar System*, `arXiv:gr-qc/0602098`, *Influence of global cosmological expansion on local dynamics and kinematics*, `arXiv:0810.2712 [gr-qc]`. For more on the on-going McVittie debate, see also N. Kaloper, M. Kleban, D. Martin, *McVittie's Legacy: Black Holes in an Expanding Universe* and references therein and thereto.

35 Cosmology III: Friedmann-Lemaître-Robertson-Walker Cosmology

So far, we have only used the kinematical framework provided by the Robertson-Walker metrics and we never used the Einstein equations. The benefit of this is that it allows one to deduce relations betweens observed quantities and assumptions about the universe which are valid even if the Einstein equations are not entirely correct, perhaps because of higher derivative or other quantum corrections in the early universe.

Now, on the other hand we will have to be more specific, specify the matter content and solve the Einstein equations for $a(t)$. We will see that a lot about the solutions of the Einstein equations can already be deduced from a purely qualitative analysis of these equations, without having to resort to explicit solutions (section 36). Exact solutions will then be the subject of section 37.

35.1 Curvature and Einstein Tensor of the Robertson-Walker Metric

Of course, the first thing we need to discuss solutions of the Einstein equations is the Ricci tensor of the Robertson-Walker metric. Since we already know the curvature tensor of the maximally symmetric spatial metric entering the Robertson-Walker metric (and its contractions), this is not difficult.

1. First of all, we write the Robertson-Walker metric as

$$ds^2 = -dt^2 + a(t)^2\, d\tilde{s}^2 = -dt^2 + a(t)^2\, \tilde{g}_{ij} dx^i dx^j \quad . \tag{35.1}$$

   In this section all objects with a tilde, $\tilde{}$, will refer to 3-dimensional quantities calculated with respect to the maximally symmetric metric $\tilde{g}_{ij}$.

2. One can then calculate the Christoffel symbols in terms of $a(t)$ and $\tilde{\Gamma}^i{}_{jk}$. The non-vanishing components are (we had already established that $\Gamma^\mu{}_{00} = 0$)

$$\Gamma^i{}_{jk} = \tilde{\Gamma}^i{}_{jk} \quad , \quad \Gamma^i{}_{j0} = \frac{\dot{a}}{a}\delta^i_j \quad , \quad \Gamma^0{}_{ij} = \dot{a}a\tilde{g}_{ij} \quad . \tag{35.2}$$

3. The non-zero components of the Riemann tensor are

$$\begin{aligned} R^i{}_{0j0} &= -\frac{\ddot{a}}{a}\delta^i_j \quad \Leftrightarrow \quad R^0{}_{i0j} = a\ddot{a}\tilde{g}_{ij} \\ R^k{}_{ilj} &= \tilde{R}^k{}_{ilj} + \dot{a}^2(\delta^k_l \tilde{g}_{ij} - \delta^k_j \tilde{g}_{il}) \quad . \end{aligned} \tag{35.3}$$

   where, by maximal symmetry, $\tilde{R}_{kilj}$ has the form (14.10)

$$\tilde{R}^k{}_{ilj} = k(\delta^k_l \tilde{g}_{ij} - \delta^k_j \tilde{g}_{il}) \quad . \tag{35.4}$$

4. The partial contraction of the purely spatial components of the Riemann tensor over the spatial indices is thus

$$\begin{aligned} R^k_{\ ikj} &= \tilde{R}_{ij} + 2\dot{a}^2\tilde{g}_{ij} \\ &= 2k\tilde{g}_{ij} + 2\dot{a}^2\tilde{g}_{ij} \ \ . \end{aligned} \tag{35.5}$$

5. Therefore the non-zero components of the space-time Ricci tensor are

$$\begin{aligned} R_{00} &= -3\frac{\ddot{a}}{a} \\ R_{ij} &= (a\ddot{a} + 2\dot{a}^2 + 2k)\tilde{g}_{ij} \\ &= (\frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} + \frac{2k}{a^2})g_{ij} \ \ . \end{aligned} \tag{35.6}$$

6. Thus the Ricci scalar is

$$R = \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k) \ \ , \tag{35.7}$$

7. Finally, therefore, putting everything together, we find that the Einstein tensor has the components

$$\begin{aligned} G_{00} &= 3(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}) \\ G_{0i} &= 0 \\ G_{ij} &= -(\frac{2\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2})g_{ij} \ \ . \end{aligned} \tag{35.8}$$

<u>Remarks:</u>

1. It follows on symmetry grounds (spatial maximal symmetry) alone

   - that the only potentially non-vanishing components of the Einstein tensor are G_{00} and G_{ij};
   - that $G_{ij} \sim g_{ij}$
   - and that the coefficients are only functions of t, not functions of the spatial coordinates.

   A formal proof of this is given in section 35.3. It is phrased there as a statement about the energy-momentum tensor in a Robertson-Walker metric, but the result is a general statement about the structure of spatially maximally symmetric space-time tensors. Thus in a sense the only non-trivial content of the above calculation is in the precise form of the t-dependent coefficients of G_{00} and G_{ij}.

2. We already know that in a maximally symmetric space not only can we express the Ricci tensor in terms of the Riemann tensor (namely as a contraction thereof) but we can also write the Riemann tensor algebraically in terms of the Ricci tensor (and even just in terms of the Ricci scalar), as is obvious from (35.4).

Even though the Robertson-Walker metrics are not space-time maximally symmetric, it is nevertheless true that even in this case the Riemann tensor can be expressed algebraically in terms of the Ricci tensor. Indeed, it is easy to see that the components of the Riemann tensor given in (35.3) can be written in terms of the components of the Ricci tensor in (35.6) simply as

$$\begin{aligned} R^i{}_{0j0} &= \tfrac{1}{3} R_{00} \delta^i{}_j \\ R^k_{ilj} &= \tfrac{1}{2}(\delta^k_l R_{ij} - \delta^k_j R_{il}) \ . \end{aligned} \tag{35.9}$$

On general grounds this follows from the fact, established in section 34.6, that the Robertson-Walker metrics are conformally flat so that the Weyl tensor vanishes.

3. The significance of this statement lies in the fact that it shows that a vacuum solution of the Einstein equations with spatial maximal symmetry is necessarily flat Minkowski space. This is perhaps as it should be, and at least vaguely Machian, but it is still good to have established this here once and for all since by just solving the vacuum equations one may (and will) find a solution that at first sight appears to be non-trivial, namely the *Milne universe* to be discussed in section 37.1, but which can then be shown to be just Minkowski space written in some non-inertial coordinates. The above result (35.9) shows that this had to be true.

4. Occasionally, in particular for the canonical analysis (i.e. developing the Hamiltonian formalism), it is useful to know the Ricci scalar (i.e. the Einstein-Hilbert Lagrangian) for the slightly more general metric

$$ds^2 = -N^2(t')(dt')^2 + a^2(t')d\tilde{s}^2 \ , \tag{35.10}$$

where the function $N(t')$ is known as the *lapse function* (cf. (21.18)). Instead of redoing the calculation of the scalar curvature in this case, one can simply use the change of variable

$$dt = N(t')dt' \quad \Rightarrow \quad \frac{d}{dt} = \frac{1}{N}\frac{d}{dt'} \tag{35.11}$$

to rewrite the final result (35.7) as (a prime on a or N denoting a derivative with respect to t')

$$R = \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k) = \frac{6}{a^2 N^3}(N(aa'' + (a')^2) - aa'N' + kN^3) \ . \tag{35.12}$$

We will come back to and make use of this result in section 35.8.

5. The results for the Christoffel symbols (35.2) and the Riemann tensor (35.3) are true in any dimension, i.e. for a general n-dimensional maximally symmetric space, and the first time that a dimension-dependence enters is in the factors of 2 and 3 in equations (35.5) and (35.6), which arise from taking traces. If one replaces the

spatial dimension $3 \to n$, equations (35.6) - (35.8) take the form

$$\begin{aligned}
R_{00} &= -n\frac{\ddot{a}}{a} \\
R_{ij} &= \left(\frac{\ddot{a}}{a} + (n-1)\frac{k+\dot{a}^2}{a^2}\right) g_{ij} \\
R &= 2n\frac{\ddot{a}}{a} + n(n-1)\frac{k+\dot{a}^2}{a^2} \\
G_{00} &= \frac{n(n-1)}{2}\frac{k+\dot{a}^2}{a^2} \\
G_{ij} &= -(n-1)\left(\frac{\ddot{a}}{a} + \frac{(n-2)}{2}\frac{k+\dot{a}^2}{a^2}\right) g_{ij} \quad .
\end{aligned} \tag{35.13}$$

35.2 Matter Content: A Perfect Fluid

Next we need to specify the matter content. On physical grounds one might perhaps like to argue that in the approximation underlying the cosmological principle galaxies (or clusters) should be treated as non-interacting particles or a perfect fluid (first discussed in section 7.2). As it turns out, we do not need to do this as either the symmetries of the metric or comparison with the Einstein tensor determined above fix the energy-momentum tensor to be that of a perfect fluid anyway.

In section 35.3 I will give a formal argument for this using Killing vectors. Informally we can already deduce this from the structure of the Einstein tensor obtained above. Comparing (35.8) with the Einstein equation $G_{\alpha\beta} = 8\pi G_N T_{\alpha\beta}$, we deduce that the Einstein equations can only have a solution with a Robertson-Walker metric if the energy-momentum tensor is of the form

$$\begin{aligned}
T_{00} &= \rho(t) \\
T_{0i} &= 0 \\
T_{ij} &= p(t) g_{ij} \quad ,
\end{aligned} \tag{35.14}$$

where $p(t)$ and $\rho(t)$ are some functions of time. A covariant way of writing this tensor is as

$$T_{\alpha\beta} = (p+\rho)u_\alpha u_\beta + p g_{\alpha\beta} \quad , \tag{35.15}$$

where $u^\alpha = (1,0,0,0)$ in a comoving coordinate system. This is precisely the energy-momentum tensor of a *perfect fluid* (cf. sections 7.2 and 7.5). In this context u^α is known as the *velocity field* of the fluid, and the comoving coordinates are those with respect to which the fluid is at rest. ρ is the energy-density of the perfect fluid and p is the pressure.

In general, this matter content has to be supplemented by an equation of state. This is usually assumed to be that of a *barytropic fluid*, i.e. one whose pressure depends only on its density, $p = p(\rho)$. The most useful toy-models of cosmological fluids arise from considering a linear relationship between p and ρ, of the type

$$p = w\rho \quad , \tag{35.16}$$

where w is known as the equation of state parameter. Occasionally also more exotic equations of state are considered, but the above covers a wide variety of commonly considered fluids and gases and other simple thermodynamic systems.

Consider e.g. a system whose entropy S is some function of the (internal) energy E and the (spatial) volume V,

$$S = S(E, V) \quad . \tag{35.17}$$

Then the 1st law of thermodynamics

$$TdS = dE + pdV \tag{35.18}$$

implies

$$T = \left(\frac{\partial S}{\partial E}\right)_V^{-1} \quad , \quad p = T\left(\frac{\partial S}{\partial V}\right)_E \quad , \tag{35.19}$$

and thus

$$p = w\rho = wE/V \quad \Leftrightarrow \quad V\partial_V S = wE\partial_E S \quad . \tag{35.20}$$

Thus the condition $p = w\rho$ is simply the statement that S is a function of $V^w E$,

$$p = w\rho \quad \Rightarrow \quad S = S(V^w E) \quad . \tag{35.21}$$

Here are the most common and useful special cases of the equation of state $p = w\rho$.

1. Dust

   For non-interacting particles, there is no pressure, $p = 0$, i.e. $w = 0$, the energy-momentum tensor has the simple form

   $$T_{\alpha\beta} = \rho u_\alpha u_\beta \tag{35.22}$$

   and such matter is usually referred to as *dust*,

   $$\text{dust:} \qquad p = 0 \quad \Rightarrow \quad w = 0 \quad . \tag{35.23}$$

   This is generally considered to be a good description of baryonic matter (and cold dark matter) today.

2. Radiation

   This corresponds to $w = 1/3$ (in 1+3 dimensions). One way to see this is to note that the trace of a perfect fluid energy-momentum tensor is

   $$T^\alpha_{\ \alpha} = -\rho + 3p \quad . \tag{35.24}$$

For electro-magnetic radiation, for example, the energy-momentum tensor is that of Maxwell theory and hence traceless (7.121). Therefore electromagnetic *radiation* in an FLRW universe (in particular compatibility with the symmetries implies neglecting all anisotropies) has the equation of state

$$\text{radiation:} \qquad p = \rho/3 \quad \Rightarrow \quad w = 1/3 \ . \tag{35.25}$$

Alternatively, this can be deduced from familiar statements about the thermodynamics of electromagnetic radiation (i.e. the photon gas). E.g. $S \sim E/T$ and $\rho \sim T^4$ imply

$$S \sim E/T \sim E(E/V)^{-1/4} = E^{3/4}V^{1/4} = (V^{1/3}E)^{3/4} \ , \tag{35.26}$$

which (by (35.21)) also implies $w = 1/3$.

As an aside, note that one generalisation of this equation of state for radiation in $d = 3$ spatial dimensions to general spatial dimension d is a perfect fluid with a traceless energy-momentum tensor (describing what one might call a "Weyl invariant" or "conformal" fluid - cf. the discussion in section 7.7). Thus the energy-momentum tensor has to satisfy

$$g^{\alpha\beta}T_{\alpha\beta} = T^{\alpha}_{\ \alpha} = -(\rho + p) + (d+1)p = 0 \ , \tag{35.27}$$

leading to the equation of state

$$\text{conformal fluid:} \qquad p = \rho/d \quad \Rightarrow \quad w = 1/d \ . \tag{35.28}$$

and

$$T_{\alpha\beta} = \frac{\rho}{d}(g_{\alpha\beta} + (d+1)u_\alpha u_\beta) \ . \tag{35.29}$$

3. Cosmological Constant

A cosmological constant Λ, on the other hand, corresponds, as we will see, to a matter contribution with $p = -\rho$, i.e. $w = -1$,

$$\text{cosmological constant:} \qquad p = -\rho \quad \Rightarrow \quad w = -1 \ . \tag{35.30}$$

Thus either ρ is negative or p is negative.

To see the relation between a $w = -1$ perfect fluid and a cosmological constant, note that the Einstein equations with a cosmological constant (19.46) give a contribution to the energy-momentum tensor proportional to $g_{\mu\nu}$. Comparing this with the covariant form (35.15) of the energy-momentum tensor, one deduces that a cosmological constant Λ is tantamount to adding matter with $p = -\rho$. Specifically, one has

$$\begin{aligned} G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \quad &\Leftrightarrow \quad G_{\mu\nu} = 8\pi G_N \left(T_{\mu\nu} - (\Lambda/8\pi G_N) g_{\mu\nu}\right) \\ &\Rightarrow \quad T^{\Lambda}_{\mu\nu} = -\frac{\Lambda}{8\pi G_N} g_{\mu\nu} \\ &\Rightarrow \quad T^{\Lambda}_{00} \equiv \rho_\Lambda = \frac{\Lambda}{8\pi G_N} = -p_\Lambda \ . \end{aligned} \tag{35.31}$$

This is the identification anticipated in (19.50).

Remarks:

1. In section 22.1 we had introduced various *energy conditions*, the *null energy condition* (NEC), the *weak energy condition* (WEC), the *dominant energy condition* (DEC), and the *strong energy condition* (SEC), and had also analysed their implications for a perfect fluid energy-momentum tensor. In particular, the conditions that we found were

$$\begin{aligned}\text{NEC:}\quad & \rho + p \geq 0 \\ \text{WEC:}\quad & \rho \geq 0 \quad , \quad \rho + p \geq 0 \\ \text{DEC:}\quad & \rho \geq |p| \\ \text{SEC:}\quad & \rho + p \geq 0 \quad , \quad \rho + 3p \geq 0 \ .\end{aligned} \tag{35.32}$$

   With the equation of state $p = w\rho$ these energy conditions can now be written as conditions on w. For physical (gravitating instead of anti-gravitating) matter one usually requires at least the condition $\rho > 0$ (positive energy density). With this condition, the NEC and the WEC are equivalent and require

$$\rho > 0 \quad \text{and NEC/WEC} \quad \Rightarrow \quad w \geq -1 \ , \tag{35.33}$$

   while the SEC requires (the 2nd condition is then stronger than the 1st)

$$\rho > 0 \quad \text{and SEC} \quad \Rightarrow \quad w \geq -1/3 \ . \tag{35.34}$$

   Finally, the DEC implies $\rho > 0$ as well as

$$\text{DEC} \quad \Rightarrow \quad |w| \leq 1 \ . \tag{35.35}$$

   Some of the conclusions about the qualitative behaviour of the solutions to the Einstein equations in section 36 rely on the strict validity of the SEC, i.e. on the assumption that (at least in the era of interest) the matter content of the universe is dominated by stuff with $w > -1/3$.

   On the other hand, as we had seen, a cosmological constant has $w = -1$, and thus either ρ is negative or p is negative. Therefore this violates either $\rho > 0$ (the WEC and the DEC) or $\rho + 3p > 0$ (the SEC).

2. The equation of state parameter need not necessarily be a constant. Consider for instance a scalar field ϕ. Such a field will respect the symmetries of a Robertson-Walker metric (and hence can potentially give rise to a solution of the Einstein equations of the Robertson-Walker form) if it depends only on t and not on the spatial coordinates, $\phi = \phi(t)$. For such a scalar field, the energy-momentum tensor (7.75),

$$T_{\alpha\beta} = \partial_\alpha\phi\partial_\beta\phi - \tfrac{1}{2}g_{\alpha\beta}\, g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - g_{\alpha\beta}V(\phi) \ , \tag{35.36}$$

reduces to

$$T_{00} = \tfrac{1}{2}\dot{\phi}^2 + V(\phi) \quad , \quad T_{ij} = \left(\tfrac{1}{2}\dot{\phi}^2 - V(\phi)\right) g_{ij} \ , \tag{35.37}$$

which is thus indeed of the general perfect fluid form (35.14). The energy and pressure density are

$$\rho(t) = \tfrac{1}{2}\dot{\phi}^2 + V(\phi) \quad , \quad p(t) = \tfrac{1}{2}\dot{\phi}^2 - V(\phi) \ , \tag{35.38}$$

leading to the time-dependent equation of state parameter

$$w(t) = \frac{p(t)}{\rho(t)} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)} \ . \tag{35.39}$$

Note that here we are treating the scalar field as a source term for the Einstein equations. This should be contrasted with (and should not be confused with) what what we did in section 6.3, where we developed the general formalism for scalar fields in a fixed gravitational field, neglecting the backreaction of the matter fields on the gravitational field / metric.

Note also that (35.39) mimics a cosmological constant, i.e. $w = -1$, during periods (of the scalar field slowly rolling down a very flat potential) where the kinetic energy term is negligible compared with the potential energy term. This can also be seen directly from the action (6.14), a constant potential leading to a constant contribution to the matter or gravity Lagrangian, a.k.a. as a cosmological constant, and plays an important role in models of *inflation*.

3. Occasionally, more exotic equations of state are also considered in cosmology. For example, extended objects like (cosmic) strings, or membranes (domain walls) or their higher-dimensional generalisations typically have a positive tension, and therefore a negative (contribution to the) pressure.[150] In particular, the equation of state of a gas of non-relativistic strings in d spatial dimensions is

$$\text{non-relativistic string gas:} \qquad p = -\rho/d \quad \Rightarrow \quad w = -1/d \ . \tag{35.40}$$

4. Another exotic object that appears occasionally in the context of cosmology is the so-called *Chaplygin gas*, with equation of state

$$\text{Chaplygin gas:} \qquad p = -A/\rho \qquad (A > 0) \ . \tag{35.41}$$

We will briefly return to the properties of the Chaplygin gas at the end of section 35.6 - cf. (35.99) and (35.100), but for the most part we will concentrate on the linear equations of state $p = w\rho$ in the following.

[150] The equation of state for such extended objects, affectionately and collectively known as "p-branes" or just "branes", is derived e.g. in T. Boehm, R. Brandenberger, *On T-Duality in Brane Gas Cosmology*, `arXiv:hep-th/0208188`.

5. A null variant of the dust energy-momentum tensor (35.22), with u^α replaced by a null vector k^α (*null dust*) will appear as the source term for a toy model of a radiating star, the *Vaidya metric*,

$$T_{\alpha\beta} = \rho k_\alpha k_\beta \tag{35.42}$$

to be discussed in sections 40 - 42 (cf. in particular section 40.4). Since this energy-momentum tensor is manifestly traceless,

$$T^\alpha_{\ \alpha} = \rho k^\alpha k_\alpha = 0 \ , \tag{35.43}$$

it can alternatively be regarded as a variant of the energy-momentum tensor for radiation (also with $\rho = 3p$).

35.3 Appendix: Space-Time Tensors with Maximal Spatial Symmetry

Here is the formal argument that the energy-momentum tensor necessarily has the form given in (35.14).[151]

It is of course a consequence of the Einstein equations that any symmetries of the Ricci (or Einstein) tensor also have to be symmetries of the energy-momentum tensor. Now we know that the metric $\tilde{g}_{ij}$ has six Killing vectors $K^{(a)}$ and that (in the comoving coordinate system) these are also Killing vectors of the Robertson-Walker metric,

$$L_{K^{(a)}}\tilde{g}_{ij} = 0 \ \ \Rightarrow \ \ L_{K^{(a)}}g_{\mu\nu} = 0 \ . \tag{35.44}$$

Therefore also the Ricci and Einstein tensors have these symmetries,

$$L_{K^{(a)}}g_{\mu\nu} = 0 \ \ \Rightarrow \ \ L_{K^{(a)}}R_{\mu\nu} = 0 \ , \ \ L_{K^{(a)}}G_{\mu\nu} = 0 \ . \tag{35.45}$$

To prove this one can either (non-covariantly) choose, for each Killing vector, an adapted coordinate system, or one generalises the argument given in section 13.4 for the Ricci scalar, $L_K R = 0$, to the Ricci tensor.

The Einstein equations then imply that $T_{\mu\nu}$ should have these symmetries,

$$L_{K^{(a)}}G_{\mu\nu} = 0 \ \ \Rightarrow \ \ L_{K^{(a)}}T_{\mu\nu} = 0 \ . \tag{35.46}$$

Moreover, since the $L_{K^{(a)}}$ act like three-dimensional coordinate transformations, in order to see what these conditions mean we can make a $(3+1)$-decomposition of the energy-momentum tensor. From the three-dimensional point of view, T_{00} transforms like a scalar under coordinate transformations (and Lie derivatives), T_{0i} like a vector, and T_{ij} like a symmetric tensor. Thus we need to determine what are the three-dimensional

[151] This elementary argument, which requires no "higher knowledge" about Lie groups and invariants, is adapted from the discussion in section 13.4 of S. Weinberg, *Gravitation and Cosmology*.

scalars, vectors and symmetric tensors that are invariant under the full six-parameter group of the three-dimensional isometries.

For scalars ϕ we thus require (calling K now any one of the Killing vectors of $\tilde{g}_{ij}$),

$$L_K\phi = K^i\partial_i\phi = 0 \quad . \tag{35.47}$$

Since $K^i(x)$ can take any value in a maximally symmetric space (homogeneity), this implies that ϕ has to be constant (as a function on the three-dimensional space) and therefore T_{00} can only be a function of time,

$$T_{00} = \rho(t) \quad . \tag{35.48}$$

For vectors, it is almost obvious that no invariant vectors can exist because any vector would single out a particular direction and therefore spoil isotropy. The formal argument (as a warm up for the argument for tensors) is the following. We have

$$L_K V^i = K^j\tilde{\nabla}_j V^i + V^j\tilde{\nabla}_j K^i \quad . \tag{35.49}$$

We now choose the Killing vectors such that $K^i(x) = 0$ but $\tilde{\nabla}_i K_j \equiv K_{ij}$ is an arbitrary anti-symmetric matrix. Then the first term disappears and we have

$$L_K V^i = 0 \;\;\Rightarrow\;\; K_{ij}V^j = 0 \quad . \tag{35.50}$$

To make the anti-symmetry manifest, we rewrite this as

$$K_{ij}V^j = K_{kj}\delta^k_i V^j = \tfrac{1}{2}K_{kj}(\delta^k_i V^j - \delta^j_i V^k) = 0 \quad . \tag{35.51}$$

If this is to hold for all anti-symmetric matrices, we must have

$$\delta^k_i V^j = \delta^j_i V^k \quad , \tag{35.52}$$

and by contraction one obtains $nV^j = V^j$, and hence $V_j = 0$. Therefore, as expected, there is no invariant vector field and

$$T_{0i} = 0 \quad . \tag{35.53}$$

We now come to symmetric tensors. Once again we choose our Killing vectors to vanish at a given point x and such that K_{ij} is an arbitrary anti-symmetric matrix. Then the condition

$$L_K T_{ij} = K^k\tilde{\nabla}_k T_{ij} + \tilde{\nabla}_i K^k T_{kj} + \tilde{\nabla}_j K^k T_{ik} = 0 \tag{35.54}$$

reduces to

$$K_{mn}(\tilde{g}^{mk}\delta^n_i T_{kj} + \tilde{g}^{mk}\delta^n_j T_{ik}) = 0 \quad . \tag{35.55}$$

If this is to hold for all anti-symmetric matrices K_{mn}, the anti-symmetric part of the term in brackets must be zero or, in other words, it must be symmetric in the indices m and n, i.e.

$$\tilde{g}^{mk}\delta^n_i T_{kj} + \tilde{g}^{mk}\delta^n_j T_{ik} = \tilde{g}^{nk}\delta^m_i T_{kj} + \tilde{g}^{nk}\delta^m_j T_{ik} \quad . \tag{35.56}$$

Contracting over the indices n and i, one obtains

$$n\tilde{g}^{mk}T_{kj} + \tilde{g}^{mk}T_{jk} = \tilde{g}^{mk}T_{kj} + \delta^m_j T^k_k \quad . \tag{35.57}$$

Therefore

$$T_{ij} = \frac{\tilde{g}_{ij}}{n} T^k_k \quad . \tag{35.58}$$

Now we already know that the scalar T^k_k has to be a constant. Thus we conclude that the only invariant tensor is the metric itself, and therefore the T_{ij}-components of the energy-momentum tensor can only be a function of t times $\tilde{g}_{ij}$. Writing this function as $p(t)a^2(t)$, we arrive at

$$T_{ij} = p(t)g_{ij} \quad . \tag{35.59}$$

We thus see that the energy-momentum tensor is determined by two functions, $\rho(t)$ and $p(t)$, precisely as in (35.14).

35.4 Conservation Laws for Perfect Fluids

The same arguments as above show that a current J^μ in a Robertson-Walker metric has to be of the form $J^\mu = (n(t), 0, 0, 0)$ in comoving coordinates, or

$$J^\mu = n(t)u^\mu \tag{35.60}$$

in covariant form. Here $n(t)$ could be a number density like a galaxy number density. It gives the number density per unit proper volume. The conservation law $\nabla_\mu J^\mu = 0$ is equivalent to

$$\nabla_\mu J^\mu = 0 \quad \Leftrightarrow \quad \partial_t(\sqrt{g}n(t)) = 0 \quad . \tag{35.61}$$

Thus we see that $n(t)$ is not constant, but the number density per unit coordinate volume is (as we had already anticipated in the picture of the balloon, Figure 44). For a Robertson-Walker metric, the time-dependent part of $\sqrt{g}$ is $a(t)^3$, and thus the conservation law says

$$n(t)a(t)^3 = const. \tag{35.62}$$

Let us now turn to the conservation laws associated with the energy-momentum tensor,

$$\nabla_\mu T^{\mu\nu} = 0 \quad . \tag{35.63}$$

The spatial components of this conservation law,

$$\nabla_\mu T^{\mu i} = 0 \quad , \tag{35.64}$$

turn out to be identically satisfied, by virtue of the fact that the u^μ are geodesic and that the functions ρ and p are only functions of time.

This could hardly be otherwise because $\nabla_\mu T^{\mu i}$ transforms as a vector under spatial rotations, and rotational invariance (isotropy) now implies that this vector has to be zero identically. Nevertheless it is of course instructive to check this explicitly.

The only interesting conservation law is thus the zero-component

$$\nabla_\mu T^{\mu 0} = \partial_\mu T^{\mu 0} + \Gamma^\mu_{\ \mu\nu} T^{\nu 0} + \Gamma^0_{\ \mu\nu} T^{\mu\nu} = 0 \quad , \tag{35.65}$$

which for a perfect fluid with $T_{00} = \rho(t)$ and $T_{ij} = p(t) g_{ij}$ becomes

$$\partial_t \rho(t) + \Gamma^\mu_{\ \mu 0} \rho(t) + \Gamma^0_{\ 00} \rho(t) + \Gamma^0_{\ ij} T^{ij} = 0 \quad . \tag{35.66}$$

Inserting the explicit expressions (35.2) for the Christoffel symbols, one finds

$$\dot\rho = -3(\rho + p)\frac{\dot a}{a} \quad . \tag{35.67}$$

Introducing some fixed comoving volume $\sim v$ and its associated proper volume

$$V(t) = a(t)^3 v \quad , \tag{35.68}$$

and noting that

$$\frac{dV(t)}{dt} = 3a(t)^2 \dot a(t) v = 3H(t) V \quad , \tag{35.69}$$

the conservation equation (35.67) can be written in a perhaps more suggestive and familiar (mechanical or thermodynamical) form as

$$\frac{dE}{dt} = -p\frac{dV}{dt} \tag{35.70}$$

where

$$E(t) = \rho(t) V(t) \tag{35.71}$$

is the total energy in the volume V. Comparing with the 1st law of thermodynamics (35.18), this equation thus encodes the statement that the time evolution of the perfect fluid is *adiabatic* in the sense that its entropy remains constant,

$$\frac{dS(t)}{dt} = 0 \quad . \tag{35.72}$$

35.5 Conservation Laws and Comoving Congruences

Before discussing some special cases of the solutions of (35.67) in section 35.6 below, it is instructive to rederive the above results in a somewhat more general and covariant manner.[152] Thus we consider a general velocity field $u^\mu(x)$ with $u^\mu u_\mu = -1$, and the perfect fluid energy-momentum tensor (35.15),

$$T_{\mu\nu} = (p + \rho) u_\mu u_\nu + p g_{\mu\nu} \quad , \tag{35.73}$$

[152] For a detailed exposition of cosmology in this covariant framework see e.g. G. Ellis, H. van Elst, *Cosmological Models* (Cargèse Lectures 1998), `arXiv:gr-qc/9812046`, or the even more detailed monograph by G. Ellis, R. Maartens, M. MacCallum, *Relativistic Cosmology*.

with for the time being ρ and p arbitrary functions of the space-time coordinates. Note that $u^\mu\nabla_\mu$ is the covariant derivative along the integral curves of u^μ, the object we denoted D_τ or $D/D\tau$ in section 5.7. Acting on scalars we will simply denote it, as usual, by an overdot, i.e.

$$u^\mu\nabla_\mu\rho \equiv \dot{\rho} \tag{35.74}$$

etc. Let us now see what the conditon $\nabla_\mu T^{\mu\nu} = 0$ (which has to hold if this energy-momentum tensor is to give us a solution to the Einstein equations) tells us.

1. We first consider the case $p = 0$, so this corresponds to a pressure free perfect fluid (and is used in cosmology to e.g. model cold dark or baryonic matter). Then one has

$$T_{\mu\nu} = \rho u_\mu u_\nu \quad \Rightarrow \quad \nabla_\mu T^{\mu\nu} = (\dot{\rho} + \rho\nabla_\mu u^\mu)u^\nu + \rho u^\mu\nabla_\mu u^\nu \ \ . \tag{35.75}$$

Here $\nabla_\mu u^\mu \equiv \theta$ is (and measures) the *expansion* of the velocity field u^μ (and was introduced previously, in the context of the Raychaudhuri equation, in (12.34)), and the last term $u^\mu\nabla_\mu u^\nu \equiv a^\nu$ is its acceleration (5.99), so that we can also write this equation as

$$(\dot{\rho} + \theta\rho)u^\nu + \rho a^\nu = 0 \ \ . \tag{35.76}$$

Since u^ν and a^ν are orthogonal to each other,

$$u_\mu u^\mu = -1 \quad \Rightarrow \quad u_\mu a^\mu = 0 \ \ , \tag{35.77}$$

this equation breaks up into two independent pieces,

$$\nabla_\mu T^{\mu\nu} = 0 \quad \Leftrightarrow \quad \dot{\rho} + \theta\rho = 0 \quad \text{and} \quad a^\nu = u^\mu\nabla_\mu u^\nu = 0 \ \ . \tag{35.78}$$

Its time (energy flow) component is a continuity equation, while its space (momentum flow) part tells us that the particles have to move on geodesics.

2. Now what happens if we include pressure p? This corresponds to adding $p(g^{\mu\nu} + u^\mu u^\nu) \equiv ph^{\mu\nu}$, but this tensor is orthogonal to u^μ (cf. again the discussion in section 12.2 on the Raychaudhuri equation),

$$u^\mu h_{\mu\nu} \equiv u^\mu(g_{\mu\nu} + u_\mu u_\nu) = u_\nu - u_\nu = 0 \ \ . \tag{35.79}$$

Therefore the equation $\nabla_\mu T^{\mu\nu} = 0$, with

$$T_{\mu\nu} = (\rho+p)u_\mu u_\nu + pg_{\mu\nu} \quad \Rightarrow \quad \nabla_\mu T^{\mu\nu} = (\dot{\rho}+\theta(\rho+p))u^\nu + (\rho+p)a^\nu + (\nabla_\mu p)h^{\mu\nu} \tag{35.80}$$

again breaks up nicely into two orthogonal pieces. The part tangent to u^μ tells us that

$$\dot{\rho} + \theta(\rho + p) = 0 \ \ , \tag{35.81}$$

so this is a conservation law, and the part orthogonal to u^μ gives

$$(\rho + p)a^\nu + (\nabla_\mu p)h^{\mu\nu} = 0 \quad , \tag{35.82}$$

which is a curved-space relativstic generalisation of the Euler equations of a perfect fluid.

In particular, now the velocity field is not composed of geodesics unless the derivative of p in the directions orthogonal to u^μ (i.e. the spatial derivative) is zero, $h^{\nu\mu}\nabla_\mu p = 0$. This is precisely the situation we are considering in cosmology, where $\rho = \rho(t)$ and $p = p(t)$ depend only on t, which is the proper time of the comoving observers described by the velocity field u^μ.

Returning thus to the cosmological setting, where we have (correctly, and uniquely as we now know) chosen the matter to move along geodesics, we are left with the continuity equation (35.81), which is now the same as (35.67) because for $u^\mu = (1, 0, 0, 0)$ in comoving coordinates one has

$$\theta = \nabla_\mu u^\mu = \frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}u^\mu) = a(t)^{-3}\partial_t(a(t)^3) = 3\dot{a}(t)/a(t) \quad . \tag{35.83}$$

As an aside, note that this equation implies

$$\frac{d}{dt}\theta = 3(\ddot{a}/a - \dot{a}^2/a^2) \quad , \tag{35.84}$$

which in terms of the Hubble parameter $H = \dot{a}/a$ and the deceleration parameter $q = -\ddot{a}a/\dot{a}^2$ (34.96) can be written as

$$\frac{d}{dt}\theta = -3H^2(q+1) \quad . \tag{35.85}$$

This equation is a special case of the Raychaudhuri equation (12.36) for timelike geodesic congruences,

$$\frac{d}{d\tau}\theta = -\tfrac{1}{3}\theta^2 - \sigma^{\mu\nu}\sigma_{\mu\nu} + \omega^{\mu\nu}\omega_{\mu\nu} - R_{\mu\nu}u^\mu u^\nu \quad . \tag{35.86}$$

Indeed, specialising (35.86) to the family of comoving observers in a Robertson-Walker geometry and noting that

- the proper time τ is the cosmological time t
- the rotation is zero (either by explicit calculation or, more ot the point, because the geodesics are orthogonal to the hypersurfaces $t = const.$, or on symmetry grounds)
- the shear is zero (either by explicit calculation or on symmetry grounds)
- the relevant component of the Ricci tensor is $R_{00} = -3\ddot{a}/a$,

oner sees that (35.86) reduces to

$$\frac{d}{dt}\theta = -\tfrac{1}{3}\theta^2 - R_{00} = -3\dot{a}^2/a^2 + 3\ddot{a}/a \tag{35.87}$$

which is identical to (35.84). Thus we see how the parameters $H(t)$ and $q(t)$, originally introduced to characterise the first terms in a Taylor expansion of the cosmic scale factor also govern the local behaviour of freely falling observers like (clusters of) galaxies. In particular, in an expanding universe the congruences of comoving observers diverge / expand ($\theta > 0$) but the rate of expansion decreases ($\dot{\theta} < 0$) provided that $q > -1$.

35.6 Conservation Laws for Specific Equations of State

We now look at the solutions of (35.67),

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a} \quad , \tag{35.88}$$

for some specific equations of state.

1. For instance, when the pressure of the cosmic matter is negligible, like in the universe today, and we can treat the galaxies (without disrespect) as dust, then one has

   $$w = 0 \quad \Rightarrow \quad \frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a} \quad , \tag{35.89}$$

   and this equation can trivially be integrated to

   $$\rho(t)a(t)^3 = \text{const.} \tag{35.90}$$

   Thus the (proper) density is proportional to the inverse (proper) spatial volume, an unsurprising (and reassuring) result. This of course also follows on the nose from the conservation law in the form (35.70):

   $$p = 0 \quad \Rightarrow \quad \frac{d}{dt}E = 0 \quad . \tag{35.91}$$

2. On the other hand, if the universe is dominated by, say, radiation, then one has the equation of state $p = \rho/3$, and the conservation equation reduces to

   $$w = 1/3 \quad \Rightarrow \quad \frac{\dot{\rho}}{\rho} = -4\frac{\dot{a}}{a} \quad , \tag{35.92}$$

   and therefore

   $$\rho(t)a(t)^4 = \text{const.} \tag{35.93}$$

   The reason why the energy density of photons decreases faster with $a(t)$ than that of dust is of course ... the redshift:

   $$\frac{1}{a(t)^4} = \frac{1}{a(t)^3}(\text{number density of photons}) \times \frac{1}{a(t)}(\text{redshift of individual photons}) \quad . \tag{35.94}$$

3. More generally, for matter with equation of state parameter w one finds

$$p = w\rho \quad \Rightarrow \quad \rho(t)a(t)^{3(1+w)} = \text{const.} \tag{35.95}$$

The interpretation of this equation becomes somewhat more transparent in terms of the thermodynamic variant (35.70) of the conservation law. For the linear equation of state $p = w\rho$, it reduces to

$$\frac{dE}{dt} = -wE\frac{dV}{dt} \quad \Leftrightarrow \quad \frac{d}{dt}(V^w E) = 0 \ . \tag{35.96}$$

By (35.21), this is simply the statement that the entropy, which is a function of $V^w E$, is constant (cf. (35.72)).

4. In particular, for $w = -1$, $\rho = E/V$ itself is constant,

$$w = -1 \quad \Rightarrow \quad \rho(t) = \text{const.} \tag{35.97}$$

as it should be, in view of its identification (35.31) with a cosmological constant.

5. Quite generally, we see from (35.88) that in an expanding universe (i.e. $\dot{a} > 0$), the energy density of matter satisfying the null energy condition (NEC) $\rho + p \geq 0$ cannot increase (and will necessarily decrease for matter satisfying the strict NEC, with $\rho + p > 0$),

$$\dot{a} > 0 \quad , \quad \rho + p \geq 0 \quad (\text{NEC}) \quad \Rightarrow \quad \dot{\rho} \leq 0 \ . \tag{35.98}$$

6. As the final example, consider the peculiar Chaplygin gas with equation of state $p = -A/\rho$ (35.41) with A constant. In this case (35.88) reads

$$p = -A/\rho \quad \Rightarrow \quad \dot{\rho} = -3\left(\rho - \frac{A}{\rho}\right)\frac{\dot{a}}{a} \ , \tag{35.99}$$

which has the exact solution[153]

$$\rho(t) = \sqrt{A + B/a(t)^6} \tag{35.100}$$

(with integration constant B). In the context of a cosmology of an expanding universe, this has the remarkable property of interpolating between what appears to be a dust-filled universe at early times (small $a(t)$),

$$a(t) \quad \text{small} \quad \Rightarrow \quad \rho(t) \sim \sqrt{B}/a(t)^3 \ , \tag{35.101}$$

and a universe filled with a cosmological constant at late times (large $a(t)$),

$$a(t) \quad \text{large} \quad \Rightarrow \quad \rho(t) \sim \sqrt{A} \quad \Rightarrow \quad p(t) \sim -\sqrt{A} \ . \tag{35.102}$$

Thus this mimics the evolution of an ordinary (matter-filled) universe to one dominated by *dark energy*, resembling the evolution of our universe (cf. section 38.1).

[153] A. Kamenshchik, U. Moschella, V. Pasquier, *An alternative to quintessence*, `arXiv:gr-qc/0103004`.

35.7 Einstein and Friedmann-Lemaître Equations

After these preliminaries, we are now prepared to tackle (hence first to determine) the Einstein equations in this setting.

Allowing for the presence of a cosmological constant, we thus consider the equations

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \quad . \tag{35.103}$$

Alternatively we can write these equations as

$$R_{\mu\nu} = 8\pi G_N (T_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} T^{\lambda}_{\lambda}) + \Lambda g_{\mu\nu} \quad . \tag{35.104}$$

Because of isotropy, there are only two independent equations, e.g. the (00)-component and any one of the non-zero (ij)-components. Using (35.6), we find

$$\begin{aligned} -3\frac{\ddot{a}}{a} &= 4\pi G_N(\rho + 3p) - \Lambda \\ \frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} + 2\frac{k}{a^2} &= 4\pi G_N(\rho - p) + \Lambda \quad . \end{aligned} \tag{35.105}$$

Alternatively, instead of the spatial components of (35.104) one could have used the (00)-component of (35.103) and the expression in (35.8) for G_{00} to deduce

$$G_{00} - \Lambda = 8\pi G_N \, T_{00} \quad \Rightarrow \quad \frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{8\pi G_N}{3}\rho + \frac{\Lambda}{3} \tag{35.106}$$

(and this equation could have also been obtained by appropriate subtraction of the previous 2 equations). Either way, we supplement these by the conservation equation

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a} \quad . \tag{35.107}$$

and thus end up with the set of equations

$$\boxed{\begin{array}{llcl} (F1) & \frac{\dot{a}^2}{a^2} + \frac{k}{a^2} & = & \frac{8\pi G_N}{3}\rho + \frac{\Lambda}{3} \\ (F2) & -3\frac{\ddot{a}}{a} & = & 4\pi G_N(\rho + 3p) - \Lambda \\ (F3) & \dot{\rho} & = & -3(\rho + p)\frac{\dot{a}}{a} \end{array}} \tag{35.108}$$

Together, these are known as the *Friedmann-Lemaître equations* or *Friedmann equations*. These equations form the basis, and govern every aspect, of the standard Friedmann-Lemaître-Robertson-Walker models of cosmology.

Remarks:

1. I will usually refer to the collection of these equations, without disrespect to Lemaître but just out of habit, to the Friedmann equations. When referring to individual equations, rather than quoting equation numbers I will simply refer to them as the equations (F1), (F2), and (F3) respectively.

2. In terms of the Hubble parameter $H(t)$ and the deceleration parameter $q(t)$, these equations can also be written as

$$\begin{array}{llcl} (F1') & H^2 & = & \frac{8\pi G_N}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \\ (F2') & q & = & \frac{1}{3H^2}[4\pi G_N(\rho + 3p) - \Lambda] \\ (F3') & \frac{d}{dt}(\rho a^3) & = & -3Hpa^3 \ . \end{array}$$

3. Commonly, (F1) is referred to as *the Friedmann equation.* As mentioned above, this is precisely the (00)-component of the Einstein equations.

4. (F2), on the other hand, can be interpreted as the Raychaudhuri equation (35.87). Indeed, with $\theta = 3\dot{a}/a$ and

$$R_{00} = 8\pi G_N(T_{00} + \tfrac{1}{2}T^\alpha_{\ \alpha}) - \Lambda = 4\pi G_N\ (\rho + 3p) - \Lambda \tag{35.109}$$

one finds that

$$\frac{d}{dt}\theta = -\tfrac{1}{3}\theta^2 - R_{00} \quad \Leftrightarrow \quad -3\ddot{a}/a = 4\pi G_N(\rho + 3p) - \Lambda \tag{35.110}$$

which is precisely (F2).

5. In writing the above equations, I have separated out the cosmological constant Λ from the remaining matter contributions. Of course, using (35.31), it could have just been treated as one other perfect fluid contribution (with $w = -1$). Occasionally either one or the other way of writing these equations is (marginally) more convenient.

6. Note that because of the Bianchi identities, the Einstein equations and the conservation equations should not be independent, and indeed they are not:

   (a) It is easy to see that (F1) and (F3) imply the second order equation (F2) so that, a pleasant simplification, in practice one only has to deal with the two first order equations (F1) and (F3). Sometimes, however, (F2) is easier to solve than (F1), because it is linear in $\ddot{a}(t)$, and then (F1) is just used to fix one constant of integration.

   (b) It is also easy to see that (F1) and (F2) imply (F3), i.e. that the gravity equations of motion imply the matter equations of motion, a general and fundamental feature of general relativity.

   (c) Finally, formally (F2) and (F3) also imply (F1), with k (which only appears in (F1)) arising as an integration constant.

7. Note also that the parameter k appears in the Friedmann equations only in the combination k/a^2. Thus a scaling of k can be compensated by a corresponding scaling of $a(t)$. This reflects the fact, mentioned several times already, that only this particular combination has individual meaning, and that one can exploit the scaling freedom to e.g. set $a_0 = 1$ or $k = 0, \pm 1$.

8. One can use (35.13) to determine the analogue of the Friedmann equations in any space-time dimension $D = n + 1$.

   - The indepedent components of the Einstein equation can then be chosen to be, from $G_{00} = \kappa T_{00}$ (for the reason for now calling the coupling constant κ and not just $8\pi G_N$, as in four dimensions, see the discussion around (19.44)),

$$\frac{k + \dot{a}^2}{a^2} = \frac{2\kappa}{n(n-1)}\rho \quad , \tag{35.111}$$

     and from R_{00}, say, using (19.45) and the fact that the trace of a perfect fluid energy-momentum tensor is now

$$T^\lambda_\lambda = -\rho + np \quad , \tag{35.112}$$

     one finds

$$\frac{\ddot{a}}{a} = -\frac{\kappa}{n(n-1)}\left[(n-2)\rho + np\right] \quad . \tag{35.113}$$

   - Finally the continuity equation takes the form

$$\dot{\rho} = -n(\rho + p)\frac{\dot{a}}{a} \quad . \tag{35.114}$$

9. As an aside note that, as the Robertson-Walker metrics are, in particular, spherically symmetric, and written in the manifestly spherically-symmetric form (24.80), we have the notion of the Misner-Sharp mass (24.82) for spherical symmetry at our disposal. In terms of the area radius (34.42),

$$\tilde{r}(t,r) = a(t)r \quad , \tag{35.115}$$

   the Misner-Sharp mass function is given by (24.82)

$$M_{MS}(t,r)m(t,r) = \tfrac{1}{2}\tilde{r}\left(1 - g^{ab}\partial_a\tilde{r}\partial_b\tilde{r}\right) \quad . \tag{35.116}$$

   With $g^{tt} = -1$ and $g^{rr} = (1 - kr^2)/a(t)^2$, this reduces to

$$m(t,r) = \tfrac{1}{2}a(t)r^3(\dot{a}(t)^2 + k) \quad . \tag{35.117}$$

   Using the Friedmann equation (F1), this can be written in a more informative way as

$$\begin{aligned} m(t,r) &= \tfrac{1}{2}a(t)^3 r^3\left(\frac{8\pi G_N}{3}\rho + \frac{\Lambda}{3}\right) \\ &= \frac{4\pi}{3}\tilde{r}^3 G_N\left(\rho + \rho_\Lambda\right) \quad , \end{aligned} \tag{35.118}$$

   with $\rho_\Lambda = \Lambda/8\pi G_N$ the energy density (35.31) associated with the cosmological constant. Note that this result, which again, as in section 24.6, has the interpretation as "mass = coordinate volume (not proper volume) $\times$ density", does not depend explicitly on either the pressure p or the curvature k. In particular it is independent of the equation of state relating p and ρ.

35.8 Lagrangian Formulation of the Friedmann Equations

For many purposes it is useful to cast the above set of Friedmann-Lemaître equations into a Lagrangian or Hamiltonian form. In particular (and this is the motivation for doing this here), this system of equations is sufficiently simple to provide a concrete illustration of some of the general features of the canonical Hamiltonian formulation of general relativity discussed in a somewhat cursory way in section 21.

Rather than specialising the general results of that section to the case at hand, we will adopt a more pedestrian approach here and derive these results from scratch. This will make contact with and hopefully shed some light on a variety of different issues that have arisen in various parts of these notes, e.g.

- the characteristic constraints of general relativity (sections 19.7 and 21), in particular the *Hamiltonian constraint*
- the *Gibbons-Hawking-York boundary term* discussed in section 20.5
- the *ADM action* discussed in section 23.4
- the role and significance of the *lapse function* (cf. (21.18)) introduced into the Robertson-Walker metric (for what appeared to be no good reason at that point) at the end of section 35.1.

In the previous section, in order to arrive at the Friedmann equations, we made an ansatz for the metric based on symmetry considerations (spatial maximal symmetry), namely

$$ds^2 = -dt^2 + a(t)^2 d\tilde{s}^2 \tag{35.119}$$

and then plugged this (together with a compatible ansatz for the energy-momentum tensor) into the full Einstein equations. Such a reduction procedure is always possible at the level of the equations of motion. In order to develop the Hamiltonian formulation of the Friedmann equations, it is convenient to start with the Lagrangian formulation. However, such a reduction procedure (plugging an ansatz for the solution into the action and then solving the resulting Euler-Lagrange equations) is not necessarily consistent at the level of the Lagrangian or action, i.e. the Euler-Lagrange equations of the reduced Langrangian will not necessarily agree with the reduced equations of motion and will not necessarily give rise to solutions of the original unreduced equations.

Indeed, in the present context it is pretty obvious that, with the single gravitational degree of freedom $a(t)$, associated with the size of the spatial metric, it is impossible to derive both the (one independent) spatial component of the Einstein equations (the Friedmann equation F2, say) and the time-time component of the Einstein equations (the Friedmann equation F1) from a Lagrangian depending just on $a(t)$ (and the matter variables).

It turns out, however, that just by introducing as an additional variable a non-trivial time-time component of the metric (whose variation can then be used to impose the Friedmann equation F1), the reduction is consistent also at the Lagrangian level. We will introduce this additional variable in the form of the lapse function $N(t)$ of (35.10), but we will now simply write t for the time-coordinate and not t', so that the ansatz for the metric is

$$ds^2 = -N^2(t)dt^2 + a^2(t)d\tilde{s}^2 \quad , \tag{35.120}$$

and we can always choose the $N(t) = 1$ "t is comoving proper time" gauge at the end of the calculations.

In section 35.1 we had already determined the Ricci scalar of this metric, namely (now again, consistent with $t' \to t$ denoting time-derivatives by overdots rather than primes) (35.12)

$$R = \frac{6}{a^2N^3}(N(a\ddot{a} + \dot{a}^2) - a\dot{a}\dot{N} + kN^3) \quad . \tag{35.121}$$

The Einstein-Hilbert Lagangian density is therefore

$$\sqrt{g}R = Na^3\sqrt{\tilde{g}}R = \sqrt{\tilde{g}}\frac{6a}{N^2}(N(a\ddot{a} + \dot{a}^2) - a\dot{a}\dot{N} + kN^3) \quad . \tag{35.122}$$

The only dependence on the spatial coordinates is in the spatial volume element $\sqrt{\tilde{g}}$. Therefore, integrating this Lagrangian density over the space-time, one obtains a (potentially infinite) volume factor from the integration over the spatial coordinates which we will simply drop. Thus in the infinite-volume case ($k = -1$ or $k = 0$ without periodic toroidal identifications) this is not really a reduction in the strict technical sense. However, this is not our main concern here. Our aim is simply to obtain a Lagrangian formulation of the Friedmann equations and (as we will see) this can be accomplished by just dropping the integration over the spatial coordinates.

We are thus left with the 1-dimensional (mechanics) gravitational Lagrangian

$$L_{EH} = \frac{6a}{N^2}(N(a\ddot{a} + \dot{a}^2) - a\dot{a}\dot{N} + kN^3) \quad . \tag{35.123}$$

Note that this depends not only on $a(t)$ and $\dot{a}(t)$, and $N(t)$ and $\dot{N}(t)$, but also (linearly) on $\ddot{a}(t)$. This of course reflects the general property of the Einstein-Hilbert Lagrangian that it depends on 2nd derivatives of the metric, thus necessitating the introduction of boundary terms etc., as discussed in sections 20.4 and 20.5.

Naively, to eliminate this $\ddot{a}$-term, we can integrate the first term of the Lagrangian by parts,

$$6a^2\ddot{a}N^{-1} = \frac{d}{dt}(6a^2\dot{a}N^{-1}) + 6aN^{-2}(-2\dot{a}^2N + a\dot{a}\dot{N}) \quad . \tag{35.124}$$

We see that this has the effect of changing the sign of the 2nd term of (35.123) and cancelling the 3rd term $\sim \dot{N}$, so that we have

$$L_{EH} = \frac{6a}{N^2}(-N\dot{a}^2 + kN^3) + \frac{d}{dt}(6a^2\dot{a}N^{-1}) \quad . \tag{35.125}$$

It turns out that the total derivative term is cancelled precisely by the Gibbons-Hawking-York boundary term discussed in section 20.5. This should not come as a surprise: after all, that was its purpose. To see this explicitly, note that with any of the definitions or characterisations of the extrinsic curvature tensor given in section 18, one finds that the extrinsic curvature of the constant time t hypersurfaces (with unit normal vector $(1/N)\partial_t$) in the ambient space-time is

$$K_{ij} = \frac{1}{2N}\partial_t g_{ij} = \frac{\dot{a}}{aN} g_{ij} \tag{35.126}$$

so that the trace is

$$K = g^{ij}K_{ij} = \frac{3\dot{a}}{aN} \quad . \tag{35.127}$$

Thus the Gibbons-Hawking-York boundary term (20.61) reduces to (again dropping the spatial volume element $\sqrt{g}$ and using $\epsilon = -1$ for spacelike hypersurfaces)

$$2\epsilon\sqrt{h}\, K \to L_{GHY} = -2a^3(3\dot{a}/aN) = -6a^2\dot{a}N^{-1} \quad . \tag{35.128}$$

The complete gravitational Lagrangian is now, in analogy with the complete standard gravitational action (20.67),

$$S_g[g_{\alpha\beta}] = S_{EH}[g_{\alpha\beta}] + S_{GHY}[g_{\alpha\beta}] \quad , \tag{35.129}$$

the Lagrangian

$$\begin{aligned} L_g &= L_{EH} + \frac{d}{dt}L_{GHY} \\ &= \frac{6a}{N^2}(-N\dot{a}^2 + kN^3) \\ &= 6N(-a(\dot{a}/N)^2 + ka) \quad . \end{aligned} \tag{35.130}$$

Inclusion of the cosmological constant Λ,

$$\sqrt{g}R \to \sqrt{g}(R - 2\Lambda) \quad , \tag{35.131}$$

leads to

$$L_g = 6N(-a(\dot{a}/N)^2 + ka - \Lambda a^3/3) \quad . \tag{35.132}$$

<u>Remarks:</u>

1. Note that, as desired, the Lagrangian now only depends on the "fields" $a(t)$ and $N(t)$ and (at most) their 1st derivatives.

2. Actually, we see that the Lagrangian depends only on $N(t)$, not its time-derivative $\dot{N}(t)$. This will also turn out to be the case for the matter action (which typically does not depend on any derivatives of the metric at all). Thus the role of $N(t)$ is just that of a Lagrange multiplier, and as we will see the constraint it imposes is simply the Friedmann equation F1.

3. One can also obtain (35.130) directly from the Gauss-Codazzi (21.6) or ADM form (21.42) of the action,

$$S_{ADM}[g_{\alpha\beta}] = \int \sqrt{g} d^4x \, (\bar{R} + K^{ij}K_{ij} - K^2) \ . \tag{35.133}$$

Indeed, the 3-dimensional scalar curvature is simply (this follows e.g. from (14.19))

$$\bar{R} = 6k/a^2 \ \ , \tag{35.134}$$

while from (35.126) one has

$$K^{ij}K_{ij} - K^2 = 3(\dot{a}/aN)^2 - 9(\dot{a}/aN)^2 = -6(\dot{a}/aN)^2 \tag{35.135}$$

so that (again dropping the spatial volume element $\sqrt{\tilde{g}}$), one finds

$$L_{ADM} = a^3 N(-6(\dot{a}/aN)^2 + 6k/a^2) = L_g \ . \tag{35.136}$$

We now need to add matter. The phenomenological description of matter in terms of the energy and pressure densities $\rho(t)$ and $p(t)$ is not sufficient for a Lagrangian formulation. The simplest matter model with a "microscopic" Lagrangian description is that of a scalar field in the gravitational field described by the class of metrics (35.120), and this is the model we will consider.

The starting point is thus the action (6.14) for a scalar field with a potential $V(\phi)$,

$$S[\phi, g_{\alpha\beta}] = \int \sqrt{g} d^4x \left[-\tfrac{1}{2} g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi - V(\phi) \right] \ . \tag{35.137}$$

Compatibility with the symmetries of the gravitational field, in particular the spatial homogeneity, requires that the scalar field is spatially constant and is thus only a function of time t. With this assumption (and again dropping $\sqrt{\tilde{g}}$), the matter Lagrangian reduces to

$$L_m = Na^3(\dot{\phi}^2/2N^2 - V(\phi)) = N(a^3\dot{\phi}^2/2N^2 - a^3V(\phi)) \ . \tag{35.138}$$

Therefore the total gravitational + matter Lagrangian and action are (reinstating the gravitational coupling constant)

$$\begin{aligned} L_{tot} &= \frac{1}{16\pi G_N} L_g + L_m \\ &= N\left(-\frac{3}{4\pi G_N}\frac{a\dot{a}^2}{2N^2} + \frac{a^3\dot{\phi}^2}{2N^2} + \frac{3ka}{8\pi G_N} - \frac{\Lambda a^3}{8\pi G_N} - a^3V(\phi) \right) \end{aligned} \tag{35.139}$$

and

$$\begin{aligned} S_{tot}[a,\phi,N] &= \int dt \ L_{tot} \\ &= \int dt \ N\left(-\frac{3}{4\pi G_N}\frac{a\dot{a}^2}{2N^2} + \frac{a^3\dot{\phi}^2}{2N^2} + \frac{3ka}{8\pi G_N} - \frac{\Lambda a^3}{8\pi G_N} - a^3V(\phi) \right) \end{aligned} \tag{35.140}$$

Before analysing the equations of motion and the Hamiltonian arising from this Lagrangian and action, let us note the following points:

1. From this form of the action it is evident that one role of $N(t)$ is to ensure time reparametrisation invariance of the action. Indeed, dt and $N(t)$ only appear in the combinations $N(t)dt$ or $(1/N(t))(d/dt)$. Hence the action remains unchanged under time reparametrisations $t \to \tilde{t}(t)$ if one simultaneously transforms $N(t)$ to a new lapse function $\tilde{N}(\tilde{t})$, say, according to

$$N(t)dt = \tilde{N}(\tilde{t})d\tilde{t} \quad \Leftrightarrow \quad \tilde{N} = N(dt/d\tilde{t}) \ . \tag{35.141}$$

   This is the remnant of the general covariance of the original action, and the mechanism here is the same as that which rendered the "parent" geodesic action (2.111) reparametrisation invariant.

2. We see that this action does not involve the time-derivative of the lapse function $N(t)$. Hence $N(t)$ acts as a Lagrange multiplier enforcing a constraint. This constraint can be regarded as the constraint associated to this reparametrisation invariance, and such constraints are thus a characteristic feature of any reparametrisation invariant or generally covariant system.

3. In this combined gravity plus matter action, we now see very clearly that the gravitational kinetic term has the opposite sign of the matter kinetic term. This is a particular (and particularly obvious) manifestation of the general fact (mentioned in connection with the DeWitt metric at the end of section 21.2) that the gravitational kinetic term is not positive definite and that the negative "direction" in field space is associated with overall spatial volume deformations.

4. Finally, this form of the action makes it particularly obvious that the cosmological constant term can also be regarded as leading to (or arising from) a constant shift of the potential for the matter fields,

$$\frac{\Lambda a^3}{8\pi G_N} + a^3 V(\phi) = a^3(V(\phi) + \rho_\Lambda) \ . \tag{35.142}$$

   We will therefore absorb the cosmological constant term into the scalar potential in the following and not carry it around explicitly.

Now let us look at the Euler-Lagrange equations arising from this action.

- The Euler-Lagrange Equation for $N(t)$

  As mentioned before, $N(t)$ acts as a Lagrange multiplier enforcing the constraint

$$\frac{\delta S_{tot}}{\delta N} = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial N(t)} = 0 \ . \tag{35.143}$$

  Explicitly this is the condition

$$-\frac{3}{4\pi G_N}\frac{a\dot{a}^2}{2N^2} + \frac{a^3\dot{\phi}^2}{2N^2} - \frac{3ka}{8\pi G_N} + a^3 V(\phi) = 0 \ . \tag{35.144}$$

This really is a constraint rather than an equation of motion, because it only depends on the fields and their 1st derivatives, not their 2nd derivatives, and thus constitutes a condition on initial data on some constant time initial hypersurface.

In the gauge $N(t) = 1$ (which we can now choose, after having determined the equation of motion arising from varying N in the action), this constraint becomes

$$\frac{3}{8\pi G_N}(a\dot{a}^2 + ak) = \frac{a^3\dot{\phi}^2}{2} + a^3 V(\phi) \ , \tag{35.145}$$

or

$$\frac{\dot{a}^2 + k}{a^2} = \frac{8\pi G_N}{3}(\frac{1}{2}\dot{\phi}^2 + V(\phi)) \ . \tag{35.146}$$

Now recall that for such a scalar field the energy density and pressure are given by (35.38)

$$\rho(t) = \tfrac{1}{2}\dot{\phi}^2 + V(\phi) \quad , \quad p(t) = \tfrac{1}{2}\dot{\phi}^2 - V(\phi) \ . \tag{35.147}$$

Hence we can write the constraint as

$$\frac{\dot{a}^2 + k}{a^2} = \frac{8\pi G_N}{3}\rho \ , \tag{35.148}$$

and this we now recognise as precisely the Friedmann equation F1.

Having derived the constraint arising from the variation of the lapse function N, we can now simplify our life by using the reparametrisation invariance to set $N(t) = 1$. Thus we can work with the simpler action

$$S_{tot}[a,\phi] = \int dt \left(-\frac{3}{4\pi G_N}\frac{a\dot{a}^2}{2} + \frac{a^3\dot{\phi}^2}{2} + \frac{3ka}{8\pi G_N} - a^3 V(\phi) \right) \ . \tag{35.149}$$

- The Euler-Lagrange equation for $a(t)$

  This is the equation

$$\frac{d}{dt}\left(-\frac{3a\dot{a}}{4\pi G_N} \right) = 3\left(-\frac{1}{4\pi G_N}\frac{\dot{a}^2}{2} + \frac{a^2\dot{\phi}^2}{2} + \frac{k}{8\pi G_N} - a^2 V(\phi) \right) \ , \tag{35.150}$$

  or

$$\frac{\ddot{a}}{a} = -4\pi G_N p - \frac{\dot{a}^2 + k}{2a^2} \ , \tag{35.151}$$

  where $p = p(t)$ is the pressure (35.147). Using the constraint (the Friedmann equation F1) derived above, this becomes

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3}(\rho + 3p) \ , \tag{35.152}$$

  which is precisely the Friedmann equation F2.

- The Euler-Lagrange equation for $\phi(t)$

  Finally, the equation of motion for the scalar field is

$$\frac{d}{dt}(a^3\dot{\phi}) = -a^3 V'(\phi) \quad \Leftrightarrow \quad \ddot{\phi} + 3(\dot{a}/a)\dot{\phi} + V'(\phi) = 0 \;\; , \tag{35.153}$$

  exhibiting the same friction term coupling to the Hubble parameter as (34.113). Multiplying by $\dot{\phi}$ and noting that

$$\rho + p = \dot{\phi}^2 \tag{35.154}$$

  and

$$\dot{\rho} = \dot{\phi}\ddot{\phi} + V'(\phi)\dot{\phi} \;\; , \tag{35.155}$$

  it is evident that this equation can equivalently be written as

$$\ddot{\phi} + 3(\dot{a}/a)\dot{\phi} + V'(\phi) = 0 \quad \Leftrightarrow \quad \dot{\rho} + 3(\rho + p)(\dot{a}/a) = 0 \;\; , \tag{35.156}$$

  which is the Friedmann equation F3.

We have thus verified that we have indeed obtained a Lagrangian description of the complete set of Friedmann equations.

35.9 Hamiltonian Formulation of the Friedmann Equations

Having obtained the Lagrangian formulation, it is now quite straightforward to pass to a Hamiltonian formulation, and this despite the presence of constraints which, in general, can significantly complicate the Hamiltonian formulation (and subsequent canonical quantisation, say) of such systems.[154]

Our starting point is the Lagrangian (35.139) (with Λ absorbed into $V(\phi)$), i.e.

$$L_{tot} = N\left(-\frac{3}{4\pi G_N}\frac{a\dot{a}^2}{2N^2} + \frac{a^3\dot{\phi}^2}{2N^2} + \frac{3ka}{8\pi G_N} - a^3 V(\phi)\right) \quad . \tag{35.157}$$

In order to streamline the following discussion, it is convenient to consider the fields $a(t)$ and $\phi(t)$ as the two coordinates

$$Q^A = (a, \phi) \tag{35.158}$$

of a 2-dimensional dynamical system, with metric

$$G_{AB} = \mathrm{diag}(-3a/4\pi G_N, a^3) \;\; . \tag{35.159}$$

[154] For detailed textbook discussions of this subject, see e.g. M. Henneaux, C. Teitelboim, *Quantisation of Gauge Systems*, or H. Rothe, K. Rothe, *Classical and Quantum Dynamics of Constrained Hamiltonian Systems*.

With this notation, the Lagrangian can be written as

$$L_{tot} = N\left(\frac{1}{2N^2}G_{AB}\dot{Q}^A\dot{Q}^B - V(Q)\right) \tag{35.160}$$

where the potential is

$$V(Q) = a^3V(\phi) - \frac{3ka}{8\pi G_N} \quad . \tag{35.161}$$

Were it not for the presence of the lapse function N as an addditional dynamical variable, this would be a completely standard classical mechanics Lagrangian, with an obvious corresponding Hamiltonian. In particular, the momenta conjugate to the variables Q^A take the almost standard form

$$P_A = \frac{\partial L_{tot}}{\partial \dot{Q}^A} = G_{AB}\dot{Q}^B/N \quad , \tag{35.162}$$

i.e.

$$P_a = -\frac{3a\dot{a}}{4\pi G_N N} \quad , \quad P_\phi = \frac{a^3\dot{\phi}}{N} \quad . \tag{35.163}$$

These relations can, as usual, be used to eliminate the velocities in favour of the momenta. The major novelty is thus the presence of N, whose conjugate momentum vanishes,

$$P_N = \frac{\partial L_{tot}}{\partial \dot{N}} = 0 \quad . \tag{35.164}$$

This is our first constraint (and a *primary constraint* in the terminology of constrained systems). While this relation does not allow us to eliminate $\dot{N}$ in terms of the momentum P_N, this is not an issue here since the Lagrangian does not depend on $\dot{N}$ in the first place.

We can now follow the standard procedure to construct the Hamiltonian, via (in the case at hand, whether or not we include the $P_N\dot{N}$-term evidently makes no difference)

$$H_{tot} = P_N\dot{N} + P_A\dot{Q}^A - L_{tot} \quad . \tag{35.165}$$

Then one finds the Hamiltonian.

$$H_{tot}(N,Q,P) = N\left(\frac{1}{2}G^{AB}P_AP_B + V(Q)\right) \equiv N\mathcal{H}(Q,P) \quad . \tag{35.166}$$

As far as its dependence on (Q,P) is concerned, this presents no surprises: time evolution is given by the Hamilton equations

$$\begin{aligned} \dot{Q}^A &= +\frac{\partial H_{tot}}{\partial P_A} = \{Q^A, H_{tot}\} \\ \dot{P}_A &= -\frac{\partial H_{tot}}{\partial Q_A} = \{P_A, H_{tot}\} \quad , \end{aligned} \tag{35.167}$$

and the 1st of these just reproduces the definition of the momenta (35.162), while the 2nd then reproduces the Euler-Lagrange equations for the fields $Q^A = (a,\phi)$ discussed in the previous section.

Turning to N, consistency of the primary constraint $P_N = 0$ (i.e. the condition that it be preserved in time) requires

$$\dot{P}_N = \{P_N, H_{tot}\} = -\mathcal{H} = 0 \ . \tag{35.168}$$

Thus the primary constraint $p_N = 0$ gives rise to the *secondary constraint*

$$\mathcal{H}(Q, P) = 0 \tag{35.169}$$

(and there are no further constraints in this class of examples). This *Hamiltonian constraint* is precisely the Friedmann equation F1, i.e. the condition that was imposed in the Lagrangian formulation by the Lagrange multipler N,

$$\frac{\partial L_{tot}}{\partial N} = 0 \quad \Leftrightarrow \quad \mathcal{H} = 0 \ . \tag{35.170}$$

A painless way to see this is to note that the N-dependence in the Lagrangian (35.160),

$$L_{tot} = \frac{1}{2N} G_{AB} \dot{Q}^A \dot{Q}^B - N V(Q) \ , \tag{35.171}$$

is precisely such that differentiation with respect to N changes the relative sign between the 2 terms and thus essentially implements the Legendre transformation from the Lagrangian to the Hamiltonian (expressed as a function of the velocities),

$$-\frac{\partial L_{tot}}{\partial N} = \frac{1}{2N^2} G_{AB} \dot{Q}^A \dot{Q}^B + V(Q) = \mathcal{H} \ . \tag{35.172}$$

This structure

$$H = N\mathcal{H} \tag{35.173}$$

of the Hamiltonian, with the constraint $\mathcal{H} = 0$, is again characteristic of parametrisation invariant or generally covariant systems. In particular, this provides a concrete illustration of the general features of the gravitational Hamiltonian mentioned in section 21.7.

36 Cosmology IV: Qualitative Analysis

A lot can be deduced about the solutions of the Friedmann-Lemaître equations, i.e. the evolution of the universe in the Friedmann-Lemaître-Robertson-Walker cosmologies, without solving the equations directly and even without specifying a precise equation of state, i.e. a relation between p and ρ. In the following we will, in turn, discuss the Big Bang, the age of the universe, and its long term behaviour, from this qualitative point of view. I will then introduce the notions of critical density and density parameters, and discuss some global and causal aspects of these cosmological models (Penrose diagrams, horizons, ...).

36.1 The Past I: Big Bang (Existence of an Initial Singularity)

One amazing thing about the Friedmann-Lemaître-Robertson-Walker models is that all of them (provided that the matter content is reasonably physical - I will be more precise about this below) predict an initial singularity, commonly known as a Big Bang. This is very easy to see.

The Friedmann equation (F2),

$$-3\frac{\ddot{a}}{a} = 4\pi G_N(\rho + 3p) \quad \Leftrightarrow \quad q = \frac{4\pi G_N}{3H^2}(\rho + 3p) \ , \tag{36.1}$$

shows that, as long as the right-hand side is positive, one has $q > 0$, i.e. $\ddot{a} < 0$ so that the universe is decelerating due to gravitational attraction. This is the case for standard matter ($\rho > 0$) when it satisfies the *strong energy condition* (SEC) strictly (cf. the discussion in section 35.2, in particular (35.34)),

$$\text{strict SEC} \quad \Rightarrow \quad \rho + 3p > 0 \quad \Rightarrow \quad \ddot{a} < 0 \ . \tag{36.2}$$

It is also true for a negative cosmological constant (its negative energy density being outweighed by 3 times its positive pressure). It need not be true, however, in the presence of a positive cosmological constant which provides an accelerating contribution to the expansion of the universe. We will, for the time being, continue with the assumption that Λ is zero or, at least, non-positive, even though, as we will discuss later, recent evidence (strongly) suggests the presence of a non-negligible positive cosmological constant in our universe today (which is, however, totally irrelevant for the energy budget of the early universe).

Since $a > 0$ by definition, $\dot{a}(t_0) > 0$ because we observe a redshift, and $\ddot{a} < 0$ because $\rho + 3p > 0$, it follows that there cannot have been a turning point in the past and $a(t)$ must be concave downwards. Therefore $a(t)$ must have reached $a = 0$ at some *finite* time in the past. We will call this time $t = 0$, $a(0) = 0$.

As ρa^4 is constant for radiation (an apppropriate description of earlier periods of the universe), this shows that the energy density grows like $1/a^4$ as $a \to 0$ so this leads to quite a singular situation.

Once again, as in our discussion of black holes, it is natural to wonder at this point if the singularities predicted by General Relativity in the case of cosmological models are generic or only artefacts of the highly symmetric situations we were considering. And again there are singularity theorems applicable to these situations which state that, under reasonable assumptions about the matter content, singularities will occur independently of assumptions about symmetries.

36.2 The Past II: Age of the Universe

With the normalisation $a(0) = 0$, it is fair to call t_0 the age of the universe. If $\ddot{a}$ had been zero in the past for all $t \leq t_0$, then we would have

$$\ddot{a} = 0 \quad \Rightarrow \quad a(t) = a_0 t / t_0 \ , \tag{36.3}$$

and

$$\dot{a}(t) = a_0 / t_0 = \dot{a}_0 \ . \tag{36.4}$$

This would determine the age of the universe to be

$$\ddot{a} = 0 \quad \Rightarrow \quad t_0 = \frac{a_0}{\dot{a}_0} = H_0^{-1} \ , \tag{36.5}$$

where H_0^{-1} is the *Hubble time*. However, provided that $\ddot{a} < 0$ for $t \leq t_0$ (as discussed above, this holds under suitable conditions on the matter content - which may or may not be realised in our universe), the actual age of the universe must be smaller than this,

$$\ddot{a} < 0 \quad \Rightarrow \quad t_0 < H_0^{-1} \ . \tag{36.6}$$

Thus the Hubble time H_0 sets an upper bound on the age of the universe. See Figure 46 for an illustration of this. In particular, this means that one can obtain an upper limit on the age of the universe by determining the leading (linear) term in the Hubble relation (34.104) from observations of redshifts of galaxies!

36.3 The Future: Long Term Behaviour

Let us now try to take a look into the future of the universe. Again we will see that it is remarkably simple to extract relevant information from the Friedmann equations without ever having to solve an equation.

We will assume that $\Lambda = 0$ and that we are dealing with matter with $\rho > 0$ and $w > -1/3$ (the SEC). The Friedmann equation (F1) can be written as

$$\dot{a}^2 = \frac{8\pi G_N}{3} \rho a^2 - k \ . \tag{36.7}$$

The left-hand side is manifestly non-negative. Let us see what this tells us about the right-hand side. Focus on the first term $\sim \rho a^2$. This term is strictly positive and, according to (35.95), behaves as

$$\rho a^2 \sim a^{-3(1+w)+2} = a^{-(1+3w)} \quad . \tag{36.8}$$

Thus for $w > -1/3$ the exponent is negative, so that if and when the cosmic scale factor $a(t)$ goes to infinity, one has

$$\lim_{a\to\infty} \rho a^2 = 0 \quad . \tag{36.9}$$

Now let us look at the second term on the right-hand side of (36.7), and analyse the 3 choices for k. For $k = -1$ or $k = 0$, the right hand side of (36.7) is strictly positive. Therefore $\dot{a}$ is never zero and since $\dot{a}_0 > 0$, we must have

$$\dot{a}(t) > 0 \;\; \forall\, t \quad . \tag{36.10}$$

Thus we can immediately conclude that open and flat universes must expand forever, i.e. they are open in space *and* time.

By taking into account (36.9), we can even be somewhat more precise about the long term behaviour. For $k = 0$, we learn that

$$k = 0: \quad \lim_{a\to\infty} \dot{a}^2 = 0 \quad . \tag{36.11}$$

Thus the universe keeps expanding but more and more slowly as time goes on. By the same reasoning we see that for $k = -1$ we have

$$k = -1: \quad \lim_{a\to\infty} \dot{a}^2 = 1 \quad . \tag{36.12}$$

Thus the universe keeps expanding, reaching a constant limiting velocity.

For $k = +1$, validity of (36.9) would lead us to conclude that $\dot{a}^2 \to -1$, but this is obviously a contradiction. Therefore we learn that the $k = +1$ universes never reach $a \to \infty$ and that there is therefore a maximal radius a_{max}. This maximal radius occurs for $\dot{a} = 0$ and therefore

$$k = +1: \quad a_{max}^2 = \frac{3}{8\pi G_N \rho} \quad . \tag{36.13}$$

Note that intuitively this makes sense. For larger ρ or larger G_N the gravitational attraction is stronger, and therefore the maximal radius of the universe will be smaller. Since we have $\ddot{a} < 0$ also at a_{max}, again there is no turning point and the universe recontracts back to zero size leading to a *Big Crunch*. Therefore, spatially closed universes ($k = +1$) with physical matter are also closed in time. All of these findings are summarised in Figure 46.

If the cosmological constant Λ is not zero, this poetic correspondence "(open/closed) in space $\Leftrightarrow$ (open/closed) in time" is no longer necessarily true since the cosmological constant (with $\rho_\Lambda a^2 \sim a^2$) will dominate over other forms of matter for sufficiently large $a(t)$ - the asymptotic solutions in this case are given in section 37.5.

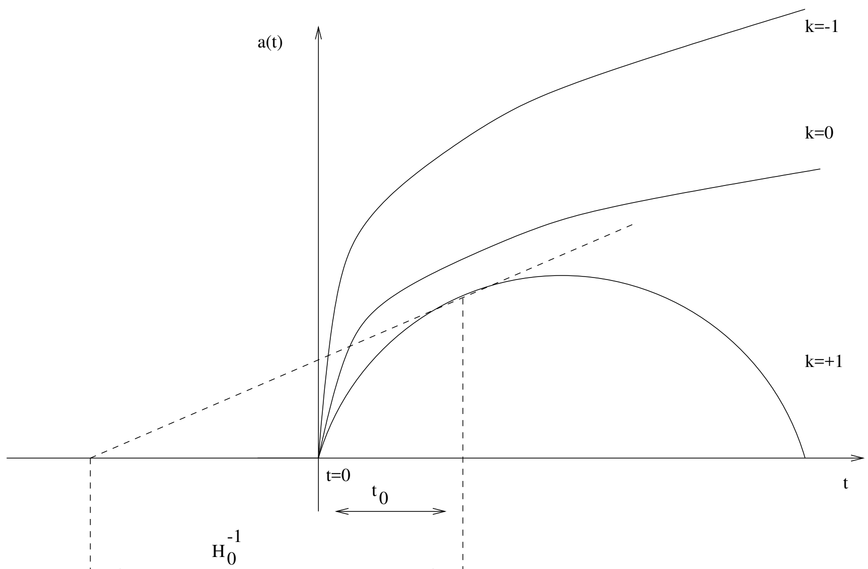


Figure 46: Qualitative behaviour of the Friedmann-Lemaître-Robertson-Walker models for $\Lambda = 0$. All models start with a Big Bang. For $k = +1$ the universe reaches a maximum radius and recollapses after a finite time. For $k = 0$, the universe keeps expanding but the expansion velocity tends to zero for $t \to \infty$ or $a \to \infty$. For $k = -1$, the expansion velocity approaches a non-zero constant value. Also shown is the significance of the Hubble time for the $k = +1$ universe showing clearly that H_0^{-1} gives an upper bound on the age of the universe.

In order to make the dependence of the Friedmann equation (F1) on the equation of state parameters w_b and on $a(t)$ more manifest, it is useful to use the conservation law (35.95,36.8) to write

$$\frac{8\pi G_N}{3}\rho_b(t)a(t)^2 = C_b a(t)^{-(1+3w_b)} \quad \Leftrightarrow \quad \Omega_b \dot{a}(t)^2 = C_b a(t)^{-(1+3w_b)} \tag{36.14}$$

for some constant C_b. Then the Friedmann equation takes the more explicit (in the sense that all the dependence on the cosmic scale factor $a(t)$ is explicit) form

$$\dot{a}^2 = \sum_b C_b a^{-(1+3w_b)} - k + \frac{\Lambda}{3}a^2 \quad . \tag{36.15}$$

In addition to the vacuum energy (and pressure) provided by Λ, there are typically two other kinds of matter which are relevant in our approximation, namely matter in the form of dust ($w = 0$) and radiation ($w = 1/3$). Denoting the corresponding constants by C_m and C_r respectively, the Friedmann equation that we will be dealing with takes the form

$$(F1'') \quad \dot{a}^2 = \frac{C_m}{a} + \frac{C_r}{a^2} - k + \frac{\Lambda}{3}a^2 \quad , \tag{36.16}$$

illustrating the qualitatively different contributions to the time-evolution.

One can then characterise the different eras in the evolution of the universe by which of the above terms dominates, i.e. gives the leading contribution to the equation of motion for a. This already gives some insight into the physics of the situation. We will call a universe

1. *matter dominated* if C_m/a dominates
2. *radiation dominated* if C_r/a^2 dominates
3. *curvature dominated* if k dominates
4. *cosmological constant dominated* if Λa^2 dominates

Here are some immediate consequences of the Friedmann equation (F1"):

1. No matter how small C_r is, provided that it is non-zero, for sufficiently small values of a that term will dominate and one is in the radiation dominated era. In that case, one finds the characteristic behaviour

$$\dot{a}^2 = \frac{C_r}{a^2} \quad \Rightarrow \quad a(t) = (4C_r)^{1/4}t^{1/2} \quad . \tag{36.17}$$

It is more informative to trade the constant C_r for the condition $a(t_0) = a_0$, which leads to

$$a(t) = a_0(t/t_0)^{1/2} \quad . \tag{36.18}$$

2. On the other hand, if C_m dominates, one has the characteristic behaviour

$$\dot{a}^2 = \frac{C_m}{a} \quad \Rightarrow \quad a(t) = (9C_m/4)^{1/3} t^{2/3} \tag{36.19}$$

or

$$a(t) = a_0 (t/t_0)^{2/3} \tag{36.20}$$

Both this and the previous example illustrate the characteristic late-time behaviour $a(t) \to \infty$ with $\dot{a}(t) \to 0$ of the $k = 0$ models.

3. For general equation of state parameter $w \neq -1$, one similarly has

$$a(t) = a_0 (t/t_0)^h \quad , \quad h = \frac{2}{3(1+w)} \quad . \tag{36.21}$$

This describes a decelerating universe ($h(h-1) < 0 \Rightarrow 0 < h < 1$) for $w > -1/3$ and an accelerating universe ($h > 1$) for $-1 < w < -1/3$. This evidently reflects, and agrees with, the fact that from (F2) one has $\ddot{a} < 0$ for $\rho + 3p > 0$ and $\ddot{a} > 0$ for $\rho + 3p < 0$.

4. For the special case $w = -1/3$, one has $h = 1$ and thus the linear evolution $a(t) \sim t$. Since, as noted in (36.38) below, one can formally attribute an equation of state parameter $w_k = -1/3$ to the curvature contribution to the Friedmann equation, this solution arises not only for an exotic matter component with $w = -1/3$ and $k = 0$, but also for an empty universe with $k = -1$. We will look at the latter (the *Milne universe*) in more detail in section 37.1 (evidently an empty universe with $k = +1$, governed by $\dot{a}^2 = -1$, is not possible).

   In spite of sharing the same Friedmann equation (F1') and the same solution $a(t)$, these two universes with $w = -1/3$ are decidedly not identical for the obvious reason that one is empty and the other one is not, and thus they solve the Einstein equations with very different energy-momentum tensors (alternatively, e.g. their Misner-Sharp masses (35.118) are different). More dramatically, the scalar curvature $R(t)$ for $a(t) = t$ is

$$R(t) = \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k) = \frac{6(1+k)}{t^2} \quad . \tag{36.22}$$

   Thus there is a singularity at $t = 0$ for $k = 0$ while $R(t) = 0$ for the Milne metric with $k = -1$ (which turns out to be just Minkowski space in disguise).

   Note, as an side, that one cannot conclude from the fact alone that one has $k = -1$ and the other one has $k = 0$ that they are different, since it is possible that a given universe can be foliated in different ways by spatial hypersurfaces with different curvatures. An example of this is provided by the de Sitter universe, the solution to the Friedmann equations with a positive cosmological constant (and no other matter) - see section 37.5.

5. For sufficiently large a, the cosmological constant Λ, if not identically zero, will always dominate, no matter how small the cosmological constant may be, as all the other energy-content of the universe gets more and more diluted. In particular, for $k = 0$, the Friedmann equation for a positive cosmological constant reduces to
$$\dot{a}^2 = (\Lambda/3)a^2 \quad \Rightarrow \quad a(t) = a(t_0)\mathrm{e}^{\pm\sqrt{\Lambda/3}(t-t_0)} \quad , \tag{36.23}$$
with Hubble parameter
$$H(t) = \frac{\dot{a}}{a} = \pm\sqrt{\Lambda/3} = H_0 \quad . \tag{36.24}$$
This gives the $k = 0$ metric of de Sitter space,
$$ds^2 = -dt^2 + \mathrm{e}^{2H_0 t} d\vec{x}^2 \quad . \tag{36.25}$$

6. Only for $\Lambda = 0$ does k dominate for large a and one obtains, as we saw before, a constant expansion velocity (for $k = 0, -1$).

7. We will find and discuss various other exact solutions in section 37.

36.5 Density Parameters and the Critical Density

The primary purpose of this section is to introduce some convenient and commonly used notation and terminology in cosmology associated with the Friedmann equation (F1'). We will now include the cosmological constant in our analysis. For starters, however, let us again consider the case $\Lambda = 0$ (or include ρ_Λ as one contribution to ρ). (F1') can be written as
$$\frac{8\pi G_N \rho}{3H^2} - 1 = \frac{k}{a^2 H^2} \quad . \tag{36.26}$$
If one defines the *critical density* ρ_{cr} by
$$\rho_{cr} = \frac{3H^2}{8\pi G_N} \quad , \tag{36.27}$$
and the *density parameter* Ω by
$$\Omega = \frac{\rho}{\rho_{cr}} = \frac{8\pi G_N\, \rho}{3H^2} \quad , \tag{36.28}$$
then (F1') becomes
$$\Omega - 1 = \frac{k}{a^2 H^2} \tag{36.29}$$
Thus the sign of k is determined by whether the actual energy density ρ in the universe is greater than, equal to, or less than the critical density,
$$\begin{array}{ccccccc} \Omega < 1 & \Leftrightarrow & \rho < \rho_{cr} & \Leftrightarrow & k = -1 & \Leftrightarrow & \text{open} \\ \Omega = 1 & \Leftrightarrow & \rho = \rho_{cr} & \Leftrightarrow & k = 0 & \Leftrightarrow & \text{flat} \\ \Omega > 1 & \Leftrightarrow & \rho > \rho_{cr} & \Leftrightarrow & k = +1 & \Leftrightarrow & \text{closed} \end{array}$$

Note that, in particular, $\rho_{cr} = 0$ at a turning point (maximal radius, $H = 0$), and correspondingly $\Omega \to \infty$ there.

This can be generalised to several species of (not mutually interacting) matter, characterised by equation of state parameters w_b, subject to the condition $w_b > 0$ or $w_b > -1/3$, with density parameters

$$\Omega_b = \frac{\rho_b}{\rho_{cr}} \quad . \tag{36.30}$$

The total matter contribution Ω_M is then

$$\Omega_M = \sum_b \Omega_b \quad . \tag{36.31}$$

Along the same lines we can also include the cosmological constant Λ. Inspection of the Friedmann equations reveals that the presence of a cosmological constant is equivalent to adding matter $(\rho_\Lambda, p_\Lambda)$ with

$$\Lambda \quad \Leftrightarrow \quad \rho_\Lambda = -p_\Lambda = \frac{\Lambda}{8\pi G_N} \quad w_\Lambda = -1 \quad . \tag{36.32}$$

in agreement with what we had already deduced in (35.31). Note that this identification is consistent with the conservation law (F3), since Λ is constant.

Then the Friedmann equation (F1') with a cosmological constant can be written as

$$(F1') \quad \Leftrightarrow \quad \Omega_M + \Omega_\Lambda = 1 + \frac{k}{a^2H^2} \quad , \tag{36.33}$$

where

$$\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_{cr}} = \frac{\Lambda}{3H^2} \quad . \tag{36.34}$$

One can also formally attribute a density parameter

$$\Omega_k = -\frac{k}{a^2H^2} \tag{36.35}$$

to the curvature, so that the Friedmann equation (F1) can now succinctly (if somewhat obscurely) be written as the condition that the sum of all density parameters be equal to 1,

$$(F1') \quad \Leftrightarrow \quad \Omega_M + \Omega_\Lambda + \Omega_k = 1 \quad . \tag{36.36}$$

Note that the Friedmann equation can thus be regarded as a kind of balance equation among the contributions of matter, dark energy and (negative) curvature: if you have more of one, you need less of the other and vice versa.

The corresponding curvature energy density ρ_k can now be determined from

$$\Omega_k = \frac{8\pi G_N}{3H^2}\rho_k \stackrel{!}{=} -\frac{k}{a^2H^2} \quad , \tag{36.37}$$

and from (F2), which does not depend on k, one deduces that the corrsponding pressure is $p_k = w_k \rho_k$ with $w_k = -1/3$, so that $\rho_k + 3p_k = 0$. Thus the curvature contribution can be described as

$$k \quad \Leftrightarrow \quad \rho_k = -3p_k = \frac{-3k}{8\pi G_N a^2} \quad w_k = -1/3 \ . \tag{36.38}$$

The continuity equation (F3) is identically satisfied in this case (or, if you prefer, requires that k is constant).

Finally, the 2nd order equation (F2') can also be written in terms of the density parameters,

$$q = \tfrac{1}{2}\sum_b (1+3w_b)\Omega_b - \Omega_\Lambda \ . \tag{36.39}$$

Denoting the values of the parameters today, at time $t = t_0$, by a subscript 0, the two key equations relating the cosmological parameters are

$$\begin{aligned} (\Omega_M)_0 + (\Omega_\Lambda)_0 + (\Omega_k)_0 &= 1 \\ \tfrac{1}{2}(1+3w_0)(\Omega_M)_0 - (\Omega_\Lambda)_0 &= q_0 \end{aligned} \tag{36.40}$$

Clearly, it is of utmost importance to determine the various contributions ρ_a to the matter density ρ of the universe (and to determine ρ_{cr} e.g. by measurements of the Hubble parameter $H(t)$, i.e. of H_0, q_0 etc.). From H_0 and q_0 and $(\Omega_M)_0$ one can then in principle determine $(\Omega_\Lambda)_0$ and $(\Omega_k)_0$.

36.6 Causal Structure and Penrose Diagrams for $k = 0$

In order to gain a better understanding of a space-time, it is always useful to study its null geodesics, the behaviour of lightcones and thus its causal structure. In this section, as a first step towards this, we will shed light on the global structure of (spatially flat, $k = 0$) cosmological space-times by constructing Penrose diagrams for them. The reason for the restriction to $k = 0$ is that we already know the exact solution to the Friedmann equations in this case, and with a single matter component characterised by the equation of state parameter w, namely

$$ds^2 = -dt^2 + a(t)^2 d\vec{x}^2 \tag{36.41}$$

with (36.21)

$$k = 0 \quad , \quad p = w\rho \quad \Rightarrow \quad a(t) = a_0 (t/t_0)^h \quad , \quad h = \frac{2}{3(1+w)} \tag{36.42}$$

(for $w \neq -1$). The case of a cosmological constant ($w = -1$), i.e. (anti-)de Sitter space, needs to be treated separately. Penrose diagrams for some other ($k = +1$) solutions will be given in section 37.

One crucial feature that will emerge from this analysis is the possible presence of *cosmological horizons* which delimit the regions of space-time that can be in causal contact at a given time or that are visible at a given time or in principle in the infinite future. This will be discussed further in section 36.7 below. This will also allow us to then, in sections 36.8 and 36.9, better understand the significance (or lack thereof) of the Hubble sphere or Hubble radius $R_H(t)$ (34.40) introduced in section 34.3.

For later reference, we note that the acceleration is

$$\ddot{a}(t) = h(h-1)\frac{a(t)}{t^2} \quad , \quad h(h-1) = -\frac{2(w+1/3)}{3(1+w)^2} \quad , \tag{36.43}$$

so that

$$h(h-1) \qquad \begin{cases} <0 \;\; \text{(decelerating)} & \text{for } \; 0<h<1 \quad w>-1/3 \\ >0 \;\; \text{(accelerating)} & \text{for } \; h>1 \qquad\;\; w<-1/3 \end{cases} \tag{36.44}$$

Now the 1st step will be the introduction of the conformal time coordinate η (cf. section 34.5) defined by

$$d\eta(t) = dt/a(t) \tag{36.45}$$

in terms of which the metric can be written as

$$ds^2 = a(\eta)^2(-d\eta^2 + d\vec{x}^2) \ . \tag{36.46}$$

This makes it manifest that the metric is conformally flat, and thus its causal structure is particularly easy to analyse. All we need to pay attention to is the range of η.

In the case at hand, we have

$$a(t) = a_0 t_0^{-h} t^h \quad \Rightarrow \quad d\eta = a_0^{-1} t_0^h t^{-h} dt \tag{36.47}$$

and therefore

$$\eta(t) = a_0^{-1} t_0^h t^{1-h}/(1-h) \tag{36.48}$$

for $h \neq 1$ and

$$\eta(t) = a_0^{-1} t_0^h \log|t| \tag{36.49}$$

for $h=1, w=-1/3$. Since the range of t is $(0,\infty)$, the range of η is

$$t \in (0,+\infty) \quad \Rightarrow \quad \begin{cases} \eta \in (0,+\infty) & \text{for } \; 0<h<1 \\ \eta \in (-\infty,+\infty) & \text{for } \; h=1 \\ \eta \in (-\infty,0) & \text{for } \; h>1 \ . \end{cases} \tag{36.50}$$

This shows that the causal structure and Penrose diagram of the space-time for $h = 1, w = -1/3$ are identical to that of Minkowski space in Figure 23, with the only crucial difference that there is now a singularity at i^-, i.e. at $t \to 0, \eta \to -\infty$ (recall that the scalar curvature is $R(t) \sim t^{-2}$ (36.22)).

For $h \neq 1$, on the other hand, as indicated in Figure 48, the Penrose diagram is given by the upper (lower) half of the Minkowski Penrose diagram respectively, with the addition of the initial spacelike singularity. This has drastic implications for the global causal structure (and the existence of cosmological horizons) to be explored below.

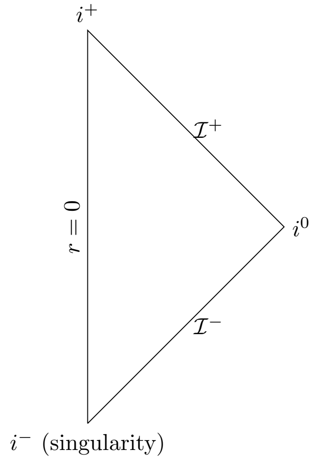


Figure 47: Penrose Diagram for the $k = 0, w = -1/3$ solution.

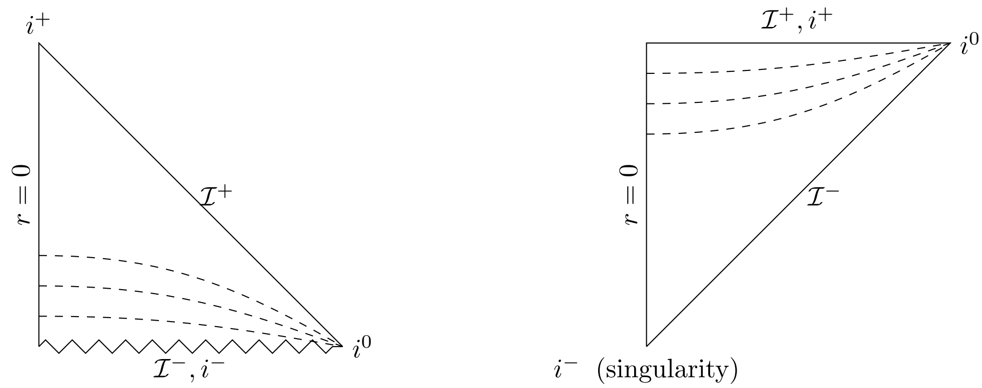


Figure 48: Penrose Diagram for the $k = 0, w \neq -1/3, -1$ solutions. On the left, the diagram for decelerating solutions with $0 < h < 1$, $w > -1/3$, on the right that for accelerating solutions with $h > 1$. Also schematically indicated (as dashed lines) are some surfaces of constant time t or η.

36.7 Lightrays and Cosmological Particle and Event Horizons

We now turn to one particularly relevant implication of the preceding analysis, namely the existence of (observer-dependent) *horizons* in cosmological space-times. Thus we will mainly consider the spatially flat case $k = 0$ in the following, but in principle everything can easily and immediately be extended to $k \neq 0$, e.g. by working in the polar coordinates (34.10), with $r \to \psi$. In the following, the term "comoving distance" will refer to the coordinate distance as measured by the comoving radial coordinate r.

Let us first (re-)analyse the evolution of lightrays in the Robertson-Walker geometry. We have already studied one aspect of light propagation in Robertson-Walker geometries in section 34.8, namely the cosmological redshift. Here we will look at the propagation of light from a more global perspective (or at least with more global intentions).

By spatial isotropy, we can without loss of generality choose the lightrays to be purely radial. Then they are governed by the equation

$$dt^2 = a(t)^2 dr^2 \quad \Leftrightarrow \quad \frac{dr_\gamma(t)}{dt} = \pm\frac{1}{a(t)} \quad . \tag{36.51}$$

Here, in order to avoid a proliferation of objects with the same anonymous name r, we call the comoving photon path $r_\gamma(t)$. Since $a(t)$ is non-negative, lightrays propagate either in the direction of increasing comoving radial coordinate distance r, or in the direction of decreasing r. In particular, lightrays coming towards us (at $r = 0$) and reaching $r = 0$ at $t = t_0$ are governed by the equation

$$\frac{dr_\gamma(t)}{dt} = -\frac{1}{a(t)} \quad \Rightarrow \quad r_\gamma(t) = \int_t^{t_0} \frac{dt}{a(t)} \quad . \tag{36.52}$$

We will occasionally also denote this solution by

$$r_\gamma(t) \equiv r_\gamma(t; t_0) \quad \text{with} \quad r_\gamma(t_0, t_0) = 0 \quad , \tag{36.53}$$

when we want to make the dependence of the solution on t_0 more explicit. We see that this solution can be directly expressed in terms of conformal time η as

$$r_\gamma(t; t_0) = \eta(t_0) - \eta(t) \quad . \tag{36.54}$$

One can also, if one wishes, express light propagation in terms of the (instantaneous) proper spacelike distance

$$R_\gamma(t) = a(t) r_\gamma(t) \quad , \tag{36.55}$$

i.e. the spatial proper distance to the photon as measured by the maximally symmetric spatial geometry on the slice of constant time t but, as we will see in sections 36.8 and 36.9, this is not without its pitfalls.

As an example, for matter with the equation of state parameter $w \neq -1$ and cosmic scale factor $a(t) = a_0(t/t_0)^h$ (36.21), the equation for lightrays is

$$dr_\gamma(t)/dt = -t_0^h t^{-h}/a_0 \quad , \quad h = \frac{2}{3(1+w)} \quad . \tag{36.56}$$

Either from this, or from the result for $\eta(t)$ from (36.48), one finds that for $h \neq 1$

$$r_\gamma(t) = \frac{t_0}{a_0(1-h)}(1 - (t/t_0)^{1-h}) \quad . \tag{36.57}$$

and

$$R_\gamma(t) = \frac{1}{1-h}(t_0^{1-h} t^h - t) \tag{36.58}$$

(while for $h = 1$ one finds a logarithmic behaviour). We now look at 2 limiting cases of these lightrays.

1. Cosmological Particle Horizon

Backtracking lightrays that reach us today, the largest comoving distance of an object about which we can have received any (causal) information by today or that can have had any causal influence on us is given by

$$r_{ph}(t_0) = \lim_{t \to 0_+} r_\gamma(t; t_0) = \eta(t_0) - \eta(0) \quad . \tag{36.59}$$

(with $t = 0$ the time of the initial singularity). If this integral converges, this means that there is a region of the universe, namely that at comoving distance $r > r_{ph}(t_0)$, with which we could not have had any causal contact until today. In that case $r_{ph}(t_0)$ defines the *cosmological particle horizon* (or just *particle horizon*).

Expressed in terms of conformal time, whether or not there is a particle horizon is equivalent to the question if $\eta(t)$ is finite or diverges as $t \to 0$. From the analysis of the previous section for the single-component models with equation of state $p = w\rho$, $w \neq -1$, we can thus conclude that

- for the non-decelerating models, $h \geq 1$, one has $\eta(t) \to -\infty$, and therefore the space-time is causally equivalent to Minkowski space in the past, and there is no particle horizon;
- for the decelerating models with $0 < h < 1$, $w > -1/3$, or those that are dominated by ordinary matter or radiation at early times, say, $\eta(0)$ is finite, the causal structure is as in the left-hand diagram of Figure 48 and there is a particle horizon.

Remarks:

(a) It is also common to express the particle horizon in terms of

$$R_{ph}(t_0) = a(t_0) r_{ph}(t_0) = a(t_0) \int_0^{t_0} \frac{dt}{a(t)} \quad , \tag{36.60}$$

giving the present instantaneous proper distance to the particle horizon. At any time t_0, the particle horizon is a spatial 2-sphere with radius $R_{ph}(t_0)$ around an observer, no information being available about objects outside that 2-sphere. This particle horizon is observer-dependent, i.e. other comoving observers at other values of r will also have a particle horizon, but it will be a 2-sphere of radius $R_{ph}(t_0)$ around them.

(b) As (36.59) shows, $r_{ph}(t_0)$ grows monotonically with t_0 (so that more and more of the universe becomes visible as time goes on). A good way of visualising the particle horizon is to draw $r_{ph}(t)$ as a function of t in a (t, r)- or (η, r)-diagram.[155] For a comoving observer at $r = r_0$, this is simply the future

[155] Note that often the particle horizon is depicted in a different and somewhat less informative, or at

lightcone of the point $(0, r_0)$, and the particle horizon size at any given time t_0 is given by the intersection of this lightcone null surface with the spacelike constant $t = t_0$ surface.

(c) This is indicated in a Penrose diagram in Figure 49. It is evident from this diagram that the past lightcone of an event at some finite time $\eta > \eta(0)$ cannot cover the entire space, and that at early times suffficiently spatially separated events have non-intersecting past lightcones, i.e. have never been in causal contact in the past. We will come back to one problematic aspect and consequence of this in section 38.3 in our brief discussion of the (aptly) so-called horizon problem in cosmology.

In interpreting this diagram, it is also good to keep in mind the highly distorted nature of constant time hypersurfaces at late times, which all end up at i^0. Thus the particle horizon really reaches infinity $\mathcal{I}^+$ at $t \to \infty$ or $\eta \to +\infty$.

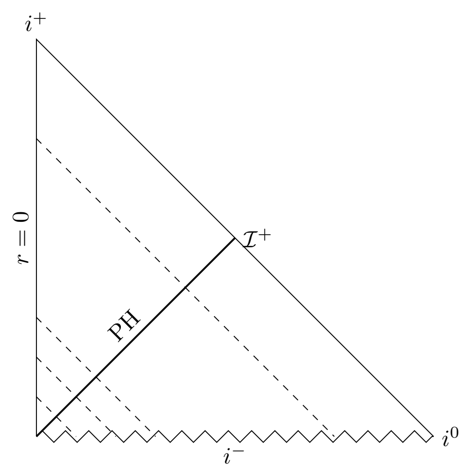


Figure 49: Cosmological Particle Horizon: indicated are the past lightcones of a comoving observer at $r = 0$ (dashed lines) as well as that observer's particle horizon (thick solid line).

This diagram also shows that the particle horizon can be regarded as the boundary of the region that can be influenced by the observer.

(d) One can of course also define these horizons for the models with $k \neq 0$. The case $k = +1$ is different in the sense that space is compact. E.g. the range of ψ in

$$ds^2 = a(\eta)^2(-d\eta^2 + d\psi^2 + \sin^2\psi d\Omega^2) \tag{36.61}$$

least potentially more confusing, way - see Figures 1 and 3 and section 3.4 of T. Davis, C. Lineweaver, *Expanding Confusion: common misconceptions of cosmological horizons and the superluminal expansion of the Universe*, `arXiv:astro-ph/0310808` for a nice discussion and illustration of this.

is finite, so that even if there is a particle horizon initially, it may be possible to see all of space (and later on all around the space at the back of one's head) at a later time, namely when the η-life-time of the universe is $\geq \pi$ (respectively $\geq 2\pi$). An example of this is provided by the matter dominated $k = +1$ universe, whose conformal diagram is given in Figure 51 of section 37.3.

(e) The particle horizon is a horizon in the sense that it is null and acts as a one-way membrane: any object that is inside the particle horizon or has crossed the particle horizon towards us will remain inside the particle horizon at all future times. This is also evident from the diagram.

2. Cosmological Event Horizon

From the above discussion it is also clear that the cosmological space-time is causally inequivalent to Minkowski space in the far future if η is bounded from above as $t \to \infty$. If it is bounded, then the past lightcone of an observer will never be able to cover all of space-time even as $t \to \infty$, i.e. there will be regions of space-time from which that observer can never ever receive any information, and in this case the past lightcone is known as a *cosmological event horizon.*

The radial coordinate size of the past lightcone from $t_0 = \infty$ at some time t is given by the future partner of (36.59), namely

$$r_{eh}(t) = \lim_{t_0 \to \infty} r_\gamma(t; t_0) = \eta(\infty) - \eta(t) \quad , \tag{36.62}$$

and there is an event horizon iff this integral is finite, i.e. iff $\eta(\infty)$ is finite.

Going back to our $k = 0$ single-component matter example, it is now precisely the accelerating cosmologies with $w < -1/3$ that exhibit an event horizon. This is iluustrated in Figure 50.

<u>Remarks:</u>

(a) Since current observations suggest that our universe will be dominated by a positive cosmological constant in the future, i.e. that it will be asymptotically de Sitter in the future, while it was dominated by standard types of matter in the past, the standard hot big bang / Λ-CDM model of our universe has both a particle horizon and an event horizon.

(b) This cosmological event horizon (or the issue whether or not one exists) is not particularly relevant for observational cosmology today (or in the foreseeable future), but it is of great theoretical interest since, at least in the (asymptotically) de Sitter case, in some respects this (again observer-dependent) horizon appears to have more in common with a true black hole horizon

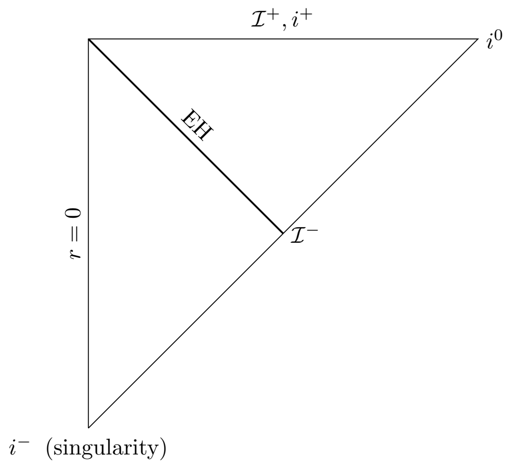


Figure 50: Cosmological Event Horizon: the region outside the event horizon (above the line EH) is invisible to a comoving observer at $r = 0$.

(like a temperature and an entropy) than one would have had the right to expect.[156]

(c) The event horizon is a horizon in the sense that it is null and acts as a one-way membrane: any object that is outside the event horizon or has crossed the event horizon away from us will remain outside the event horizon at all future times. Given that the event horizon is the past lightcone at a point at future infinity, these statements are almost a tautology, and are also evident from the diagram.

36.8 Lightrays and the Hubble Sphere

Recall that the Hubble radius $R_H(t)$ was introduced in section 34.3 as the (instantaneous) proper distance at which the recessional velocity (34.29)

$$V_{rec}(t) \equiv \frac{d}{dt} R_p(t) = H(t) R_p(t) \tag{36.63}$$

of a comoving object with $R_p(t) = a(t) r_1$ equals the speed of light,

$$V_{rec}(t) = c \quad \Leftrightarrow \quad R_H(t) = c/H(t) = ca(t)/\dot{a}(t) \tag{36.64}$$

For a contracting universe this should of course be replaced by $R_H = -c/H = c/|H|$, but until further notice we will consider the expanding case in the following (and we will from now on set $c = 1$ again).

[156]See e.g. M. Spradlin, A. Strominger, A. Volovich, *Les Houches Lectures on de Sitter Space*, `arXiv:hep-th/0110007` and references therein.

I also mentioned in section 34.3 that the Hubble sphere is occasionally also referred to as the Hubble horizon. Let us see if, or to which extent, it has such a role and how it is related to the cosmological particle and event horizons discussed in the previous section.

The most direct connection between the Hubble radius and the propgation of light arises from describing a photon path via its proper distance $R_\gamma(t)$ (36.55). This proper distance of an (out- or ingoing) lightray changes with time according to

$$\frac{d}{dt}R_\gamma(t) = \dot{a}(t)r_\gamma(t) \pm 1 = \frac{R_\gamma(t)}{R_H(t)} \pm 1 \quad . \tag{36.65}$$

Thus expressed (or plotted) in terms of proper distance, an ingoing lightray exhibits (or appears to exhibit) quite a different behaviour from the strictly monotonous $\dot{r}_\gamma < 0$ evolution of $r_\gamma(t)$. In particular, for such ingoing lightrays the proper distance reaches a maximum at a time $t = t_m$ when

$$r_\gamma(t_m) = 1/\dot{a}(t_m) = r_H(t_m) \quad \Leftrightarrow \quad R_\gamma(t_m) = R_H(t_m) \quad , \tag{36.66}$$

i.e. when the lightray crosses the Hubble sphere, and then decreases towards $r = 0$. Here

$$r_H(t) = R_H(t)/a(t) = 1/\dot{a}(t) \quad . \tag{36.67}$$

is the time-dependent comoving radial coordinate of the Hubble radius. This behaviour can intuitively (but with caution) be attributed to the fact that at the time when the lightray crosses the Hubble radius $R_H(t)$ it has recessional velocity (i.e. away from us) equal to c and is thus momentarily at rest with respect to us.

Before discussing how (not) to interpret this result, let us determine how the Hubble radius evolves with time. The evolution equation for $R_H(t)$ is

$$\frac{d}{dt}R_H(t) = \frac{d}{dt}(a(t)/\dot{a}(t)) = 1 + q(t) \quad , \tag{36.68}$$

while that for $r_H(t)$ is

$$\frac{d}{dt}r_H(t) = \frac{d}{dt}(1/\dot{a}(t)) = q(t)/a(t) \quad . \tag{36.69}$$

Using (36.39) in the schematic form

$$q(t) = \tfrac{1}{2}(1 + 3w)\Omega_M - \Omega_\Lambda \tag{36.70}$$

and

$$k = 0 \quad \Rightarrow \quad \Omega_M + \Omega_\Lambda = 1 \quad , \tag{36.71}$$

these equations can be written as

$$\begin{aligned} \frac{d}{dt}R_H(t) &= \tfrac{3}{2}(1 + w)\Omega_M \\ a(t)\frac{d}{dt}r_H(t) &= \tfrac{3}{2}(1 + w)\Omega_M - 1 \quad . \end{aligned} \tag{36.72}$$

In particular, if one has a single species of matter (or just a cosmological constant), i.e. $\Omega_M = 1$ (or $\Omega_M = 0$) one has

$$q = \tfrac{1}{2}(1+3w) \ , \tag{36.73}$$

which is constant, and

$$\begin{aligned} \frac{d}{dt}R_H(t) &= \tfrac{3}{2}(1+w) \\ a(t)\frac{d}{dt}r_H(t) &= \tfrac{1}{2}(1+3w) \ . \end{aligned} \tag{36.74}$$

This implies that the Hubble radius increases monotonically in time for

$$\begin{aligned} w > -1 \quad &\Rightarrow \quad \frac{d}{dt}R_H(t) > 0 \\ w > -\frac{1}{3} \quad &\Rightarrow \quad \frac{d}{dt}r_H(t) > 0 \ . \end{aligned} \tag{36.75}$$

Evidently for $w > -1$ and $a(t) \sim t^h$ the solution for $R_H(t)$ reproduces the definition $R_H = a/\dot{a}$,

$$w > -1 \quad \Rightarrow \quad R_H(t) = \tfrac{3}{2}(1+w)t = t/h \ , \tag{36.76}$$

while for $w = -1$,

$$w = -1 \quad \Rightarrow \quad R_H(t) = R_H(t_0) = 1/H_0 \ . \tag{36.77}$$

With these results at our disposal, we can now make the following elementary observations:

1. Since $R_H(t)$ is non-decreasing for $w \geq -1$, it is a triviality that the maximal proper distance $R_\gamma(t_m) = R_H(t_m)$ a lightray reached in the past that arrives today is bounded from above by the value of the Hubble radius today,

$$R_\gamma(t_m) = R_H(t_m) < R_H(t_0) \ . \tag{36.78}$$

(It is also easy to sharpen this upper bound somewhat for $w > -1/3$, but we will not need this).

2. Thus we see that the objects whose light we receive today at $r = 0$ cannot be further away from us today than at a proper distance $R_H(t_0)$ so that the Hubble radius $R_H(t_0)$ provides some kind of (rough) upper bound on the distance of such objects. This may suggest (to some) that therefore the Hubble radius provides a limit to what we can see (or can have seen) of the universe at time t_0.[157] However, this appears to me to be at best an extremely misleading way of phrasing things.

[157] See e.g. F. Melia, *The Cosmic Horizon*, `arXiv:0711.4181 [astro-ph]`, or the plots and numerical calculations performed in O. Bikwa, F. Melia, A. Shevchuk, *Photon Geodesics in FRW Cosmologies*, `arXiv:1112.4774 [astro-ph.CO]` in order to confirm (36.78) (which requires no numerical confirmation) in some examples.

In fact, even though $R_\gamma(t)$ reaches a maximum at $t = t_m$, this is not the case for the comoving radial coordinate $r_\gamma(t)$ along the photon path. The monotonous behaviour $\dot{r}_\gamma < 0$ for ingoing lightrays implies that

$$t < t_m \quad \Rightarrow \quad r_\gamma(t) > r_\gamma(t_m) \tag{36.79}$$

so that information about objects at larger distances $r > r_H(t_m)$ than the Hubble radius can easily reach us. This distance is bounded only by the *particle horizon* $r_{ph}(t)$ which is what one gets when one tracks $r_\gamma(t)$ back to $t = 0$. Therefore it is precisely the particle horizon which tells us about which points / comoving objects in the constant time spatial surface we can already have obtained information, and not the Hubble radius.

3. The metric induced on the Hubble surface Σ_H defined by $r = r_H(t)$ is

$$\begin{aligned} ds^2|_{\Sigma_H} &= -dt^2 + a(t)^2(dr_H(t)^2 + r_H(t)^2 d\Omega^2) \\ &= -(1 - q(t)^2)dt^2 + R_H(t)^2 d\Omega^2 \quad . \end{aligned} \tag{36.80}$$

Therefore this surface Σ_H is null iff $q = \pm 1$. In these two cases,

$$q = -1 \quad \Rightarrow \quad w = -1 \quad , \quad q = +1 \quad \Rightarrow \quad w = +\frac{1}{3} \quad , \tag{36.81}$$

(i.e. a pure positive cosmological constant or radiation) the Hubble radius agrees with the event horizon or the particle horizon respectively,

$$\begin{aligned} w = -1 \quad &\Rightarrow \quad R_H(t) = R_{eh}(t) = R_H(t_0) \\ w = +\frac{1}{3} \quad &\Rightarrow \quad R_H(t) = R_{ph}(t) \quad . \end{aligned} \tag{36.82}$$

In a universe which is dominated by radiation in the distant past and a cosmological constant in the far future, the Hubble radius will interpolate between the particle hoirzon in the past and the event horizon in the future, but it has no particular significance inbetween.

4. Indeed, more generally one has

$$\Sigma_H \quad \text{is} \quad \begin{cases} \text{spacelike} & \text{for} \quad w > 1/3 \\ \text{null} & \text{for} \quad w = 1/3 \\ \text{timelike} & \text{for} \quad -1 < w < 1/3 \\ \text{null} & \text{for} \quad w = -1 \\ \text{spacelike} & \text{for} \quad w < -1 \end{cases} \tag{36.83}$$

Thus for the physically relevant intermediate range $-1 < w < 1/3$ the Hubble surface is timelike, causally there is nothing strange or interesting going on there, and nothing prevents one from crossing it mulitple times in both directions. In particular it cannot and will not act like a one-way membrane.

36.9 Hubble "Horizon": Is our Universe (like) a Black Hole?

Hinchcliff's rule:

Whenever the title of a paper (section) is a question with a yes/no answer, the answer is "no".

In the previous section we collected some elementary properties of the Hubble radius (Hubble sphere). In particular, we saw that the Hubble sphere is only null in 2 special cases, in which it coincides with the particle or event horizon respectively.

Nevertheless, in spite of these elementary and well-known facts, strangely there is some debate in the current literature about the significance of the Hubble "horizon" (or lack thereof), and even the bizzarre idea that the "visible universe" (defined by $R_H(t)$) is somehow like the inside of a black hole.

As far as I can tell one of the contributing factors to this is the fact that in cosmology, distances to distant objects are commonly expressed in terms of their (somewhat fictitious) instantaneous proper spacelike distance $R(t_0) = a(t_0)r$ from us today, and not (for instance) in terms of their (approximately constant) comoving coordinate distance, or some measure of distance at the time the objects emitted the light that we receive today. In principle, this is perfectly fine, of course, and with due care myths about the Hubble radius can also be exorcised from this point of view.[158]

In practice, however, this use of $R(t)$ leads to a somewhat allegorical (and therefore potentially misleading) way of talking about perfectly mundane things:

- For example, as we have seen above in our discussion of lightrays, one can say (referring to $R_\gamma(t)$) that lightrays reach a maximum distance and then turn around to come towards us, but viewed in terms of comoving coordinates the lightrays just continue in the direction of decreasing r. Thus care needs to be taken to separate the physical motion of objects through space from artefacts arising from describing them in terms of their instantaneous proper spacelike distance.

- Let us consider the spatially flat decelerating case $k = 0$ and $w > -1/3$, so that one has $a(t) \sim t^h$ for some $0 < h < 1$. Then

$$a(t) \sim t^h \quad \Rightarrow \quad R_H(t) = t/h \quad . \tag{36.84}$$

  In this setting one frequently encounters the following kind of argument to "explain" from this point of view why light can reach us from outside the Hubble sphere: for $0 < h < 1$ the Hubble sphere expands faster than the universe; thus

[158] See e.g. P. van Oirschot, J. Kwan, G. Lewis, *Through the Looking Glass: Why the "Cosmic Horizon" is not a horizon*, `arXiv:1001.4795 [astro-ph.CO]`, G. Lewis, P. van Oirschot, *How does the Hubble Sphere limit our view of the Universe?*, `arXiv:1203.0032` for illuminating discussions and enjoyable dissections of these issues from this perspective.

light emitted from a receding galaxy initially outside the Hubble sphere can eventually be overtaken by the Hubble radius and can then become visible to us. True, and very figurative, in terms of some radius overtaking some other radius, but this totally obscures the fact that the Hubble sphere has nothing to do with this and that the only thing that matters for what is visible to us today is if the object is inside the particle horizon or not, i.e. has comoving coordinate $r < r_{ph}(t_0)$ or not.

The above argument also seems to suggest that somehow things change when one has an accelerating universe with $h > 1$, and this is again true, but the thing that changes is that (as we have seen) for $h > 1$ there is simply no particle horizon and therefore no obstruction to seeing objects at any time. This simple fact is again obscured by the above argument.

Nevertheless, as I mentioned already in section 34.3, some people (and not just laymen who understand neither cosmology nor black holes) appear to be obsessed with the idea that our visible universe, defined (counterfactually, as we saw above) as the interior of the Hubble sphere, is somehow like the inside of a Schwarzschild black hole, the Hubble sphere playing the role of its horizon.

This is usually based on some variant of one of the following arguments.[159]

1. Friedmann Equation and the Schwarzschild Radius

   The Friedmann equation
   $$H^2 = \frac{8\pi G_N}{3}\rho - \frac{k}{a^2} \tag{36.85}$$
   can (by dividing by H^3) be written as
   $$R_H = 2G_N \frac{4\pi R_H^3}{3}\rho - \frac{k}{a^2 H^3} \quad . \tag{36.86}$$
   Therefore the statement that $k = 0$ is equivalent to the statement that the Hubble radius R_H is equal to the (would-be) Schwarzschild radius $R_s = 2G_N M$ associated with the total mass
   $$M(R_H) = \frac{4\pi R_H^3}{3}\rho \tag{36.87}$$
   contained in the Hubble sphere,
   $$k = 0 \quad \Leftrightarrow \quad R_H = R_s = 2G_N\, M(R_H) \quad . \tag{36.88}$$

   There may be something profound in this, I don't know, but just saying "hey, it's the Schwarzschild radius, hence I have a black hole" is not!

[159] See e.g. F. Melia (footnote 157) or F. Melia, M. Abdelquader, *The Cosmological Spacetime*, `arXiv:0907.5394 [astro-ph.CO]`.

- In particular, as we have seen above, for reasonable equations of state the "horizon" $r = r_H$ is timelike and we (and lightrays) can simply leave this would-be interior of the black hole. The region we cannot in principle escape from is that bounded by the particle horizon, but the above numerology does not work in that case (unless $w = 1/3$, so that $R_{ph} = R_H$).
- Moreover, as discussed at length in section 32, a black hole is not defined by "mass m enclosed in a sphere of radius $r < 2m$" but by reference to the outside of the black hole and the existence of an *event horizon* and the fact that there is a region of space-time causally sealed off from the asymptotic region.
- Other quasi-local characterisations of black hole-like objects also do not refer directly to mass and radius but to trapped surfaces and the behaviour of lightrays (and more on this below).
- In particular, we will see below that for $k = 0$ (but only for $k = 0$), the Hubble "horizon" is an apparent horizon, but one with properties that are quite different to those of the apparent horizon of a black hole.[160]
- Last but not least, the interior of a black hole exhibits a future spacelike singularity while our universe appears to have emerged from an initial (past) spacelike singularity (and, no, saying "I meant *white* hole, not black hole" will not help, see the discussion below).

2. Writing the FRW Metric in Schwarzschild-like Form

In section 34.4 it was shown that the metric can be written in a form which resembles that of the Schwarzschild metric in Painlevé-Gullstrand-coordinates. Using the current notation ($R = a(t)r$ was called $\tilde{r}$ in section 34.4 and we use $R_H = 1/H$) we can write the general Robertson-Walker metric as (34.43)

$$\begin{aligned} ds^2 &= -(1 - R^2 H(t)^2)dt^2 - 2RH(t)dtdR + (dR^2 + R^2 d\Omega^2) \\ &= -(1 - (R/R_H)^2)dt^2 - 2(R/R_H)dtdR + (dR^2 + R^2 d\Omega^2) \\ &= -\Phi(dt + (R/\Phi R_H)dR)^2 + \Phi^{-1}dR^2 + R^2 d\Omega^2 \quad , \end{aligned} \tag{36.89}$$

where

$$\Phi(t, R) = 1 - (R/R_H(t))^2 \quad . \tag{36.90}$$

This form of the metric shows that there is an infinite time-dilation between the proper time of an observer at constant $R = R_0$ as $R_0 \to R_H$ and the cosmological time t, vaguely suggestive of the same phenomenon in the Schwarzschild metric. However, this is clearly a problem of the observers moving at constant $R(t) =$

[160] Moreover, there is nothing new about the observation (36.88) regarding the relation between the Hubble radius and the Schwarzschild-like radius R_s. It is e.g. already pointed out in section 10.1 of S. Hawking and G. Ellis, *The large scale structure of space-time*, that the numerology leading to (36.88) provides an "intuitive way of viewing this result".

$a(t)r$, not of the space-time itself. These observers at constant R are not the standard comoving observers (at constant r). Their 4-velocity has the form

$$u^t = \Phi(t, R_0)^{-1/2} \quad , \quad u^R = 0 \tag{36.91}$$

in the above (t, R)-coordinates, and

$$u^t = \Phi(t, R_0)^{-1/2} \quad , \quad u^r = \Phi(t, R_0)^{-1/2}(-R_0 H(t)/a(t)) \tag{36.92}$$

in the standard comoving (t, r) coordinates, with

$$\Phi(t, R_0) = 1 - R_0^2 H(t)^2 \ . \tag{36.93}$$

so that $u^t \to \infty$ for $R_0 \to R_H = a(t)/\dot{a}(t)$, while comoving observers have $u^t = 1$. Calculating the acceleration of these observers (in comoving coordinates, say), one finds

$$a^\alpha = u^\beta \nabla_\beta u^\alpha = \Phi^{-2}(q + H^2 R_0^2)(R_0 H^2)(-R_0 H, 1/a) \tag{36.94}$$

with norm

$$g_{\alpha\beta} a^\alpha a^\beta = \Phi^{-3}(q + H^2 R_0^2)^2 (R_0 H^2)^2 \ . \tag{36.95}$$

Clearly this diverges as $R_0 \to R_H = 1/H$, and therefore this is exactly like the space-time seen by ∞ acceleration observers in a Rindler space-time who detect a fictitious Rindler horizon.

Thus R_H is not like an event horizon and there is nothing like a black hole in sight. In fact, the best and most informative way of saying what $R_H(t)$ is and what its significance is (or is not) is that (for $k = 0$) $R_H(t)$ is a spherically symmetric marginally trapped tube (MTT) foliated by spherically symmetric marginally trapped surfaces (MTSs). In other words, with respect to a spherically symmetric foliation of space-time, for $k = 0$ (but only for $k = 0$) the surface $r = r_H(t)$ is an *apparent (3-)horizon* (cf. the discussion in section 32.9). Moreover, as the hypersurface $r = r_H(t)$ is timelike in the range $-1 < w < 1/3$, the horizon-terminology is really not very appropriate and, as mentioned in section 32.9, in current terminology such a hypersurface is referred to as a *timelike membrane*.

Concretely, and in elementary terms, this means the following. Consider (for any k) the equation

$$a(t)\frac{dr}{\sqrt{1 - kr^2}} = \pm dt \tag{36.96}$$

for out- respectively ingoing lightrays, and write this in terms of the area radius $R(t) = a(t)r$,

$$dR_\pm = \sqrt{1 - kR^2/a^2}\left(\pm 1 + RH/\sqrt{1 - kR^2/a^2}\right) dt \ . \tag{36.97}$$

This can of course also be deduced directly from the PG-like metric (36.89) and its $k \neq 0$ counterpart (34.46).

For sufficiently small values of R,

$$R|H|/\sqrt{1-kR^2/a^2} < 1 \quad \Leftrightarrow \quad R(t) < \frac{1}{\sqrt{H(t)^2+k/a(t)^2}} \equiv R_{ah}(t) \quad , \tag{36.98}$$

this exhibits the standard behaviour

$$R(t) < R_{ah}(t) \quad \Rightarrow \quad \begin{cases} dR_+/dt > 0 \\ dR_-/dt < 0 \end{cases} \tag{36.99}$$

that one solution, R_+, propagates to larger values of R, while the other, R_-, propagates to smaller values of R. Evidently something special happens at $R = R_{ah}$ and for larger values one has

$$R(t) > R_{ah}(t) \quad \Rightarrow \quad \begin{cases} H(t) > 0: \quad dR_+/dt > 0 \quad \text{and} \quad dR_-(t)/dt > 0 \\ H(t) < 0: \quad dR_+/dt < 0 \quad \text{and} \quad dR_-(t)/dt < 0 \end{cases} \tag{36.100}$$

Thus, for a contracting universe this means that the spheres of constant R and t are *trapped* for $R > R_{ah}(t)$, like the 2-spheres inside the event horizon of the Schwarzschild black hole, with both ingoing and (would-be) outgoing radial lightrays moving towards smaller radii. However, far from indicating the presence of a black hole in this case, these trapped spheres and the apparent horizon $R_{ah}(t)$ indicate (together with an energy condition) a future cosmological singularity.

In an expanding universe, on the other hand, for $R > R_{ah}$ both in- and outgoing lightrays move to larger values of R. The spheres with $R > R_{ah}$ are thus the opposite of trapped surfaces, i.e. anti-trapped or trapped towards the past (and reflect the existence of a big bang singularity in the past).

The fact that both in- and outgoing expansions are positive is also a characteristic feature of the (unphysical) "white hole" region of a black hole in the region before the past event horizon. However, in that case it is the region around $r = 0$, and with $r < 2m$, that contains the anti-trapped surfaces while sufficiently large spheres show normal behaviour. Here, on the contrary, it is sufficiently large spheres around the comoving observer that are anti-trapped while in a sufficiently small region around any cmoving observer there are no (anti-)trapped surfaces.

Noting that for $k = 0$ the apparent horizon is equal to the Hubble radius,

$$k = 0 \quad \Rightarrow \quad R_{ah}(t) = R_H(t) \quad , \tag{36.101}$$

all of this is simply a restatement of the discussion about the behaviour of lightrays in the previous section, and nothing seems to be gained by coining a new name for this well-established concept of an apparent horizon. Moreover, for $k \neq 0$ the Hubble sphere is not even an apparent horizon or some other marginally trapped tube, so it seems best not to associate the word "Hubble" with the word "horizon" at all.

37 Cosmology V: Some Exact Solutions

We have seen that a lot can be learnt about the Friedmann-Lemaître-Robertson-Walker models without ever having to solve a differential equation. On the other hand, more precise information can be obtained by specifying an equation of state for the matter content and solving explicitly the Friedmann equations. We have already seen some special exact solutions ($k = \Lambda = 0$ and only one species of matter) in section 36.4. Solutions with $k \neq 0$ but $\Lambda = 0$ are also easy to obtain. Together with the exact solutions for just $k \neq 0$ and $\Lambda \neq 0$, but with no other matter, these are the most useful solutions in practice, but it is also possible to solve the equations explicitly in some other cases.[161]

37.1 Milne Universe

Let us start our excursion into exact solutions by looking at the totally unphysical case of a completely empty universe, i.e. $\rho = p = \Lambda = 0$, and only k possibly not zero. As trivial as this may be, it has its pedagogical value, which is why I am including this case here.

In this case (F1) reduces to

$$\dot{a}^2 + k = 0 \quad \Rightarrow \quad k \leq 0 \ . \tag{37.1}$$

For $k = 0$, one has $\dot{a} = 0$ and thus $a(t) = a_0$, and the metric is

$$ds^2 = -dt^2 + a_0^2 d\vec{x}^2 \ . \tag{37.2}$$

It should not come as a suprise that the metric we have found is just that of Minkowski space (the constant a_0 can be absorbed into a rescaling of the spatial coordinates).

The case $k = -1$ is a bit more interesting. In that case the equation to solve is (I will call the time-coordinate τ instead of t for reasons that will perhaps become apparent below)

$$\dot{a}^2 = +1 \quad \Rightarrow \quad a(\tau) = \pm(\tau - \tau_0) \ . \tag{37.3}$$

and the resulting space-time metric is (choosing without loss of generality $\tau_0 = 0$)

$$ds^2 = -d\tau^2 + \tau^2 \, d\tilde{\Omega}_3^2 \tag{37.4}$$

where $d\tilde{\Omega}_3^2$ is the line-element on the unit 3-hyperboloid (14.33).

This appears to describe a non-trivial universe (known as the *Milne universe*) with either (for $a(\tau) = +\tau$) a big bang at $\tau = 0$ and with the universe subsequently expanding

[161] For explicit solutions of the Friedmann equations for $\Lambda \neq 0$ and non-trivial matter content (generalising the solution (38.26)) see e.g. B. Aldrovandi, R. Cuzinatto, L. Medeiros, *Analytic solutions for the Λ-FRW Model*, `arXiv:gr-qc/0508073`.

linearly with $\tau > 0$, or (for $a(\tau) = -\tau$) with a linearly contracting universe for $\tau < 0$, ending in a "big crunch" at $\tau = 0$, and all this in spite of the fact that the universe is empty.

This is deceptive, however (not that it is empty, but that it is non-trivial). As always, we need to be careful to disentangle coordinate artefacts from genuine geometrical statements. Indeed, this space-time is again nothing other than (a part of) Minkowski space-time. We had already anticipated this in section 35.1, in connection with (35.9) which expresses the Riemann tensor of a Robertson-Walker metric in terms of its Ricci tensor (and which therefore implies that a vacuum solution will necessarily be flat).

Nevertheless, it will be instructive to see this explicitly. To that end we start with the Minkowski metric

$$ds^2 = -dt^2 + d\vec{x}^2 \tag{37.5}$$

and introduce coordinates that are adapted to the family of space-like hyperboloids

$$t^2 - \vec{x}^2 = \tau^2 > 0 \quad . \tag{37.6}$$

For $t > 0$ these hyperboloids fill (foliate) the interior of the future lightcone at the origin (and the interior of the past lightcone for $t < 0$), and if you draw these hyperboloids you can perceive what appears to be a non-trivial dynamical evolution of these surfaces. To show that this "fake" dynamics is precisely the dynamics exhibited by the Milne metric, introduce new coordinates $(\tau, \rho, \text{angles})$ via

$$t = \tau \cosh\rho \quad , \quad \vec{x} = \vec{n}\tau \sinh\rho \tag{37.7}$$

where $\vec{n}$ is a unit vector on S^2, so that

$$\vec{n}.\vec{n} = 1 \quad , \quad \vec{n}.d\vec{n} = 0 \quad , \quad d\vec{n}.d\vec{n} = d\Omega_2^2 \quad . \tag{37.8}$$

Then the Minkowski metric becomes precisely the Milne metric (37.4),

$$ds^2 = -d\tau^2 + \tau^2 d\rho^2 + \tau^2 \sinh^2\rho \, d\Omega_2^2 \quad . \tag{37.9}$$

Remarks:

1. These coordinates only cover a part of Minkowski space-time, namely the interior of the future (and past) lightcone of the origin. They are adapted to the comoving (and thus geodesic) observers of the Milne metric, i.e. to families of observers with constant values $\rho = \rho_0$, $\vec{n} = \vec{n}_0$. Thus in terms of Minkowski coordinates, the worldlines of these observers are described by

$$t(\tau) = \tau \cosh\rho_0 \quad , \quad \vec{x}(\tau) = \vec{n}_0 \tau \sinh\rho_0 \tag{37.10}$$

or

$$\vec{x}(t) = (\vec{n}_0 \tanh\rho_0)t \equiv \vec{v}_0 t \tag{37.11}$$

with

$$(v_0)^2 = \tanh^2 \rho_0 \leq 1 \ . \tag{37.12}$$

Thus these are geodesic observers emanating radially with different initial directions and velocities from the origin, the constant-time hyperboloids of the Milne metric being the surfaces which these observers reach at their proper time τ.

2. The coordinate singularity at $\tau = 0$ is simpy due to the fact that the worldlines of these observers all intersect at the origin (and thus do not provide good coordinates there).

3. These coordinates are, as may have occurred to you by now, the future and past relatives of the Rindler coordinates discussed (way back) in sections 1.3 and 3.4 (and then again in the context of the near-horizon geometry of the Schwarzschild metric in section 26.6), in particular of the spherical Rindler space discussed at the end of section 3.4:

   Introducing Rindler-like coordinates $(\tau, \rho, \text{angles})$ adapted to the timelike hyperboloids

   $$\vec{x}^2 - t^2 = \rho^2 \tag{37.13}$$

   via

   $$t = \rho \sinh \tau \quad , \quad \vec{x} = \vec{n}\rho \cosh \tau \ , \tag{37.14}$$

   with $\vec{n}$ a spatial unit vector, $\vec{n}.\vec{n} = 1$, the Minkowski metric takes the form (3.100)

   $$\begin{aligned} ds^2 &= -\rho^2 d\tau^2 + d\rho^2 + \rho^2 \cosh^2 \tau \, d\Omega_2^2 \\ &= d\rho^2 + \rho^2(-d\tau^2 + \cosh^2 \tau \, d\Omega_2^2) \ . \end{aligned} \tag{37.15}$$

   This is the spherical Rindler metric. The metric in brackets in the 2nd line will reappear below, in the guise of the (2+1)-dimensional de Sitter metric.

4. Thus, we see that we can foliate Minkowski space in different ways:

   - in the standard inertial coordinates, slices of constant Minkowski time t are Euclidean spaces $\mathbb{R}^3$, and slices of constant x^3, say, are Minkowski spaces $\mathbb{R}^{1,2}$;
   - in terms of Minkowski time and spatial spherical coordinates,

     $$ds^2 = -dt^2 + dr^2 + r^2 \, d\Omega_2^2 \tag{37.16}$$

     slices of constant r are of the form $S^2 \times \mathbb{R}$ (and we will recognise these below, in more fancy terminology, as (2+1)-dimensional versions of the Einstein static universe);
   - in Milne coordinates, slices of constant τ are hyperboloids H^3;
   - in the Rindler-like coordinates (37.15), slices of constant ρ are are de Sitter spaces dS_3.

5. The "singularity" of the Milne metric is of course just a coordinate singularity, but with some discrete identifications on the hyperboloid this becomes an honest conical singularity at $\tau = 0$. This is a nice (and much studied) toy-model of a big bang (everywhere flat except at $\tau = 0$).

37.2 Einstein Static Universe

This particular solution is mainly of historical interest. Before the discovery and understanding of the Hubble expansion of the universe, it was natural to assume that the universe can be described by a static solution compatible with the cosmological principle (homogeneity and isotropy). This led Einstein to introduce the (in-)famous cosmological constant. Let us see how this comes about.

Static means that $a(t) = a_0, \dot{a}(t) = \ddot{a}(t) = 0$. (F2) with $\ddot{a} = 0$ and a zero cosmological constant leads to

$$\text{(F2)}(\Lambda = 0) \quad \Rightarrow \quad 4\pi G_N(\rho + 3p) = 0 \ . \tag{37.17}$$

This equation is simply not satisfied for ordinary matter, and therefore no static solution of the Friedmann equations with $\Lambda = 0$ exists. (Re-)introducing Λ, however, (F2) then tells us that

$$\text{(F2)} \quad \Rightarrow \quad 4\pi G_N(\rho + 3p) = \Lambda \ . \tag{37.18}$$

Moreover, the conservation law (F3) with $\dot{a} = 0$ implies

$$\text{(F3)} \quad \Rightarrow \quad \dot{\rho} = 0 \ , \tag{37.19}$$

which now implies that also $\dot{p} = 0$. (37.18) then simply fixes the constant Λ in terms of the constants ρ and p. For a standard matter content of the universe, say $\rho = \rho_m + \rho_r$, this requires that $\Lambda > 0$. Finally, the first Friedmann equation (F1) now becomes an algebraic equation for $a(t) = a_0$, namely

$$\begin{aligned} \text{(F1)} \quad &\Rightarrow \quad \frac{k}{a^2} = \frac{8\pi G_N}{3}\rho + \frac{\Lambda}{3} = 4\pi G_N(\rho + p) \\ &\Rightarrow \quad k = +1 \quad \text{and} \quad a_0^2 = (4\pi G_N(\rho + p))^{-1} \ . \end{aligned} \tag{37.20}$$

This is thus a static solution of the Friedmann equations,

$$ds^2 = -dt^2 + a_0^2 \, d\Omega_3^2 \ , \tag{37.21}$$

the *Einstein Static Universe*.

Remarks:

1. The topology of the solution is $\mathbb{R} \times S^3$, the radius a_0 of the S^3 being smaller for larger energy and pressure density (bigger gravitational attraction) and vice-versa.

2. In this solution, the gravitational attraction is precisely balanced by the positive cosmological constant. Note that even though a positive cosmological constant has a positive energy density $\rho_\Lambda > 0$, it has a negative pressure, and the net effect of a positive cosmological constant in (F2) is that of gravitational repulsion rather than attraction,
$$\rho_\Lambda + 3p_\Lambda = -2\rho_\Lambda < 0 \quad . \tag{37.22}$$

3. The precise balance between matter and cosmological constant required by this solution also implies that it is unstable to perturbations of either Λ or ρ, p: a slight increase in Λ relative to $\rho + p$, no matter how small, will make the universe expand, and a slight decrease will make it collapse. This alone is enough to make this particular solution unphysical and shows that, even with inclusion of a cosmological constant, an expanding or collapsing universe is practically inevitable, thus undermining the original motivation for introducing the cosmological constant in the first place.

37.3 Matter Dominated Era

We now return to somewhat more realistic solutions of the Friedmann equations. In a matter dominated era (but retaining the curvature term) we have to solve the equation
$$\dot{a}^2 = \frac{C_m}{a} - k \quad . \tag{37.23}$$

We will now look in turn at the cases $k = 0, +1, -1$.

1. For $k = 0$, this is the equation we already discussed above, leading to the solution (36.20),
$$a(t) = a_0(t/t_0)^{2/3} \tag{37.24}$$
This solution is also known as the Einstein - de Sitter universe. It describes a universe expanding from an initial singularity at $t = 0$. Another solution, arising from taking the negative square root of the Friedmann equation $\dot{a}^2 = C_m/a$,
$$a^{1/2}\dot{a} = -(C_m)^{1/2} \quad \Rightarrow \quad a(t) = a_0(t_f - t)^{2/3}/(t_f - t_0)^{2/3} \quad , \tag{37.25}$$
describes a universe collapsing to a big crunch at $t = t_f$.

2. For $k = +1$, the equation is
$$\dot{a}^2 = \frac{C_m}{a} - 1 \quad . \tag{37.26}$$
We recall that in this case we will have a recollapsing universe, with $a_{max} = C_m$ attained for $\dot{a} = 0$,
$$\dot{a} = 0 \quad \Rightarrow \quad a = a_{max} = C_m \quad . \tag{37.27}$$

The equation (37.26) can be solved in closed form for t as a function of a, and the solution to

$$\frac{dt}{da} = (\frac{a}{a_{max} - a})^{1/2} \tag{37.28}$$

is

$$t(a) = \frac{a_{max}}{2} \arccos(1 - 2a/a_{max}) - \sqrt{aa_{max} - a^2} \quad , \tag{37.29}$$

as can easily be verified. The universe starts at $t = 0$ with $a(0) = 0$, reaches its maximum $a = a_{max}$ at

$$t_{max} = a_{max} \arccos(-1)/2 = a_{max}\pi/2 \quad , \tag{37.30}$$

and ends in a Big Crunch at $t = 2t_{max}$.

The curve $a(t)$ is a cycloid, as is most readily seen by writing the solution in parametrised form. For this it is convenient to use the conformal time coordinate η introduced in (34.48) through

$$d\eta = dt/a(t) \quad \Rightarrow \quad \partial_t = a(\eta)^{-1}\partial_\eta \quad , \quad \partial_\eta = a(t)\partial_t \quad . \tag{37.31}$$

Denoting a derivative with respect to η by a prime, and noting that $\dot{a} = a'/a$, one then finds that for $k \neq 0$ the Friedman equation (37.23) can be written as

$$\begin{aligned} \dot{a}^2 + k = \frac{C_m}{a} \quad &\Leftrightarrow \quad (a')^2 + ka^2 = C_m a \\ &\Leftrightarrow \quad ((a - C_m/2k)')^2 + k(a - C_m/2k)^2 = kC_m^2/4 \quad . \end{aligned} \tag{37.32}$$

Thus for $k = +1$, the solution to the Friedmann equation can be written as

$$a(\eta) - C_m/2 = (C_m/2)\cos(\eta - \eta_0) \quad . \tag{37.33}$$

Choosing $\eta_0 = \pi$ (so that $a(\eta = 0) = 0$), and integrating the relation $dt/d\eta = a(\eta)$ to find $t(\eta)$, one then finds the solution

$$\begin{aligned} a(\eta) &= \frac{a_{max}}{2}(1 - \cos\eta) \\ t(\eta) &= \frac{a_{max}}{2}(\eta - \sin\eta) \quad , \end{aligned} \tag{37.34}$$

which makes it transparent that the curve is indeed a cycloid, roughly as indicated in Figure 46.

The maximal radius is reached at

$$t_{max} = t(a = a_{max}) = t(\eta = \pi) = a_{max}\pi/2 \tag{37.35}$$

(with $a_{max} = C_m$), as before, and the total lifetime of the universe is $2t_{max}$.

3. Analogously, for $k = -1$ the Friedmann equation in parametrised form (37.32) can be solved in terms of hyperbolic (rather than trigonometric) functions,

$$\begin{aligned} a(\eta) &= \frac{C_m}{2}(\cosh\eta - 1) \\ t(\eta) &= \frac{C_m}{2}(\sinh\eta - \eta) \quad . \end{aligned} \tag{37.36}$$

Remarks:

1. We see that for small times (for which matter dominates over curvature) the solutions for $k \neq 0$ reduce to $t \sim \eta^3, a \sim \eta^2$ and therefore $a \sim t^{2/3}$ which is indeed the exact solution for $k = 0$.

2. Analogously, for late times in the $k = -1$ model one finds that $a(\eta) \sim t(\eta)$, reproducing the expected late-time behaviour $\dot{a} \to 1$ of section 36.3.

3. In section 29 we will use the exact solutions of this matter dominated phase to describe the interior geometry of collapsing stars.

It is instructive to display the causal structure of the $k = +1$ metric in a conformal diagram. Recall that in terms of conformal time the metric takes the form

$$ds^2 = a(\eta)^2(-d\eta^2 + d\Omega_3^2) = a(\eta)^2(-d\eta^2 + d\psi^2 + \sin^2\psi d\Omega_2^2) \quad , \tag{37.37}$$

Suppressing the transverse 2-sphere and eliminating the conformal factor $a(\eta)^2$, we are thus led to consider the metric

$$d\tilde{s}^2 = -d\eta^2 + d\psi^2 \quad . \tag{37.38}$$

Noting that in the case at hand the range of both η and ψ is finite,

$$\eta \in [0, 2\pi] \quad , \quad \psi \in [0, \pi] \quad , \tag{37.39}$$

we see that no further redefinition of the coordinates is required to obtain a conformal diagram of this solution (Figure 51).

Here are some of the characteristic features of this diagram:

1. Since this solution is spatially compact, there is no analogue here of spatial infinity i^0. As a consequence, this diagram a priori looks very different from Penrose diagrams for asymptotically flat space-times, in particular also from those for spatially flat cosmologies in section 36.6.

2. All timelike and null geodesics begin at the initial (spacelike) singularity at $\eta = 0$ and end at the final singualrity at $\eta = 2\pi$.

3. A lightray sent out at the Big Bang will reach the antipodal point of the sphere exactly at the time $\eta_{max} = \pi$ the universe reaches its maximal radius, and will have circled around the universe exactly once precisely at the time $2\eta_{max}$ of the final big crunch.

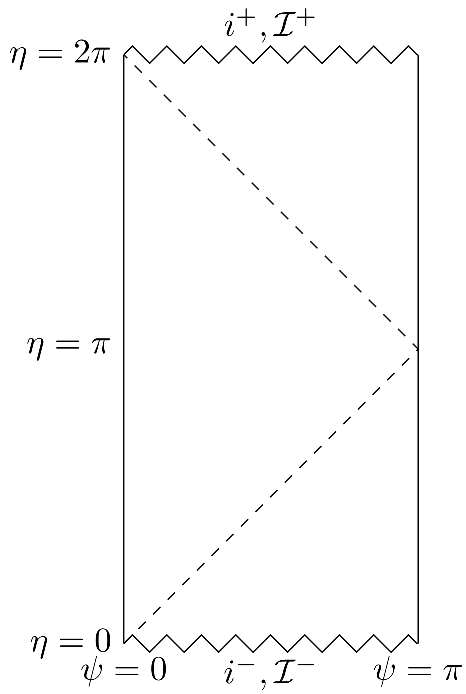


Figure 51: Conformal Diagram of the $k = +1$ matter dominated universe.

37.4 Radiation Dominated Era

We now consider the situation when radiation is dominant (as is expected during some time in the very early universe). In this case we need to solve

$$a^2\dot{a}^2 = C_r - ka^2 \quad . \tag{37.40}$$

For $k = 0$ we had already seen the solution in (36.18),

$$a(t) = a_0(t/t_0)^{1/2} \tag{37.41}$$

(and there is also evidently a corresponding collapsing solution). For $k = +1$, on the other hand, there will be a maximal radius a_{max} at

$$k = +1 : \quad \dot{a} = 0 \quad \Rightarrow \quad a_{max} = C_r^{1/2} \quad . \tag{37.42}$$

Here are 2 ways to solve the Friedmann equation for $k \neq 0$:

1. Because $a(t)$ appears only quadratically, it is convenient to make the change of variables $b = a^2$. Then one obtains
$$\frac{\dot{b}^2}{4} + kb = C_r \quad . \tag{37.43}$$
For $k = \pm 1$, one necessarily has $b(t) = b_0 + b_1 t + b_2 t^2$. Fixing $b(0) = 0$, one easily finds the solution
$$a(t) = [2C_r^{1/2}t - kt^2]^{1/2} \quad . \tag{37.44}$$

As expected this reduces to $a(t) \sim t^{1/2}$ for small times where the curvature term is irrelevant.

- For $k = +1$ one has
$$a(0) = a(2C_r^{1/2}) = 0 \quad , \tag{37.45}$$
and $a = a_{max}$ at the time
$$t_{max} = C_r^{1/2} \quad \Rightarrow \quad a(t_{max}) = a_{max} = C_r^{1/2} \quad . \tag{37.46}$$
- For $k = -1$, on the other hand, the universe expands forever, the late-time behaviour being given by
$$a(t) \to (-kt^2)^{1/2} = t \quad , \tag{37.47}$$
again as expected.

All this is of course in agreement with the results of the qualitative discussion given earlier.

2. It is also instructive to solve these equations in terms of conformal time η. Using $\dot{a} = a'/a$, the Friedmann equation becomes
$$(a')^2 + ka^2 = C_r \quad . \tag{37.48}$$
Thus the solution for $k = +1$ is obviously (with the choice $a(\eta = 0) = 0$)
$$k = +1 \quad \Rightarrow \quad a(\eta) = C_r^{1/2} \sin\eta \tag{37.49}$$
and for $k = -1$ one has
$$k = -1 \quad \Rightarrow \quad a(\eta) = C_r^{1/2} \sinh\eta \quad . \tag{37.50}$$
Using $dt/d\eta = a(\eta)$ one can also find $t(\eta)$. E.g. for $k = +1$ one has
$$t(\eta) = C_r^{1/2}(1 - \cos\eta) \quad , \tag{37.51}$$
where the integration constant has been chosen such that
$$\eta = 0 \quad \Rightarrow \quad a(\eta = 0) = 0 \quad , \quad t(\eta = 0) = 0 \quad . \tag{37.52}$$
One also has $\eta_{max} = \pi/2$ and the recollapse to $a(\eta) = 0$ at $\eta = \pi$ corresponding to $t = 2C_r^{1/2}$, as in the 1st derivation.

Again it is instructive to display the $k = +1$ solution in a conformal diagram (Figure 52) and to compare it with that of the matter dominated solution (Figure 51). The main difference is that here the range of η is equal to the range of ψ,
$$\eta \in [0, \pi] \quad , \quad \psi \in [0, \pi] \quad . \tag{37.53}$$
As a consequence, the conformal diagram is a square, and any lightray sent out at the Big Bang can only travel half-way around the universe during the lifetime of the universe.

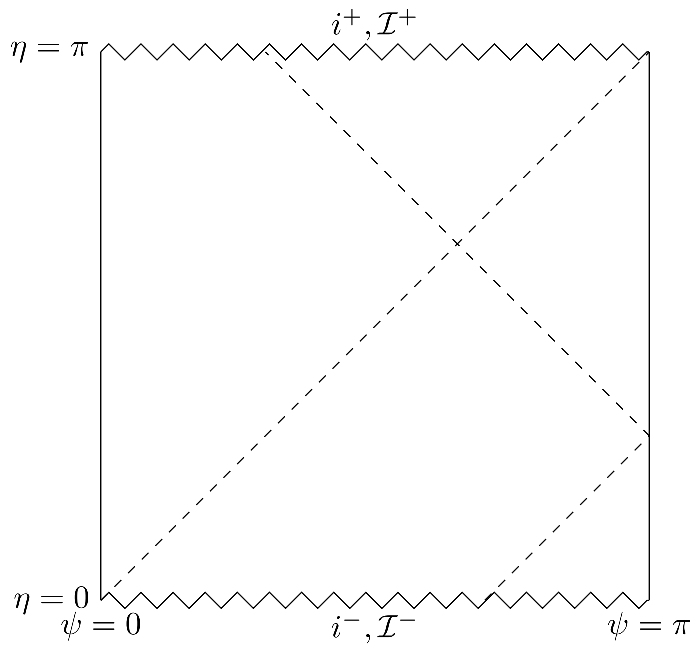


Figure 52: Conformal Diagram of the $k = +1$ radiation dominated universe.

37.5 Cosmological Constant Dominated Era: (Anti-) de Sitter Space

This case is of considerable interest for at least two reasons. On the one hand, as we know, Λ is the dominant driving force for $a(t)$ very large, and may therefore, if current observations are to be believed (see section 38.1), dominate the late-time behaviour of our universe.

On the other hand, the currently most popular cosmological models trying to also address and solve the so-called *horizon problem* and *flatness problem* (cf. the discussion in section 38.3) of the standard FLRW model of cosmology (as well as a number of other issues) use a mechanism called *inflation* based on an era of exponential expansion during some time in the very early universe. This is typically generated by something that acts effectively like a cosmological constant.

Thus the equations to solve are, setting $\rho = 0$ and $p = 0$ but retaining Λ and k,

$$\dot{a}^2 = -k + \frac{\Lambda}{3}a^2 \quad . \tag{37.54}$$

We see immediately that Λ has to be positive for $k = +1$ or $k = 0$, whereas for $k = -1$ both positive and negative Λ are possible,

$$\dot{a}^2 = -k + \frac{\Lambda}{3}a^2 \quad \Rightarrow \quad \begin{cases} \text{if } \Lambda > 0 & k = 0, \pm 1 \text{ possible} \\ \text{if } \Lambda < 0 & k = -1 \end{cases} \tag{37.55}$$

This is one instance where the solution to the second order equation (F2),

$$\ddot{a} = \frac{\Lambda}{3} a \quad , \tag{37.56}$$

is more immediate, namely trigonometric functions for $\Lambda < 0$ (only possible for $k = -1$) and hyperbolic functions for $\Lambda > 0$. The first order equation then fixes the constants of integration according to the value of k.

It will be convenient to express Λ in terms of a length-scale ℓ (which will turn out to be the curvature radius of the space-time) through

$$|\Lambda|/3 = \ell^{-2} \tag{37.57}$$

so that the two Friedmann equations read

$$\dot{a}^2 = -k \pm a/\ell^2 \quad , \quad \ddot{a} = \pm a/\ell^2 \quad , \tag{37.58}$$

the plus/minus referring to $\Lambda > 0$ and $\Lambda < 0$ respectively.

<u>SOLUTIONS FOR $\Lambda > 0$:</u>

1. $k = 0$: The solutions are evidently (cf. (36.23))

   $$a_\pm(t) \sim \mathrm{e}^{\pm t/\ell} \tag{37.59}$$

   leading to the two metrics (related by $t \to -t$)

   $$ds^2 = -dt^2 + \mathrm{e}^{\pm 2t/\ell} d\vec{x}^2 \tag{37.60}$$

2. $k = +1$: The solution is

   $$ds^2 = -dt^2 + \ell^2 \cosh^2 t/\ell \; d\Omega_3^2 \quad . \tag{37.61}$$

3. $k = -1$: The solution is

   $$ds^2 = -dt^2 + \ell^2 \sinh^2 t/\ell \; d\tilde{\Omega}_3^2 \quad . \tag{37.62}$$

<u>REMARKS:</u>

1. It turns out that all 3 metrics actually represent the same space-time metric, just written in different coordinates. This space-time is known as the *de Sitter space-time*. Thus these 3 metrics exhibit different slicings of the de Sitter space-time (or, henceforth, de Sitter space for short, or just dS space), with the $t =$ const. slices being $\mathbb{R}^3$, S^3 and H^3 respectively.

2. One way to establish this would be to directly exhibit the coordinate transformations that map one metric to the other, but this is messy and does not provide any additional insight. We will proceed in a different way in section 39 below. Indeed, it turns out that the de Sitter solution is the unique maximally symmetric spacetime with positive curvature (cf. the discussion in sections 14 and 39), and this perspective will provide us with a more efficient and insightful way of constructing different coordinate systems and exploring the relations among them.

3. All 3 metrics exhibit an exponential expansion or collapse at early and/or late times. The $k = -1$ metric also appears to exhibit a big bang singularity as $t \to 0$, but this is an illusion. Indeed, as $t \to 0$ the metric approaches the metric (37.4) of the Milne universe, and we already know that the singularity at $\tau = 0$ of the Milne metric is just a coordinate singularity.

4. The coordinates appearing in the $k = +1$ metric turn out to cover de Sitter space globally. Thus the global picture of de Sitter space is that of a 3-sphere that

   - started out with infinite radius in the infinite past $t \to -\infty$,
   - undergoes an exponential contraction to a sphere of radius ℓ at time $t = 0$
   - and then again undergoes an exponential expansion to infinite size as $t \to +\infty$.

   It is clear from this description that e.g. the 2 metrics for $k = 0$ only cover the contracting (-) or expanding (+) period of the de Sitter universe.

5. For later purposes it is useful to write the expanding $k = 0$ metric in terms of conformal time η (cf. sections 34.5 and 36.6) as

$$ds^2 = \mathrm{e}^{2t/\ell}(-dt^2\mathrm{e}^{-2t/\ell} + d\vec{x}^2) = \frac{\ell^2}{\eta^2}(-d\eta^2 + d\vec{x}^2) \tag{37.63}$$

   where

$$\eta = -\ell\mathrm{e}^{-t/\ell} \quad , \tag{37.64}$$

   with $\eta \to -\infty$ for $a(t(\eta)) \to 0$. This clearly exhibits the simple Minkowskian causal structure of de Sitter space in the expanding phase / patch.

6. The (2+1)-dimensional counterpart of the $k = +1$ metric is the metric that appeared in the Rindler-like coordinate system (37.15). Thus we can say that these Rindler-like coordinates provide a (radial) foliation of the left- and right Rindler wedges by de Sitter spaces.

7. A detailed discussion of many different coordinate systems for de Sitter space is given in section 39.2 below.

Solution for $\Lambda < 0$:

In this case, the only possibility (within the Robertson-Walker ansatz for the metric) is $k = -1$, and the solution is

$$ds^2 = -dt^2 + \ell^2 \sin^2 t/\ell \; d\tilde{\Omega}_3^2 \quad . \tag{37.65}$$

Remarks:

1. This solution is nowadays known as the *anti-de Sitter space-time* (or *anti-de Sitter space* or AdS space for short).

2. As its $\Lambda > 0, k = -1$ counterpart, it appears to exhibit a big bang singularity as $t \to 0$, but this is again an illusion since this metric also approaches the non-singular Milne metric (37.4) as $t \to 0$.

3. AdS space turns out to be the unique maximally symmetric space-time with negative curvature.

4. The above cosmological coordinates do not provide a complete covering of anti-de Sitter space. In (one choice of) global coordinates the metric turns out to take the form (39.75)
$$ds^2 = -\cosh^2 \rho d\tau^2 + d\rho^2 + \sinh^2 \rho d\Omega_2^2 \tag{37.66}$$
with $0 \leq \rho < \infty$ and $-\infty < \tau < \infty$. In these coordinates the metric is time-independent and therefore (unlike dS) AdS has a global timelike Killing vector, namely ∂_τ.

5. A detailed discussion of many different coordinate systems for anti-de Sitter space is given in section 39.3 below.

38 Cosmology VI: The Universe Today - Insights and Puzzles

38.1 Λ-CDM Cold Dark Matter + Dark Energy Model

Let us recall the key equations governing the evolution of the universe which, in the notation introduced in section 36.5, take the compact (albeit somewhat obscure) form

$$\begin{aligned} \Omega_M + \Omega_\Lambda + \Omega_k &= 1 \\ \tfrac{1}{2}(1+3w)\Omega_M - \Omega_\Lambda &= q \end{aligned} \tag{38.1}$$

Their version today (at $t = t_0$) is (36.40)

$$\begin{aligned} (\Omega_M)_0 + (\Omega_\Lambda)_0 + (\Omega_k)_0 &= 1 \\ \tfrac{1}{2}(1+3w_0)(\Omega_M)_0 - (\Omega_\Lambda)_0 &= q_0 \end{aligned} \tag{38.2}$$

In the universe today, the radiation contribution to the matter content is negligible and the only non-negligible matter content appears to be that of $w = 0$ pressureless matter, and thus

$$q_0 = \tfrac{1}{2}(\Omega_M)_0 - (\Omega_\Lambda)_0 \quad . \tag{38.3}$$

In that case Ω_M is related to $(\Omega_M)_0$ by

$$\Omega_M = \frac{\Omega_M}{(\Omega_M)_0}(\Omega_M)_0 = \frac{\rho_M}{(\rho_M)_0}\frac{(\rho_{cr})_0}{\rho_{cr}}(\Omega_M)_0 = \frac{a_0^3}{a^3}\frac{H_0^2}{H^2}(\Omega_M)_0 \quad . \tag{38.4}$$

Likewise, for the cosmological constant and curvature contributions one has

$$\Omega_\Lambda = \frac{H_0^2}{H^2}(\Omega_\Lambda)_0 \quad , \quad \Omega_k = \frac{a_0^2 H_0^2}{a^2 H^2}(\Omega_k)_0 \quad . \tag{38.5}$$

Thus the Friedmann equation can be written as

$$H^2 = H_0^2\left(\frac{a_0^3}{a^3}(\Omega_M)_0 + \frac{a_0^2}{a^2}(\Omega_k)_0 + (\Omega_\Lambda)_0\right) \quad . \tag{38.6}$$

Using the constraint $(\Omega_M)_0 + (\Omega_\Lambda)_0 + (\Omega_k)_0 = 1$ to eliminate $(\Omega_k)_0$ from this equation, one finds

$$H^2 = \frac{H_0^2 a_0^2}{a^2}\left(1 + (\frac{a_0}{a} - 1)(\Omega_M)_0 + (\frac{a^2}{a_0^2} - 1)(\Omega_\Lambda)_0\right) \quad . \tag{38.7}$$

This equation is also frequently written as an equation for H as a function of the redshift z, $H = H(z)$, which can be obtained by the substitution $a_0/a = 1 + z$,

$$H(z) = H_0(1+z)\left[1 + (\Omega_M)_0 z + (\Omega_\Lambda)_0\left(\frac{1}{(1+z)^2} - 1\right)\right]^{1/2} \quad . \tag{38.8}$$

This expression is useful because it is expressed directly in terms of observable quantities, with $H(z)$ the value of the Hubble parameter at the time an object emitted the light that we now observe with redshift z. Using

$$1 + z = \frac{a_0}{a(t)} \quad \Rightarrow \quad \frac{dz}{H(z)} = -\frac{a_0 dt}{a(t)} \tag{38.9}$$

it is also possible to convert other evolution equations in t (or integrals over t) into evolution equations in z (or integrals over z).

For a long time it was believed that $\Omega_\Lambda = 0$ (or at least negligibly small) today. In that case one would have

$$(\Omega_M)_0 + (\Omega_k)_0 = 1 \tag{38.10}$$

and

$$q_0 = \tfrac{1}{2}(\Omega_M)_0 > 0 \ . \tag{38.11}$$

Therefore the curvature parameter k would be directly related to the value q_0 of the deceleration parameter today:

$$\begin{aligned} q_0 > 1/2 \quad &\Rightarrow \quad \rho_0 > (\rho_{cr})_0 \quad \Rightarrow \quad k = +1 \\ q_0 < 1/2 \quad &\Rightarrow \quad \rho_0 < (\rho_{cr})_0 \quad \Rightarrow \quad k = -1 \end{aligned} \tag{38.12}$$

While there were large uncertainties about q_0, observations indicated a value of ρ_0 much smaller than the critical density $(\rho_{cr})_0$ (even taking into account not just visible matter but also some invisible ("dark") matter whose existence had been deduced indirectly from astrophysical observations like the rotations of galaxies). Thus this strongly suggested a decelerating open $k = -1$ universe. While perhaps not the most hospitable place in the long run, at least this scenario had the virtue of simplicity.

However, exciting recent developments and observations in cosmology and astrophysics have provided strong evidence for a very different and extremely intriguing and puzzling picture of the universe today. Since I am not an expert on these matters, I will just summarise the results here.[162]

1. Estimates for the current matter contribution $\Omega_M = \sum_b \Omega_b$ are

$$(\Omega_M)_0 \approx 0.3 \ . \tag{38.13}$$

2. Ordinary (visible, baryonic) matter only accounts for a small fraction of this, namely

$$(\Omega_{M,\text{visible}})_0 \approx 0.04 \ . \tag{38.14}$$

   Most of the matter density of the universe must therefore be due to some form of (as yet ill-understood, non-relativistic, weakly interacting) *Dark Matter* or *Cold Dark Matter* (CDM).

3. A detailed analysis of the fine structure of the anisotropies of the *Cosmic Microwave Background* suggests that the universe is spatially flat, $k = 0$, i.e. that the total density of whatever fills the universe is very close to the critical density.

[162] For excellent recent introductions to cosmology which explain the significance of these results, how they were obtained and much more, see e.g. M. Trodden and S. Carroll, *TASI Lectures: Introduction to Cosmology*, `arXiv:astro-ph/0401547`, and J. Frieman, M. Turner, D. Huterer, *Dark Energy and the Accelerating Universe*, `arXiv:0803.0982v1 [astro-ph]`.

4. Observations of redshifts of particular types of *Supernovae* in distant galaxies lead to the conclusion that the universe must currently be in a phase of *accelerated expansion*, with the data well explained by assuming that the *Dark Energy* responsible for this acceleration (thus with equation of state parameter $w < -1/3$) is due to the existence of a positive cosmological constant ($w = -1$) with

$$q_0 = \tfrac{1}{2}(\Omega_M)_0 - (\Omega_\Lambda)_0 < 0 \tag{38.15}$$

and, more specifically,

$$(\Omega_\Lambda)_0 \approx 0.7 \ . \tag{38.16}$$

Since then

$$(\Omega_M)_0 + (\Omega_\Lambda)_0 \approx 1 \quad \Leftrightarrow \quad (\Omega_k)_0 \approx 0 \ , \tag{38.17}$$

these last two (completely independent) observations are evidently compatible with each other.

<u>Remarks:</u>

1. Approximate numerical values for the parameters Λ and H_0 characterising the universe today are

$$H_0 \approx 2.3 \times 10^{-18}\ \mathsf{s}^{-1} \quad , \quad \Lambda \approx 10^{-35}\ \mathsf{s}^{-2} \ . \tag{38.18}$$

2. Somewhat more informative are perhaps the current estimates for the age of the universe,

$$t_0 \approx 13.8 \times 10^9 \text{years} \tag{38.19}$$

and the energy density of dark energy, approximately

$$\rho_\Lambda \approx (10^{-3}\ \mathsf{eV})^4 \ . \tag{38.20}$$

3. The equation of state parameter for dark energy is known to very close to $w = -1$,

$$w \approx -0.98 \pm 0.05 \ , \tag{38.21}$$

so a cosmological constant indeed appears to be a very plausible candidate for dark energy.

4. The fact that there is something like a cosmological constant is perhaps not particularly puzzling as such (in the absence of a good reason why it should not have been there in the first place), but nevertheless there are a number of puzzling issues related to the value of the cosmological constant - see section 38.4 for a brief discussion.

38.2 Exact Λ-CDM Solution

The exact solution for a spatially flat universe dominated by non-interacting (cold dark) matter was already given in (36.20),

$$a(t) = a_0(t/t_0)^{2/3} \quad . \tag{38.22}$$

and that for a spatially flat universe dominated by a cosmological constant was given in (36.23),

$$a(t) = a_0 e^{\sqrt{\Lambda/3}(t-t_0)} \quad . \tag{38.23}$$

An exact solution of the Friedmann equations can also be written down when both types of energy/matter are present simultaneously (as in the universe today). In this case the Friedmann equation is just

$$\Omega_M + \Omega_\Lambda = 1 \quad , \tag{38.24}$$

so that the Friedmann equation of the Λ-*CDM Model* (cold dark matter with a cosmological constant) can be written as

$$H^2 = H_0^2 \left(\frac{a_0^3}{a^3}(\Omega_M)_0 + (\Omega_\Lambda)_0 \right) \quad . \tag{38.25}$$

An exact solution of this equation is

$$\begin{aligned} a(t) &= a_0 \left(\frac{(\Omega_M)_0}{(\Omega_\Lambda)_0} \right)^{1/3} \left(\sinh((3/2)\sqrt{(\Omega_\Lambda)_0} H_0 t) \right)^{2/3} \\ &\equiv a_0 \left(\frac{(\Omega_M)_0}{(\Omega_\Lambda)_0} \right)^{1/3} (\sinh(t/t_\Omega))^{2/3} \quad . \end{aligned} \tag{38.26}$$

This evidently reproduces the above power-law (exponential) behaviour at early (late) times. The transition between the decelerating matter dominated and accelerating Λ-dominated phases occurs at the time $t = t_\Lambda$ at which

$$\ddot{a}(t_\Lambda) = 0 \quad . \tag{38.27}$$

Calculating

$$\frac{d^2}{dt^2} \sinh^{2/3}(t/t_\Omega) = (2/3)t_\Omega^{-2} \left(-(1/3)\sinh^{-4/3}(t/t_\Omega)\cosh^2(t/t_\Omega) + \sinh^{2/3}(t/t_\Omega) \right) \tag{38.28}$$

one finds that at $t = t_\Lambda$

$$\sinh^2(t_\Lambda/t_\Omega) = 1/2 \quad , \tag{38.29}$$

and plugging this back into (38.26), one sees that

$$\frac{a(t_\Lambda)}{a_0} = \left(\frac{(\Omega_M)_0}{2(\Omega_\Lambda)_0} \right)^{1/3} \quad . \tag{38.30}$$

With the values for the density parameters given above, $(\Omega_M)_0 = 0.3$, $(\Omega_\Lambda)_0 = 0.7$, this is roughly

$$a(t_\Lambda) \approx 0.6\, a_0 \quad . \tag{38.31}$$

Thus the transition occurred at a time when the universe was roughly speaking "half as big" as today, and at a corresponding value

$$z = \frac{a_0}{a(t_\Lambda)} - 1 \approx 0.66 \tag{38.32}$$

of the redshift parameter.

38.3 Flatness and Horizon Problems

The currently favoured Λ-CDM scenario with $(\Omega_M)_0 = 0.3$, $(\Omega_\Lambda)_0 = 0.7$ and $k = 0$ has been tested and confirmed in various independent ways. Nevertheless, it is a mystery and raises all kinds of questions and puzzles, in particular because the emergence of this particular universe appears to be somewhat unnatural (although it may perhaps be difficult to quantify this sentiment) and to require an incredible amount of fine-tuning.

I will not say anything about the dark matter component, since an explanation presumably needs to be found in the realm of particle physics (every model of physics beyond the standard model worth its salt has its own dark matter candidates), and since the story is in any case sufficiently strange and interesting even without worrying about, or having to put up with, things like WIMPS (weakly interacting massive particles), or MACHOS (massive compact halo objects), or *binos*, *winos* and other *neutralinos*.

Here, in a nutshell, are some of these problems or puzzles (and for more in-depth and far-reaching discussions of these issues please consult the extensive literature on cosmology). The first two (discussed in this section) are independent of the recent discovery of dark energy, and appear to require some fine-tuning or other mechanism to intervene in the very early universe. The third, on the other hand, has to do with dark energy / the cosmological constant today, and with the relatively recent period of accelerated expansion of the universe and is discussed in section 38.4.

I should stress that, while I generally make an attempt in these notes to present just the (well-established) facts, this is not entirely possible in this and the subsequent section, since there is no general consensus on how to resolve these issues (in particular those arising in connection with dark energy). As a consequence, these sections contain not just facts but certainly and unavoidably also some opinions (whereas there was no point in sharing with you my opinion about the Riemann curvature tensor, say - I don't even have one).

1. The Flatness Problem

The flatness problem is the problem that a universe with a value of the density parameter $\Omega = \Omega_M + \Omega_\Lambda \approx 1$ today, equivalently $(\Omega_k)_0 \approx 0$, appears to be "unnatural".

To see in simple terms what the issue is, note that from the Friedmann equation we can read off that

$$H^2 = \frac{8\pi G_N}{3}\rho - \frac{k}{a^2} \quad \Rightarrow \quad \Omega_k = -\frac{k}{a^2 H^2} \quad . \tag{38.33}$$

Now for a matter or radiation dominated universe we have $a(t) \sim t^{2/3}$ or $a(t) \sim t^{1/2}$ (with subleading corrections when $k \neq 0$), and therefore

$$a^2 H^2 = \dot{a}^2 \sim \begin{cases} t^{-2/3} & \text{for } w = 0 \\ t^{-1} & \text{for } w = 1/3 \end{cases} \tag{38.34}$$

so that

$$\Omega_k(t) \sim \begin{cases} t^{+2/3} & \text{for } w = 0 \\ t & \text{for } w = 1/3 \end{cases} \tag{38.35}$$

Thus a small value of $(\Omega_k)_0 \approx 0$ today (whether to within a few percent, as observations suggest, or even give or take a couple of orders of magnitude) seems to require $\Omega_k = 0$ with an enormous accuracy at some given time in the early universe.

It is more common to phrase this issue in terms of the critical density ρ_{cr} and the density parameter $\Omega = \Omega_M + \Omega_\Lambda$ for matter / radiation plus dark energy (but the contribution of the latter is negligible in the early universe). To that end we rewrite the Friedmann equation as

$$H^2 = \frac{8\pi G_N}{3}\rho - \frac{k}{a^2} \quad \Leftrightarrow \quad (\rho_{cr} - \rho)a^2 = -\frac{3k}{8\pi G_N} \quad , \tag{38.36}$$

the point of rewriting it in this way being that what now appears on the right-hand side is a constant. In terms of the density parameter Ω we have

$$(\Omega^{-1} - 1)\rho a^2 = -\frac{3k}{8\pi G_N} \quad . \tag{38.37}$$

Since the right-hand side is a constant, the product of the two terms on the left-hand side is a constant. However, as we have already seen (and used) in section 36.3, for ordinary matter ($w > -1/3$) the factor ρa^2 decreases (enormously) as the universe expands. This has to be (and will be, by the above equation) compensated and accompanied by a huge increase in the first factor. Thus, to have $(\Omega)_0 \approx 1$ today appears to require the density $\rho(t)$ of the universe to be equal to the critical density $\rho_{cr}(t)$ with an enormous accuracy at some given time in the early universe.

For example, in the matter dominated phase one has $\rho a^2 \sim a^{-1}$. In this case one can write (38.37) as

$$(\Omega^{-1} - 1)/a = (\Omega_0^{-1} - 1)/a_0 \quad \Leftrightarrow \quad \Omega = \frac{\Omega_0}{\Omega_0 + (a/a_0)(1 - \Omega_0)} \quad , \tag{38.38}$$

or

$$\Delta\Omega \equiv \frac{\Omega - 1}{\Omega} = \frac{a}{a_0}\Delta\Omega_0 \quad . \tag{38.39}$$

Going back to the beginning of the matter-dominated phase when radiation decoupled from matter, $z = a_0/a \approx 10^3$, one sees that

$$\Delta\Omega \approx 10^{-3}\Delta\Omega_0 \quad , \tag{38.40}$$

so that at that time the deviation from flatness (as measured by $\Delta\Omega$) was smaller by a factor of 10^{-3} than it is today.

However, this may begin to look like a more serious fine-tuning issue when one compares with the very early universe and specifies a particular ("earliest") time in the past. Concretely, let us (for the sake of argument) assume that

- the relevant early time is at the Planck scale, $t_P \approx 10^{-43}\mathsf{s}$
- and that matter consists of dust and radiation, with radiation dominating at early times, so that
$$\rho_r(t)a(t)^2 \sim a(t)^{-2} \sim t^{-1} \quad . \tag{38.41}$$

With the age of the universe t_0 taken to be (only orders of magnitude will be relevant for this argument) $t_0 \approx 10^{17}\mathsf{s}$, one finds that

$$(\Omega^{-1} - 1)_{t=t_0} \approx 10^{-2} \quad \Rightarrow \quad (\Omega^{-1} - 1)_{t=t_p} \approx 10^{-62} \quad . \tag{38.42}$$

In particular, any miniscule deviation from this tiny value at $t = t_p$, i.e. a miniscule difference between the actual and the critical density of the universe at that time, would have led to a universe completely incompatible with obervations. The universe would have most likely either recollapsed after a very short time ($\Omega \to \infty$), or expanded so quickly as to prevent the formation of any structure in the universe ($\Omega \to 0$). In fact, $\Omega = 0$ and $\Omega = \infty$ are the only attractors (attractive fixed points) of the theory in the terminology of dynamical systems, while $\Omega = 1$ is a repeller (unstable fixed point).

What is one supposed to make of this?

Now it should be borne in mind that the fact that Ω was very close to 1 in the past is not in itself a fine-tuning issue: after all, $\Omega \to 1$ in the past is *implied* by the Friedmann equations anyway, no matter what the size of Ω today. So one option is to declare that the above argument regarding finetuning is bogus, in particular since extrapolating all the way back to the Planck time (or the GUT era or ...), and maintaining the assumptions of isotropy and homogeneity that underlie the Friedmann equations requires some enormous leap of faith.[163]

[163]For various critical examinations of this flatness problem, see e.g. P. Helbig, *Is there a flatness problem in classical cosmology?*, `arXiv:1112.1666 [astro-ph.CO]` and M. Holman, *How Problematic is the Near-Euclidean Spatial Geometry of the Large-Scale Universe?*, `arXiv:1803.05148 [gr-qc]`, and references therein and thereto.

But if the flatness problem is perceived as a real problem, then no explanation for this can be found in the standard model of cosmology we have been discussing here so far (FLRW with standard matter content at early times). Occasionally, therefore, weakly anthropic arguments of the kind "if Ω had not been fine-tuned to this value by some (unknown) mechanism/deity, we couldn't be here in the first place to ask the question why Ω takes the value it does today" have been advocated.

While this cannot be ruled out (this is the whole problem with this entire anthropic reasoning business), there is actually a mechanism which naturally leads to the required tiny value of $\Omega - 1$ at early times, namely *inflation*.[164] As this mechanism is also invoked to solve other problems of the standard model of cosmology (such as the horizon problem, see below) which are (a) possibly more serious and (b) not obviously of anthropic significance, one could perhaps consider the anthropophile value of Ω_0 as an unintentional (but serendipitous) side-effect of inflation.

What inflation does is to postulate, for a number of reasons, a brief but highly significant period of exponential expansion in the very early universe, as could be triggered by the presence of a cosmological constant (36.23),

$$H^2 = (\Lambda/3) \quad \Rightarrow \quad a(t) \sim \mathrm{e}^{Ht} \ . \tag{38.43}$$

or more precisely by something that acts like a cosmological constant during a certain period, e.g. an almost constant scalar field (35.39). During such a period of exponential expansion, the factor ρa^2 grown enormously, and correspondingly $\Omega^{-1} - 1$ can be driven arbitrarily close to zero. When, after the end of inflation, the universe resumes its normal (radiation-dominated, say) decelerating expansion (this requires a process called *reheating*, since the universe has cooled down significantly during the inflationary period), the initial value of $\Omega^{-1} - 1$ for this phase can easily be small enough to account for $\Omega \approx 1$ to within a few percent today.

2. The Horizon Problem

The key issue here is the high degree of isotropy of the universe that we observe, in particular as manifested in the cosmic microwave background radiation (CMBR, see the references in section 34.8), which is isotropic to better than one part in 10^5. This is a puzzle because, as we will see below, according to the standard model of cosmology, the different regions of the spatial surface at $t = t_{ls}$ (the time of

[164] For introductions to inflation, see e.g. A. Linde, *Particle Physics and Inflationary Cosmology*, `arXiv:hep-th/0503203`; D. Baumann, *TASI Lectures on Inflation*, `arXiv:0907.5424 [hep-th]`; S. Tsujikawa, *Introductory review of cosmic inflation*, `arXiv:hep-ph/0304257`; A. Liddle, *An introduction to cosmological inflation*, `arXiv:astro-ph/9901124`; A. Linde, *Inflationary Cosmology*, `arXiv:0705.0164v2 [hep-th]`; R. Brandenberger, *Inflationary Cosmology: Progress and Problems*, `arXiv:hep-ph/9910410`.

"last scattering") from which the CMBR photons that we observe today originated were not ever in causal contact since the beginning of the universe. Thus standard causal microphysics provides no explanation for why these different regions were so precisely at the same temperature. This is schematically illustrated in Figure 53.

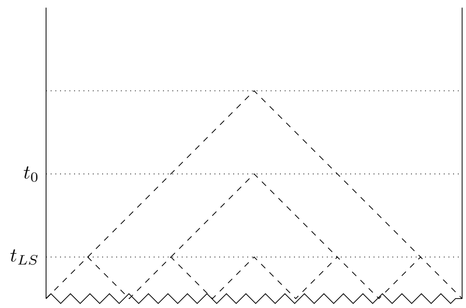


Figure 53: Horizon Problem ($k = +1$ cartoon): causally disconnected patches at the time t_{LS} of last scattering become simultaneously visible at later times.

To quantify this somewhat, note that the maximal spatial coordinate distance photons could have travelled between the big bang at $t = 0$ and $t = t_{ls}$ is (setting $k = 0$ for simplicity, the k-dependence is insignificant for this argument)

$$dt^2 = a(t)^2 dr^2 \quad \Rightarrow \quad r(t_{ls}) = r_{ph}(t_{ls}) = \int_0^{t_{ls}} \frac{dt}{a(t)} \quad , \tag{38.44}$$

the particle horizon at time $t = t_{ls}$. Converting this to proper distance at time $t = t_{ls}$, one obtains the quantity

$$R_{ph}(t_{ls}) = a(t_{ls}) \int_0^{t_{ls}} \frac{dt}{a(t)} \quad , \tag{38.45}$$

and for a radiation-dominated universe with $a(t) \sim t^{1/2}$ one finds

$$R_{ph}(t_{ls}) = t_{ls}^{1/2} \int_0^{t_{ls}} dt\, t^{-1/2} = 2t_{ls} \quad . \tag{38.46}$$

Its significance in the present context is that the past light cone of events that are further than $2R_{ph}(t_{ls})$ apart on the surface $t = t_{ls}$ do not intersect, so that they are so far apart that they were never before in causal contact. In particular this means that no causal interaction can be responsible for the temperature being the same at the two events.

In the (matter dominated) meantime (i.e. between t_{ls} and t_0, today) the size of such a causal patch of size $\sim R_{ph}(t_{ls})$ on the last scattering surface has expanded to proper size

$$\frac{a(t_0)}{a(t_{ls})} R_{ph}(t_{ls}) \sim (t_0/t_{ls})^{2/3} t_{ls} = t_0^{2/3} t_{ls}^{1/3} \quad . \tag{38.47}$$

On the other hand, the distance over which the CMBR photons could have travelled since $t = t_{ls}$ is

$$a(t_0)\int_{t_{ls}}^{t_0}\frac{dt}{a(t)} = 3t_0^{2/3}(t_0^{1/3} - t_{ls}^{1/3}) \sim t_0 \quad , \tag{38.48}$$

where we have dropped the second term since $t_0 \sim 10^{10}$ years while $t_{ls} \sim 10^5$ years (once again all numbers here and below are just meant to be order of magnitude estimates). Thus the region of the last scattering surface from which we receive the CMBR photons today is much larger than a causal patch, their ratio being (one can of course calculate this either at $t = t_{ls}$ or at $t = t_0$, here we have chosen the latter option)

$$t_0/(t_0^{2/3}t_{ls}^{1/3}) = (t_0/t_{ls})^{1/3} \approx 10^{5/3} \quad . \tag{38.49}$$

What this means is that the sky splits into roughly $4\pi 10^{10/3} \gtrsim 10^4$ disconnected patches, that were never in communication before sending light to us. In view of this the observed isotropy of the CMBR is not only astounding but utterly implausible.

Again, inflation solves this in an extremely natural way. Inflation operates at a time $t_i \ll t_{ls}$ (perhaps some time between $10^{-36}\mathsf{s}$ and $10^{-32}\mathsf{s}$ after the big bang) and it can easily inflate a tiny causally connected patch at that time $t = t_i$ to such a size that at time $t = t_{ls}$ it is (more than) large enough to explain the isotropy of the CMBR.

This is easy to understand by looking at the expanding de Sitter metric in conformal time (37.63),

$$ds^2 = \mathrm{e}^{2t/\ell}(-dt^2\mathrm{e}^{-2t/\ell} + d\vec{x}^2) = \frac{\ell^2}{\eta^2}(-d\eta^2 + d\vec{x}^2) \tag{38.50}$$

with

$$\eta = -\ell\mathrm{e}^{-t/\ell} \quad . \tag{38.51}$$

This shows that in a conformal diagram whatever came before inflation can now be pushed to arbitrarily large negative values of η, and therefore the horizon problem as depicted in Figure 53 does not arise.

The majority view among cosmologists appears to be that for these and other reasons inflation should be considered to be part of the standard model of cosmology. However, science is not decided by opinion polls and it is good scientific practice to keep an open mind. In particular, one should keep in mind the possibility that for instance the two (flatness and horizon) problems (or perhaps better: puzzles) may indicate problems with the approximation of the early universe by a FLRW cosmology and are thus not something that should (or needs to, or perhaps even can) be solved by (inflationary) modifications within the class of FLRW models.[165]

The most significant success of inflation, however, is that it not only "explains" the high degree of isotropy of the CMBR but that it also provides a mechanism for, and a precise quantitative account of, density perturbations and the small inhomogeneities exhibited by the CMBR (namely as arising from quantum vacuum fluctuations). Simple models of inflation appear to be able to account for the latest (2013) precision measurements and data from the Planck satellite[166], but the debate over this and other aspects of inflation continues.[167] I will not try to constantly update these notes as this discussion evolves.[168]

Ultimately, in order for inflation to be considered a natural solution to the above problems (and others I have not mentioned - see the literature cited above), one needs to be able to show that inflation arises fairly naturally (in some precise sense) and does not itself require a comparably huge amount of fine-tuning in order to resolve these issues. This is a complicated and intensely debated issue, and one I don't feel sufficiently competent and knowledgeable about to have an informed opinion, let alone utter an opinion in public. It will perhaps only be settled once one has a better understanding of the physics in the very early, pre-inflationary, era that is supposed to be responsible for setting the initial conditions for inflation.

38.4 Aspects of the Cosmological Constant Problem(s)

This is a complex, confusing and multi-faceted problem that has been around for a long time, and I will not remotely be able to do justice to it here.[169] It has been sharpened and brought to the forefront again by the discovery of dark energy.[170]

We have remarked before that the cosmological constant looks like a vacuum energy contribution to the energy-momentum tensor. It is perhaps better to turn this around and

[165] For example, there are claims that in certain inhomogeneous Lemaître-Tolman cosmological models the horizon problem can be avoided without invoking inflation - see e.g. chapter 18.17 of J. Plebanski, A. Krasinski, *An Introduction to General Relativity and Cosmology.*

[166] A. Linde, *Inflationary Cosmology after Planck 2013*, `arXiv:1402.0526 [hep-th]`.

[167] A. Ijjas, P. Steinhardt, A. Loeb, *Inflationary schism after Planck2013*, `arXiv:1402.6980 [astro-ph.CO]`.

[168] For a recent (at the time of writing this) assessment of inflation see J. Martin, *Cosmic Inflation: Trick or Treat?*, `arXiv:1902.05286 [astro-ph.CO]`.

[169] The classic reference for this is the authoritative and influential article by S. Weinberg, *The cosmological constant problem*, Rev. Mod. Phys. 61 (1989) 1-23.

[170] For recent reviews see e.g. S. Carroll, *The Cosmological Constant*, Living Rev. Relativity 4 (2001) 1, `http://www.livingreviews.org/lrr-2001-1`; J. Polchinski, *The Cosmological Constant and the String Landscape*, `arXiv:hep-th/0603249`; T. Padmanabhan, *Cosmological Constant - the Weight of the Vacuum*, `arXiv:hep-th/0212290`; R. Bousso, *TASI Lectures on the Cosmological Constant*, `arXiv:0708.4231 [hep-th]`, J. Martin, *Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)*, `arXiv:1205.3365 [astro-ph.CO]`.

to say that vacuum energy is one potential contribution to the cosmological constant, in the sense that the associated energy density ρ_Λ can be written as

$$\rho_\Lambda = (\rho_\Lambda)_{\text{bare}} + (\rho_\Lambda)_{\text{vac}} = \rho_{\Lambda_0} + \rho_{\text{vac}} \quad . \tag{38.52}$$

Here

$$\rho_{\Lambda_0} = \frac{\Lambda_0}{8\pi G_N} \tag{38.53}$$

is associated with a "bare" cosmological constant Λ_0, some parameter in the action, while ρ_{vac} is a quantum contribution arising from the energy in the ground state or vacuum, i.e. it is the vacuum expectation value of the Hamiltonian density operator $\hat{T}_{00}$,

$$\rho_{\text{vac}} =< \text{vac}|\hat{T}_{00}|\text{vac} > \quad . \tag{38.54}$$

In a Poincaré-invariant field theory in Minkowski space, with a Poincaré-invariant ground state, this will give rise to a contribution

$$< \text{vac}|\hat{T}_{ab}|\text{vac} >= -\rho_{\text{vac}}\eta_{ab} \quad , \tag{38.55}$$

so the equation of state

$$p_{\text{vac}} = -\rho_{\text{vac}} \tag{38.56}$$

is implied by Lorentz invariance.

This suggests that this extends to quantum field theory in a gravitational background (as described e.g. in the references in footnote 87 in section 27.7) in the form

$$< \text{vac}|\hat{T}_{\alpha\beta}|\text{vac} >= -\rho_{\text{vac}} g_{\alpha\beta} \quad . \tag{38.57}$$

In discussions of the cosmological constant problem, it is usually assumed that (38.57) holds, but it is worth bearing in mind that this is a non-trivial (and possibly incorrect) assumption:

1. First of all, the semi-classical approach of treating quantised matter fields in a classical gravitational background is not something that is completely internally consistent, but it is generally believed to be a valid approximation for sufficiently weak gravitational fields (for which one expects whatever quantum gravitational effects may be present in principle to be completely negligible).

2. Even in Minkowski space an expression like (38.55) requires some sort of regularisation procedure, and the procedures that are common or privileged in the case of Poincaré-invariant field theories may either not be available or may not be in any way privileged when one considers a general curved background. Thus there are ambiguities in the calculation of $< \hat{T}_{\alpha\beta} >$ which may effect the validity of (38.57).

One (relatively harmless) ambiguity of this kind would be the addition of a term proportional to the Einstein tensor $G_{\alpha\beta}$ to the right-hand side of (38.57),

$$< \mathrm{vac}|\hat{T}_{\alpha\beta}|\mathrm{vac} >= -\rho_{\mathrm{vac}} g_{\alpha\beta} + \frac{\gamma}{8\pi G_N} G_{\alpha\beta} \tag{38.58}$$

for some constant γ. This would be compatible with $\nabla^\alpha T_{\alpha\beta} = 0$ and with the Minkowskian limit, and would amount to a renormalisation of Newton's constant G_N.

3. Finally, and perhaps most crucially, while reasonable Poincaré-invariant quantum field theories in Minkowski space typically have a preferred and unique Poincaré-invariant ground state, the vacuum, the key feature of quantum field theory in a curved background is that quantum field theory can be extended to this case but that this uniqueness is lost (and it is precisely this non-uniqueness that is at the heart of the characteristic phenomena of quantum field theory in a gravitational field like particle production and Hawking radiation). As a consequence, the state unceremoniously called $|\mathrm{vac}>$ in (38.57) is not unique, and there is no *a priori* reason to believe that (38.57) will hold for any putative vacuum state.

In spite of all this it is usually assumed that something like (38.57) is at least approximately true for sufficiently weak fields, and we will proceed with this assumption.

With the vacuum-energy contributing to the cosmological constant, it is natural to imagine that contributions to the cosmological constant can arise from many quantum field theory processes such as phase-transitions etc, and are of the order of the momentum cut-off, say. Now in field theory in a fixed background, only energy-differences matter, not absolute energies, but the inclusion of gravity and its presumed universal coupling to all forms of energy change this. It is then not clear if one can simply drop finite (let alone infinite) energy-differences from one's equations.

Depending on the physics or physical process one is looking at, natural estimates for the energy scale of a cosmological constant produced in this way via some phase transition, say, might be in the MeV or GeV range, or perhaps (via some wild extrapolation) even at the (ultimate Planck cut-off) scale of

$$E_P = \sqrt{\frac{\hbar c^5}{8\pi G_N}} \approx 10^{18}\ \mathsf{GeV} \tag{38.59}$$

(with the inclusion of the factor of 8π in the definition this is known as the reduced Planck energy). But one does not really have to go that far to see that there is an issue: after all, any of these values are many many orders of magnitude above what was and is compatible with observation, or even with the very existence of our universe.

One cannot simply "solve" this problem by using the bare value Λ_0 of the cosmological constant to cancel the vacuum contribution by hand, because

- this would require an enormous (and unexplained) fine-tuning of this bare value
- contributions to the vacuum energy are expected to arise at various instances during the evolution of the universe while this cancellation could at best be achieved at one point in time during the evolution of the universe (if this cancellation is chosen to take place too early, it will be incompatible with observations today while if it happens too late it will be incompatible with the well-established thermal history of the universe).

It therefore seemed natural to seek some mechanism that would simply make the net cosmological constant identically zero, and a lot of effort went into finding or inventing some mechanism responsible for this. Nowadays, with the strong evidence for dark energy, this cannot be the solution and the original question/problem has morphed into at least 3 distinct questions, namely

1. Why is the cosmological constant Λ not huge?
2. What is the origin of Dark Energy, with $\rho_{DE} \sim (10^{-3}\ \mathsf{eV})^4$?
3. Why is ρ_Λ (constant in time) of the same order of magnitude as ρ_M (today)?

and we will briefly address these in turn.

1. Why is Λ not huge?

   This is the same problem (described above) as before the discovery of dark energy, and is essentially the question "why the vacuum does not gravitate" (Polchinski; see his article in footnote 170 for a very illuminating discussion of this issue).

   This whole business points to some loop-hole in the reasoning leading to (38.57) and the estimates for the size of this quantity.[171] It may even indicate some fundamental lack of (or perhaps mis-) understanding about how gravity couples to quantum fields.[172]

   One on the face of it rather radical, but actually quite conservative, possibility that has been explored (and that I mention here not because I think it is the most plausible explanation but because it is easy to explain and also quite interesting in its own right) is to modify the Einstein equations in such a way that the

[171] See e.g. Q. Wang, Z. Zhu, W. Unruh, *How the huge energy of quantum vacuum gravitates to drive the slow accelerating expansion of the Universe*, `arXiv:1703.00543 [gr-qc]` where it is argued that, because of quantum fluctuations, the quantum vacuum actually gravitates substantially differently from a cosmological constant.

[172] See S. Hollands, R. Wald, *Quantum Field Theory Is Not Merely Quantum Mechanics Applied to Low Energy Effective Degrees of Freedom*, `arXiv:gr-qc/0405082` for some reflections on this issue.

gravitational field does not couple directly to any contribution to the energy-momentum tensor $\sim g_{\alpha\beta}$, i.e. such that they are invariant under shifts

$$T_{\alpha\beta} \to T_{\alpha\beta} + c g_{\alpha\beta} \tag{38.60}$$

for constant c. To that end one postulates the so-called *trace-free Einstein equations*

$$R_{\alpha\beta} - \tfrac{1}{4} g_{\alpha\beta} R = 8\pi G_N \left(T_{\alpha\beta} - \tfrac{1}{4} g_{\alpha\beta} T\right) \ , \tag{38.61}$$

supplemented by the usual covariant energy-momentum tensor conservation law

$$\nabla_\alpha T^{\alpha\beta} = 0 \tag{38.62}$$

(which is not implied by the modified Einstein equations (38.61)).[173] Note that both sides of (38.61) are manifestly traceless. In particular, therefore, any contribution to the energy-momentum tensor $\sim g_{\alpha\beta}$ (like a cosmological constant) does not contribute to (38.61),

$$T_{\alpha\beta} \sim g_{\alpha\beta} \quad \Rightarrow \quad T_{\alpha\beta} - \tfrac{1}{4} g_{\alpha\beta} T = 0 \ , \tag{38.63}$$

and therefore does not couple directly to gravity (said differently, the source term is invariant under the shift (38.60), as required).

While this looks like a major departure from the usual (and well-tested) Einstein equations, and this might make one believe that the equations (38.61) can easily be ruled out experimentally, what actually happens is more subtle and somewhat surprising. Namely, writing (38.61) as

$$R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R = 8\pi G_N \left(T_{\alpha\beta} - \tfrac{1}{4} g_{\alpha\beta} T\right) - \tfrac{1}{4} g_{\alpha\beta} R \tag{38.64}$$

and using the contracted Bianchi idenities $\nabla_\alpha G^{\alpha\beta} = 0$ as well as the postulated $\nabla_\alpha T^{\alpha\beta} = 0$, one deduces that

$$\nabla_\alpha (8\pi G_N\, T + R) = 0 \quad \Rightarrow \quad 8\pi G_N\, T = -R + 4\Lambda \tag{38.65}$$

for some integration constant Λ. Plugging this result back into (38.64), one finds

$$\begin{aligned} R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R &= 8\pi G_N\, T_{\alpha\beta} - \tfrac{1}{4} g_{\alpha\beta} [8\pi G_N\, T + R] \\ &= 8\pi G_N\, T_{\alpha\beta} - g_{\alpha\beta} \Lambda \ . \end{aligned} \tag{38.66}$$

This is nothing other than the usual Einstein equations with a cosmological constant,

$$R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R + \Lambda g_{\alpha\beta} = 8\pi G_N\, T_{\alpha\beta} \ , \tag{38.67}$$

[173] See e.g. G. Ellis, H. van Elst, J. Murugan, J.-P. Uzan, *On the Trace-Free Einstein Equations as a Viable Alternative to General Relativity*, `arXiv:1008.1196v2 [gr-qc]` and references therein for a more detailed discussion and the history of this and other variants of *unimodular gravity*, originally even considered by Einstein himself in 1919.

the crucial difference being that here Λ is *not* determined by the matter content and its vacuum energy but arises solely as an integration constant. While this does not explain the observed tiny value of the cosmological constant, it separates this issue from that of the vacuum fluctuations. Note also that this scenario does not rule out a gravitational coupling to general quantum corrections to states of some physical system (Lamb shift, Casimir energies, ...) which are not Lorentz invariant, as required e.g. by precision tests of the equivalence principle (see the discussion regarding "do vacuum fluctuations gravitate?" in the reviews by Polchinski and Martin in footnote 170).

2. What is the origin of Dark Energy, with $\rho_{DE} \sim (10^{-3}\ \mathsf{eV})^4$?

   Perhaps this is the wrong question altogether. After all, it could just be that Λ happens to be a new constant of nature, like G_N, that has now been found to be non-zero and that has been experimentally determined, and whose precise value cannot currently (or perhaps even in principle) be derived from some underlying theory.

   But a priori it is also possible that the value can be derived, and/or that dark energy is due to something other than a cosmological constant. Again anthropic reasoning (shudder!) can be invoked to argue that the above is a plausible value for the cosmological constant. It would however of course be very desirable to find alternative explanations.

   Numerous models have been proposed that do give rise to a late-time acceleration of the universe for one reason or another (one of the buzzwords here is *quintessence*). Most of them, however, assume (explicitly or implicitly) that somehow the first problem has been solved and that there is no (bare or combined) cosmological constant source term in the Einstein equations.

   It is not clear at all, however, what kind of mechanism could produce on the nose an effect of the desired size. One curious observation, which may provide us with a clue, is that the energy scale E_Λ associated with the cosmological constant,

$$E_\Lambda \approx 10^{-3}\ \mathsf{eV} = 10^{-12}\ \mathsf{GeV}\ \ , \tag{38.68}$$

   which is much much smaller than the "natural" Planck energy scale,

$$E_\Lambda \approx 10^{-30} E_P\ \ , \tag{38.69}$$

   is however of the order of the geometric mean of the (current-day) Hubble scale and the Planck scale (I do not know who (if anybody) should be credited with this observation which has probably been made independently multiple times). In energy units this is the curious relation

$$H_0 \approx 10^{-42}\ \mathsf{GeV}\ \ ,\ \ E_P \approx 10^{18}\mathsf{GeV} \quad \Rightarrow \quad E_\Lambda \approx (E_P H_0)^{1/2}\ \ , \tag{38.70}$$

and in time-units, expressed in terms of the (reduced) *Planck time*

$$t_P = \sqrt{\frac{\hbar 8\pi G_N}{c^5}} \approx 10^{-43}\,\mathsf{s} \;\;, \tag{38.71}$$

this is the equally baffling statement that

$$(H_0)^{-1} \approx 10^{60} t_P \quad , \quad \Lambda \approx 10^{-120}(t_P)^{-2} \approx (H_0)^2 \;\; . \tag{38.72}$$

This is certainly intriguing, as it appears to relate the cosmological constant to both a fundamental UV cut-off (the Planck scale) and a current IR cut-off (the Hubble scale), which has led to some "holographic" ruminations.[174] This curious relation is also one aspect of the so-called "coincidence problem", and we will briefly come back to this below.

A more conservative realisation of the scenario described by (38.70) is also possible.[175] For concreteness, consider the scalar field in a cosmological background discussed in section 34.10, with a UV momentum cutoff at a momentum $k = k_{UV}$. Namely if, for one reason or another, one can argue that the leading quartic divergence in the vacuum energy

$$\rho_{\text{vac}} \sim \tfrac{1}{2}\int \frac{d^3k}{(2\pi)^3}\sqrt{k^2 + m_{eff}^2} = \frac{(k_{UV})^4}{16\pi^2} + \ldots \tag{38.73}$$

(which would also be present in, and hence destroy, Minkowski space) should be subtracted, then the subleading term is quite generically of the form

$$\rho_{\text{vac}} \sim (k_{UV})^2 (H_0)^2 \tag{38.74}$$

(in the model in section 34.10 this H_0^2-term would arise from the scalar curvature term in the effective mass). With a cut-off at the Planck scale, this gives a vacuum energy contribution of the order of the observed dark energy density,

$$k_{UV} \sim k_P \quad \Rightarrow \quad \rho_{\text{vac}} \sim (E_P)^2 (H_0)^2 \sim (\rho_\Lambda)_{\text{observed}} \;\; . \tag{38.75}$$

In designing such scenarios, care should be taken that the vacuum contribution thus determined really has an equation of state parameter equal (or very close to) $w = -1$, so that ρ_{vac} calculated today is really (approximately) constant (and does not behave like $\rho_{\text{vac}}(t) \sim H(t)^2$, say, which would be in conflict with many cosmological observations). Exploring scenarios of this kind is clearly a constructive alternative to anthropic incantations.

[174] See e.g. A. Cohen, D. Kaplan, A. Nelson, *Effective Field Theory, Black Holes, and the Cosmological Constant*, `arXiv:hep-th/9803132`, M. Li, *A Model of Holographic Dark Energy*, `arXiv:hep-th/0403127`, T. Padmanabhan, *Vacuum Fluctuations of Energy Density can lead to the observed Cosmological Constant*, `arXiv:hep-th/0406060`, S. Hsu, A. Zee, *A speculative relation between the cosmological constant and the Planck mass*, `arXiv:hep-th/0406142`, and references thereto for theoretical speculations along these lines.

[175] See e.g. M. Maggiore, *Zero-point quantum fluctuations and dark energy*, `arXiv:1004.1782 [astro-ph.CO]`, and D. Bernard, A. LeClair, *Scrutinizing the Cosmological Constant Problem and a possible resolution*, `arXiv:1211.4848 [hep-th]`.

3. Why is ρ_Λ (constant in time) of the same order of magnitude as ρ_M (today)?

This is weird, and one version of what is known as the *Coincidence Problem* in cosmology. To see why this is peculiar, note that ρ_Λ is constant while $\rho_M \sim a^{-3}$ so that the ratio of the two behaves as

$$\frac{\rho_\Lambda(t)}{\rho_M(t)} \sim a(t)^3 \quad . \tag{38.76}$$

It is therefore, one might reason, extremely unlikely to find oneself at a time when the two happen to be comparable in size, but nevertheless this is precisely the time we find ourselves at (how precisely depends among other things on what one means by "comparable in size"). However, any such statement has to also rely on some probability distribution for something (the probability that observers exist at some time t_0?) that allows one to quantify this statement somewhat. But it is not clear if this is an issue that really requires an explanation.

If one thinks that this coincidence is indeed a problem that requires a solution, then one can try to come up with models in which dark energy does not behave like a cosmological constant at all times. However:

> "Why obfuscate? If a poet sees something that walks like a duck and swims like a duck and quacks like a duck, we will forgive him for entertaining more fanciful possibilities, It could be a unicorn in a duck suit - who's to say! But we know that more likely, it's a duck."[176]

Fair enough.[177]

Note by the way that this coincidence issue is equivalent to the empirical observation (38.70) mentioned above. Namely, with $\rho_\Lambda \sim \rho_M$, and both of order of the critival density because $\Omega_M + \Omega_\Lambda = 1$, one has

$$\rho_\Lambda = \frac{\Lambda}{8\pi G_N} \sim (\rho_c)_0 = \frac{3H_0^2}{8\pi G_N} \sim H_0^2 E_P^2 \tag{38.77}$$

which is just (38.70). One can also eliminate any reference to Planck units and the Planck scale and write this as (38.72)

$$\Lambda \approx (H_0)^2 \quad . \tag{38.78}$$

Thus another (and perhaps more objective and constructive) way of stating the coincidence issue is as the curious fact that the two a priori completely unrelated time-scales, one set by $(H_0)^{-1}$ (which turns out to be remarkably close to the current best estimates for the age of the universe) and the other set by the energy density or curvature radius of dark energy, are approximately equal. Why should

[176] R. Bousso, *TASI Lectures on the Cosmological Constant*, `arXiv:0708.4231 [hep-th]`.

[177] See also E. Bianchi, C. Rovelli, *Why all these prejudices against a constant?*, `arXiv:1002.3966 [astro-ph.CO]` for further reflections on this and related issues.

> the cosmological constant today be related to the age of the universe today??? Perhaps this is indeed not a coincidence and is trying to tell us something.

It is not clear if a fundamental physics (non-multiverse, non-anthropic) explanation for dark energy can be found that is capable of resolving all 3 of the cosmological constant problems at once, but perhaps one should at least be humble and continue to entertain the possibility, alluded to above, that this issue is telling us that we are missing something fundamental about the coupling of gravity and quantum fields.

This is not to say that there is not a strong case to be made for what appears to be an extreme fine-tuning of other (standard model) parameters to support everything from primordial nucleosynthesis to chemistry and life as we know it.[178] However, among other things precisely because of the proviso "as we know it", and also because of the (as far as I know) limited understanding of chemistry based on other than standard model gauge theories, this begs the question and the significance of these findings is far from clear and difficult to assess.

[178] See e.g. L. Barnes, *The Fine-Tuning of the Universe for Intelligent Life*, `arXiv:1112.4647 [physics.hist-ph]` for a recent assessment of the situation.

G: Varia

Until now, our treatment of the basic structures and properties of General Relativity has been reasonably systematic and standard. This part contains a biased and varied selection of other fun topics.

Here is a list of the topics (currently) covered in this part:

38. de Sitter and anti-de Sitter Spaces
39. Vaidya Metrics I: Bondi Gauge and Radiation Fields
40. Vaidya Metrics II: Radial Null and Timelike Geodesics
41. Vaidya Metrics III: Linear Mass $m(v) = \mu v$ (a case study)
42. Exact Wave-like Solutions of the Einstein Equations
43. Kaluza-Klein Theory

39 de Sitter and anti-de Sitter Space

(Anti-) de Sitter spaces are the simplest solutions of the Einstein equations with $\Lambda \neq 0$. As such they are the curved counterparts of the $\Lambda = 0$ Minkowski space-time. In particular, they are the unique constant curvature (or maximally symmetric) space-times, just as Minkowski space is the unique flat (or Poincaré-symmetric) space-time. Thus they are the Lorentzian-signature counterparts of spheres and hyperboloids, and made a first appearance as such in section 14. Since they are thus in some sense the simplest non-trivial space-times, it is worthwhile to study them in some detail.

The de Sitter and anti-de Sitter spaces subsequently reappeared in the context of cosmology in section 37.5, but it may not be immediately apparent that what we called (anti-) de Sitter there is indeed identical to what we called (anti-) de Sitter in section 14. In order to bridge this gap, we will in particular redo the analysis of section 14.4 in the special case of interest. We will realise the (A)dS spaces via embeddings into a higher-dimensional vector space, and we will use this embedding to express the resulting induced metric in various coordinate systems. In this way we will, in particular, also recover the metrics encountered in the cosmological context above. This will complete the proof that the solution of the Friedmann equations in the cosmological constant dominated phase is unique (for a given cosmological constant Λ) and uniquely given by the maximally symmetric (A)dS space.

39.1 Embeddings, Isometries and Coset Space Structure

We will now discuss the embeddings of (A)dS space, beginning with the more familiar cases of Euclidean signature spheres and hyperboloids, already discussed at some length in section 14.

We will denote the coordinates of the 5-dimensional embedding space by z^A, the range of the indices (e.g. 1 to 5 or 0 to 4) being chosen to be whatever is convenient or suggestive in the case at hand.

1. We can realise the (unit-radius) sphere S^4 by embedding it into $\mathbb{R}^5$ via

$$S^4 : \quad (z^1)^2 + \ldots + (z^5)^2 = +1 \ . \tag{39.1}$$

   If we equip $\mathbb{R}^5$ with the standard Euclidean metric, which is invariant under translations ($\sim \mathbb{R}^5$) and rotations $SO(5)$ of $\mathbb{R}^5$, then the defining equation above is invariant under the $SO(5)$-rotations, and the induced metric on S^4 will therefore be $SO(5)$-invariant. This gives rise to the standard $SO(5)$-invariant line element $d\Omega_4^2$. The dimension of the isometry group is 5(5-1)/2=10. This is the same as the dimension of the Poincaré (or Euclidean) group in 4 dimensions, and thus the S^4 equipped with this metric is maximally symmetric.

If we wanted to discuss the sphere of radius R, we would replace the $+1$ on the right-hand side of the above equation by R^2 (and likewise for the radius R or curvature radius ℓ below).

2. Likewise, we can realise the (unit curvature radius) hyperboloid H^4 by embedding it into $\mathbb{R}^5$ via
$$H^4 : \quad -(z^0)^2 + (z^1)^2 + \ldots + (z^4)^2 = -1 \quad . \tag{39.2}$$
We will now of course correspondingly choose the embedding space to be $\mathbb{R}^{1,4}$, i.e. the space equipped with the Lorentz-singnature metric
$$ds^2 = -(dz^0)^2 + (dz^1)^2 + \ldots + (dz^4)^2 \quad . \tag{39.3}$$
Its isometry group is the $(4+1)$-dimensional Poincaré group, which has dimension 15. The equation defining H^4 is left invariant by its $SO(4,1)$ Lorentz-subgroup which has dimension 10, and thus the metric induced on H^4 by the Minkowski metric on the embedding space will have isometry-group $SO(4,1)$ and is maximally symmetric. This metric has Euclidean signature because the (-1) on the right-hand side of (39.2) allows one to completely eliminate the time-like direction z^0, and the corresponding line element is denoted by $d\tilde{\Omega}_4^2$.

3. If we change the sign on the right-hand side of (39.2), the equation will still be invariant under $SO(4,1)$, but now the signatue of the induced metric will be Lorentzian instead of Euclidean and we obtain a realisation of a maximally symmetric space-time, namely de Sitter space,
$$dS_4 : \quad -(z^0)^2 + (z^1)^2 + \ldots + (z^4)^2 = +1 \quad . \tag{39.4}$$

4. By the same token, we can obtain a maximally symmetric Lorentzian signature space-time from the embedding equation
$$AdS_4 : \quad -(z^0)^2 + (z^1)^2 + (z^2)^2 + (z^3)^2 - (z^4)^2 = -1 \quad . \tag{39.5}$$
Since this equation is $SO(3,2)$-invariant, this space-time will have isometry group $SO(3,2)$, induced from the signature (2,3) metric on the embedding-space $\mathbb{R}^{2,3}$. The dimension of $SO(3,2)$ is also 10, just like that of $SO(4,1)$ or $SO(5)$, and (39.5) defines the maximally symmetric Anti-de Sitter space (actually, for AdS we will take the universal covering space of the space described by (39.5) - we will come back to this below).

As already indicated in section 14, the statements about the isometries of maximally symmetric space(-time)s can be compactly summarised by writing them as homogeneous spaces of the isometry groups.

This generalises the statement that the 2-sphere can be written as the homogeneous space (or coset space) $SO(3)/SO(2)$, which itself comes about as follows:

1. It is clear that the $SO(3)$ rotations act transitively on the 2-sphere, i.e. that any point can be mapped to any other point (this is the property of homogeneity discussed in section 14).

2. Moreover, given any point p on the 2-sphere, there is an $SO(2)$ subgroup of $SO(3)$, $SO(2)_p \subset SO(3)$, consisting of rotations around the axis passing through that point, that leaves the point p invariant but acts on the vectors at that point by 2-dimensional rotations (isotropy).

3. Since this $SO(2)_p$-transformation must also be a symmetry of the metric at that point, this shows in particular that the metric has a Euclidean signature at each point.

4. Putting all this together, given any point p, we can establish a 1:1 correspondence between points on S^2 and elements of $SO(3)$ modulo elements of $SO(2)_p$, and we write this as
$$S^2 \cong SO(3)/SO(2) \quad , \tag{39.6}$$
the set on the right-hand side considered as the set of equivalence classes $[g]$ with $g \in SO(3)$ and $[gh] = [g]$ for $h \in SO(2)$, say (this defines right-cosets, $SO(2)$ acting on the right on an $SO(3)$-element).

These statements generalise straightforwardly to the 4-sphere, and also to the other maximally symmetric space-times discussed above, and we summarise these facts in the following table, adding also the notation we will occasionally use for the corresponding line-element (when we do not just write it anonymously as ds^2), and giving the embedding (with $\vec{z}^2 = (z^1)^2 + (z^2)^2 + (z^3)^2$):

$M \cong G/H$	G	H	$d\Omega^2$	embedding
$M = S^4$	$SO(5)$	$SO(4)$	$d\Omega_4^2$	$+\vec{z}^2 + (z^4)^2 + (z^5)^2 = +1$
$M = H^4$	$SO(4,1)$	$SO(4)$	$d\tilde{\Omega}_4^2$	$-(z^0)^2 + \vec{z}^2 + (z^4)^2 = -1$
$M = dS_4$	$SO(4,1)$	$SO(3,1)$	$d\Omega_{1,3}^2$	$-(z^0)^2 + \vec{z}^2 + (z^4)^2 = +1$
$M = AdS_4$	$SO(3,2)$	$SO(3,1)$	$d\tilde{\Omega}_{1,3}^2$	$-(z^0)^2 + \vec{z}^2 - (z^4)^2 = -1$

(39.7)

The G-isometries are generated by the 5(5-1)/2 = 10 rotational Killing vectors (cf. (10.27))
$$J_{AB} = \eta_{AC} z^C \partial_B - \eta_{BC} z^C \partial_A = -J_{BA} \tag{39.8}$$
of the metric η_{AB} of the embedding space, and they satisfy the Lie bracket algebra
$$[J_{AB}, J_{CD}] = \eta_{AD} J_{BC} + \eta_{BC} J_{AD} - \eta_{AC} J_{BD} - \eta_{BD} J_{AC} \tag{39.9}$$
which provides a realisation of the Lie algebra of G.

This generalises in an obvious way to higher dimensions, so that e.g. $SO(4,2)$ is the isometry group of AdS_5. As we saw before (section 10.3), this group also happens to be the *conformal group* of 4-dimensional Minkowski space, and this is one of the fundamental ingredients in the so-called *AdS/CFT correspondence* relating gravitational theories in 5-dimensional (asymptotically) anti-de Sitter space-times to conformal field theories in $3+1$ dimensions.[179]

39.2 Some Coordinate Systems for de Sitter space

For de Sitter space, we have the defining equation (39.4),

$$-(z^0)^2 + (z^1)^2 + \ldots + (z^4)^2 = +1 \quad . \tag{39.10}$$

This describes a time-like hyperboloid with topology $\mathbb{R} \times S^3$, the S^3 arising from the slicing at fixed z^0,

$$(z^1)^2 + \ldots + (z^4)^2 = +1 + (z^0)^2 = \text{const.} > 0 \quad . \tag{39.11}$$

We can introduce suitable coordinates by solving (39.10), i.e. by parametrising the hyperboloid, either globally or at least locally. As we will see in the following, different coordinatisations are naturally suggested by grouping the terms in (39.10) in different ways.

Here is an overview of the coordinate systems and topics discussed in this section:

1. Global Coordinates
2. Conformal Time, Penrose Diagrams and Horizons
3. Hyperbolic Slicing Coordinates
4. de Sitter Slicing Coordinates
5. Planar Coordinates
6. Static Coordinates
7. Eddington-Finkelstein and Kruskal-Szekeres Coordinates
8. Interlude on (A)dS Schwarzschild
9. Painlevé-Gullstrand-like Coordinates

[179]See e.g. J. Polchinski, *Introduction to Gauge/Gravity Duality*, `arXiv:1010.6134` and H. Nastase, *Introduction to AdS-CFT*, `arXiv:0712.0689 [hep-th]` for accessible introductions to this by now vast subject.

39.2.1 Global Coordinates

Writing (39.10) as in (39.11) as

$$(z^1)^2 + \ldots + (z^4)^2 = +1 + (z^0)^2 \quad , \tag{39.12}$$

we see that the general solution is

$$z^0 = \sinh\tau \quad , \quad z^a = n^a \cosh\tau \quad , \tag{39.13}$$

where n^a, with $a = 1, \ldots, 4$, is a unit vector on S^3. This means that

$$\delta_{ab} n^a n^b = 1 \quad , \quad \delta_{ab} n^a dn^b = 0 \quad , \quad \delta_{ab} dn^a dn^b = d\Omega_3^2 \tag{39.14}$$

(analogous identities will be used repeatedly in the following). Then one finds the metric

$$ds^2 = \left(-(dz^0)^2 + \sum_a (dz^a)^2\right)|_{(39.10)} = -d\tau^2 + \cosh^2\tau\, d\Omega_3^2 \quad . \tag{39.15}$$

Remarks:

1. This is the $k = +1$, $\Lambda > 0$ solution (37.61) of the Friedmann equations, thus confirming that the solution found there is maximally symmetric.
2. The manifest symmetries in this coordinate system are the symmetries of S^3, i.e. the subgroup $SO(4) \subset SO(4,1)$ of the total isometry group (thus 6 out of 10 isometries are manifest).
3. It can be read off from (39.10) that these coordinates cover the hyperboloid globally (modulo the usual, and utterly harmless, issues with spherical coordinates at the poles of a sphere).
4. These are the Lorentzian-signature counterparts of the standard (hyper-)spherical coordinates on S^4, and are related to these by "Wick rotation" (continuation to imaginary time),

$$\tau = i\theta \quad \Rightarrow \quad -d\tau^2 + \cosh^2\tau\, d\Omega_3^2 \to d\theta^2 + \cos^2\theta\, d\Omega_3^2 = d\Omega_4^2 \tag{39.16}$$

(with $-\pi/2 \leq \theta \leq \pi/2$).

39.2.2 Conformal Time, Penrose Diagrams and Horizons

From the global coordinates introduced above we can pass to conformal time in order to then construct the Penrose diagram for de Sitter space. Thus we write

$$ds^2 = -d\tau^2 + \cosh^2\tau\, d\Omega_3^2 = \cosh^2\tau(-d\tau^2/\cosh^2\tau + d\Omega_3^2) \tag{39.17}$$

and introduce conformal time (usually called η, but I will call it T here) by the relation

$$dT = \frac{d\tau}{\cosh\tau} \quad . \tag{39.18}$$

Surprisingly this (hyperbolic) equation has the simple (trigonometric) solution

$$\cosh\tau = \frac{1}{\cos T} \tag{39.19}$$

with

$$\tau \in (-\infty, +\infty) \quad \Rightarrow \quad T \in (-\pi/2, +\pi/2) \quad . \tag{39.20}$$

Just as an aside, and for the record: with cosh replaced by sinh, the equation has a hyperbolic solution,

$$dR = \frac{d\rho}{\sinh\rho} \quad \Rightarrow \quad \sinh\rho = -\frac{1}{\sinh R} \quad . \tag{39.21}$$

Returning to the case at hand, we see that in terms of the new variable T the metric takes the form

$$ds^2 = \frac{1}{\cos^2 T}\left(-dT^2 + d\Omega_3^2\right) \quad . \tag{39.22}$$

Therefore de Sitter space is conformal to an interval $(-\pi/2, +\pi/2) \times S^3$ of the Einstein static universe (ESU) of section 37.2.

Writing the metric on the 3-sphere as usual as $d\psi^2 + \sin^2\psi d\Omega_2^2$, we see that de Sitter space is conformal to

$$d\tilde{s}^2 = -dT^2 + d\psi^2 + \sin^2\psi d\Omega_2^2 \quad , \tag{39.23}$$

with $\psi \in [0, \pi]$ and $T \in (-\pi/2, +\pi/2)$. Suppressing the transverse 2-sphere and adding the points with $T = \pm\pi/2$ (future and past infinity), we end up with the simple Penrose diagram of de Sitter space in Figure 54.

Now consider a comoving observer at the north pole $\psi = 0$. Such an observer will have both an event horizon (the boundary of that part of the universe that this observer can in principle obtain information about or be influenced by), and a particle horizon, which here we interpret as forming the boundary of the region which this observer can inprinciple have influence on. These horizons and "regions of influence" are indicated in the diagrams in Figure 55.

The intersection of these 2 regions is called the (northern) *causal diamond* and is the only region of de Sitter space that is fully accessible to an observer at the north pole in the sense that this is the region this observer can send signals to and receive signals from. This northern causal diamond is completely causally disconnected from the (southern) causal diamond of a comoving observer at the south pole $\psi = \pi$ (Figure 56).

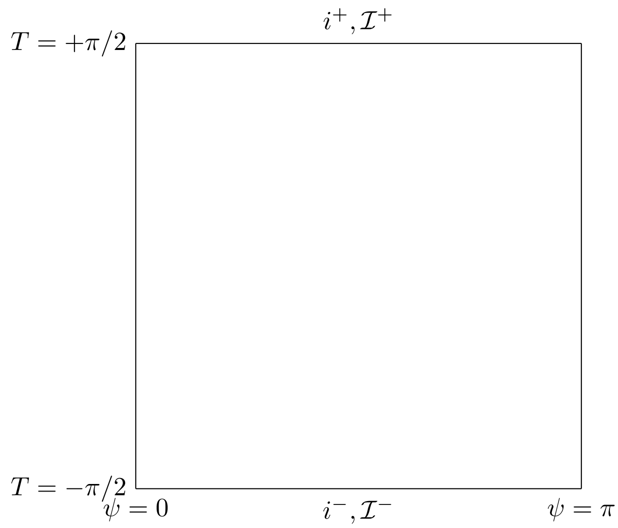


Figure 54: Penrose diagram for de Sitter space.

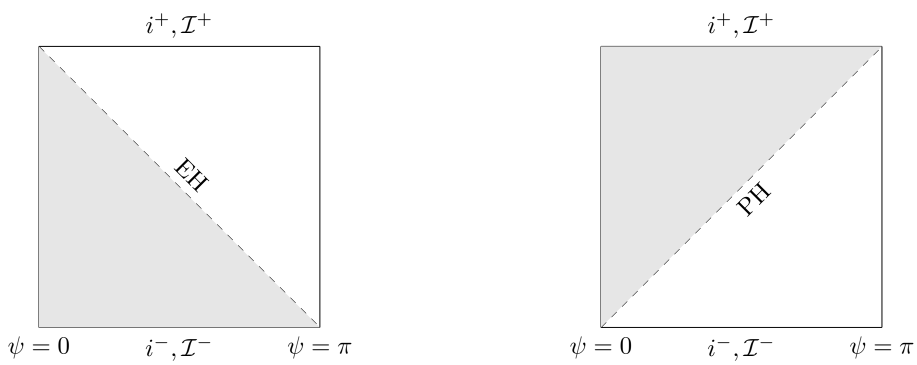


Figure 55: Event and Particle Horizons and "Regions of Influence" for a comoving observer at the north pole $\psi = 0$ in de Sitter space.

39.2.3 Hyperbolic Slicing Coordinates

We observe that (39.10) can also be written as

$$-(z^0)^2 + \sum_{k=1}^{3}(z^k)^2 \equiv \eta_{\alpha\beta} z^\alpha z^\beta = +1 - (z^4)^2 \quad . \tag{39.24}$$

For $z^4 > 1$, the right-hand side is negative and slices of constant z^4 are therefore hyperboloids H^3. In that case it is natural to introduce

$$z^4 = \cosh\tau \quad , \quad z^\alpha = n^\alpha \sinh\tau \quad \text{with} \quad \eta_{\alpha\beta} n^\alpha n^\beta = -1 \quad . \tag{39.25}$$

The n^α thus parametrise H^3, and one finds that the de Sitter metric can be written as

$$ds^2 = -d\tau^2 + \sinh^2\tau \, d\tilde{\Omega}_3^2 \quad , \tag{39.26}$$

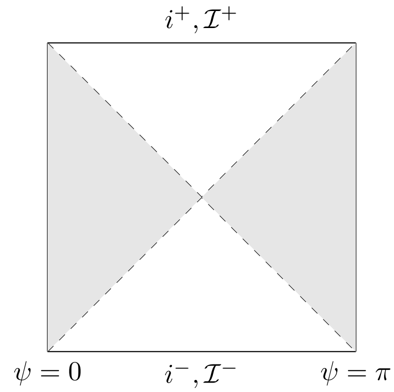


Figure 56: Northern and Southern Causal Diamonds in de Sitter space.

with $d\tilde{\Omega}_3^2$ the line-element on the unit hyperboloid in any coordinate system.

<u>Remarks:</u>

1. This is evidently the $k = -1$ metric (37.62).
2. The manifest symmetries in this coordinate system are the symmetries of H^3, i.e. the subgroup $SO(3,1) \subset SO(4,1)$ of the total isometry group (thus again 6 out of 10 isometries are manifest).
3. Note that these coordinates only cover the region $z^4 > 1$, which is only a part of de Sitter space. We will discuss coordinates in the range $|z^4| < 1$ below.

39.2.4 de Sitter Slicing Coordinates

Curiously, de Sitter space can be foliated by de Sitter spaces of one dimension less. To see this note that when $|z^4| < 1$ the right-hand side of (39.24) is positive. Thus the slices of constant z^4 are then indeed dS_3 spaces, and adapted coordinates are

$$z^\alpha = rn^\alpha \quad \text{with} \quad \eta_{\alpha\beta}n^\alpha n^\beta = 1 \quad , \quad z^4 = (1 - r^2)^{1/2} \tag{39.27}$$

with $0 < r < 1$, or

$$z^\alpha = \sin\psi n^\alpha \quad \text{with} \quad \eta_{\alpha\beta}n^\alpha n^\beta = 1 \quad , \quad z^4 = \cos\psi \quad . \tag{39.28}$$

The n^α parametrise dS_3, and one finds the metric

$$ds^2 = \frac{dr^2}{1 - r^2} + r^2 d\Omega_{1,2}^2 = d\psi^2 + \sin^2\psi\, d\Omega_{1,2}^2 \quad , \tag{39.29}$$

with $d\Omega_{1,2}^2$ the line-element on the unit curvature radius $(2 + 1)$-dimensional de Sitter space dS_3.

39.2.5 Planar Coordinates

In order to reproduce the spatially flat $k = 0$ metric (37.60), we introduce (admittedly with a certain amount of hindsight) coordinates t and x^k through

$$z^4 - z^0 = \mathrm{e}^{-t} \quad , \quad z^4 + z^0 = \mathrm{e}^{t} - \vec{x}^2 \mathrm{e}^{-t} \quad , \quad z^k = \mathrm{e}^{-t} x^k \quad . \tag{39.30}$$

It is evident that this also solves (39.10). Then one finds the metric

$$ds^2 = -dt^2 + \mathrm{e}^{-2t} d\vec{x}^2 \quad . \tag{39.31}$$

<u>Remarks:</u>

1. This is indeed the $k = 0$ metric (37.60).
2. By sending $t \to -t$, one sees that the metric can be written in either of the two ways
$$ds^2 = -dt^2 + \mathrm{e}^{\pm 2t} d\vec{x}^2 \quad . \tag{39.32}$$
3. The manifest symmetries in this coordinate system are the Euclidean group, i.e. the translational and rotational symmetries of $\mathbb{R}^3$, as well as the time-translation plus scaling symmetry
$$t \to t + \lambda \quad , \quad \vec{x} \to \mathrm{e}^{-\lambda} \vec{x} \quad . \tag{39.33}$$
Thus in this coordinate system, at least 7 of the 10 isometries are reasonably manifest.
4. These coordinates only cover the half-space
$$z^4 - z^0 = \mathrm{e}^{-t} \geq 0 \quad . \tag{39.34}$$
This is precisely one-half of de Sitter space, as indicated in Figure 57. To see this, note that from (39.13), with ψ the polar angle, one has
$$z^0 = \sinh\tau \quad , \quad z^4 = \cos\psi \cosh\tau \quad . \tag{39.35}$$
Together with (39.19)
$$\cosh\tau = \frac{1}{\cos T} \quad , \tag{39.36}$$
the condition $z^4 = z^0$ then translates into
$$z^4 = z^0 \quad \Leftrightarrow \quad \sin T = \cos\psi \quad , \tag{39.37}$$
which, in the range of (T, ψ) that we are considering, has the solution
$$T + \psi = \pi/2 \quad . \tag{39.38}$$
This is the diagonal line $t = \infty$ indicated in Figure 57.

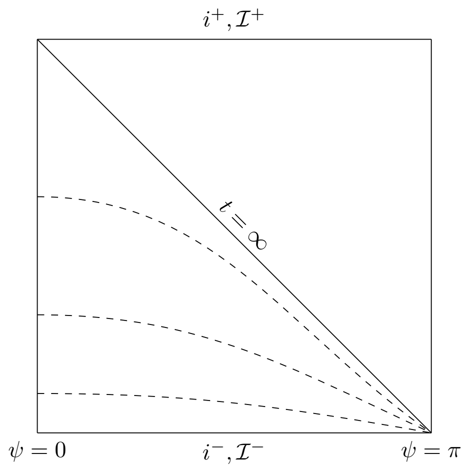


Figure 57: Planar Coordinates for de Sitter space cover "half" of de Sitter space. Indicated (schematically) are some lines of constant t. The event horizon of the comoving observer at the north pole is at $t = +\infty$.

5. Writing the metric (39.32) as

$$ds^2 = \mathrm{e}^{\pm 2t}(-\mathrm{e}^{\mp 2t}dt^2 + d\vec{x}^2) \tag{39.39}$$

and introducing a new time coordinate $\tau = \mp \exp \mp t$, the metric takes the form

$$ds^2 = \tau^{-2}(-d\tau^2 + d\vec{x}^2) \tag{39.40}$$

with $\tau \in (-\infty, 0)$ or $\tau \in (0, \infty)$ respectively. This is the positive curvature space-time counterpart of the generalised Poincaré upper half plane metric (11.74) and the de Sitter counterpart of the anti-de Sitter Poincaré coordinates (39.117) to be discussed below.

39.2.6 Static Coordinates

So far, the metric in all the coordinate systems was explicitly time-dependent (with a different time-coordinate in each case). It is possible to locally introduce a coordinate system that is time-independent. Namely, let us write (39.10) as

$$(z^1)^2 + (z^2)^2 + (z^3)^2 \equiv \sum_k (z^k)^2 = 1 + (z^0)^2 - (z^4)^2 \tag{39.41}$$

and introduce a spatial radial coordinate r via $\sum_k (z^k)^2 = r^2$. Then one has

$$\sum_k (z^k)^2 = r^2 \quad \Rightarrow \quad (z^4)^2 - (z^0)^2 = 1 - r^2 \quad . \tag{39.42}$$

Provided that $r^2 \leq 1$, a natural choice for the coordinates is

$$z^0 = (1-r^2)^{1/2}\sinh t \quad , \quad z^k = rn^k \quad , \quad z^4 = (1-r^2)^{1/2}\cosh t \quad , \tag{39.43}$$

leading to the metric

$$ds^2 = -(1-r^2)dt^2 + (1-r^2)^{-1}dr^2 + r^2 d\Omega_2^2 \quad , \tag{39.44}$$

the coordinate r taking the values $0 \leq r \leq 1$.

REMARKS:

1. In this coordinate system the metric is time-independent, i.e. ∂_t is a Killing vector.

2. It is adapted to a static and geodesic observer at $r = 0$ (corresponding to an observer at the north or south pole of the sphere in global coordinates, say) and covers that observer's causal diamond (Figure 56). Indeed, the equations for a radial geodesic read

$$(1-r^2)\dot{t} = E \quad , \quad \dot{r}^2 - r^2 = E^2 - 1 \quad . \tag{39.45}$$

In particular, for $E = 1$, one has $\dot{r} = \pm r$, and one particular solution is $r = \dot{r} = 0$, so $\dot{t} = 1$ and $t = \tau$ is proper time for this observer.

More generally, for the $E = 1$ (and $\dot{r} = +r$, say) observers one has

$$r(\tau) = r_0 \mathrm{e}^{\tau} \quad , \quad t(\tau) = \tau - \tfrac{1}{2}\ln(1 - r(\tau)^2) \quad . \tag{39.46}$$

These reach $r = 1$ in finite proper time $\tau = -\ln r_0$, while $t(\tau)$ diverges as $r \to 1$.

3. Of course $r = 1$ is not a real singularity but signals the location of an observer-dependent horizon beyond which a static coordinate system cannot be extended. Indeed, in spite of the many symmetries of the de Sitter metric, it has no everywhere timelike Killing vector, and the above is the best that one can do.

In fact, the above form of the static metric already suggests that the metric becomes time-dependent, with ∂_t a spacelike Killing vector, for $r > 1$. This can be confirmed explicitly by introducing coordinates that are appropriate in that range, namely instead of (39.43)

$$z^0 = (T^2-1)^{1/2}\cosh R \quad , \quad z^k = T n^k \quad , \quad z^4 = (T^2-1)^{1/2}\sinh R \quad , \tag{39.47}$$

leading to the metric

$$ds^2 = -(T^2-1)^{-1}dT^2 + (T^2-1)dR^2 + T^2 d\Omega_2^2 \quad . \tag{39.48}$$

Since $T^2 > 1$, it is also natural to parametrise $T = \cosh\tau$, so that

$$ds^2 = -d\tau^2 + \sinh^2\tau dR^2 + \cosh^2\tau d\Omega_2^2 \quad . \tag{39.49}$$

The existence of this alternative form of the metric beyond the horizon of the static metric is not surprising: we did (but then quickly dismissed as not particularly insightful) something analogous in the Schwarzschild case, introducing the time coordinate $T = r$ and the radial coordinate $R = t$ in the region $0 < r < 2m$ (27.1), and we will also briefly consider something analogous in the anti-de Sitter case below - see (39.144).

39.2.7 Eddington-Finkelstein and Kruskal-Szekeres Coordinates

As the de Sitter metric in static coordinates has the standard form of a static spherically symmetric metric, one can follow the general recipe outlined in sections 31.7 and 31.9 to construct the counterpart of Eddington-Finkelstein and Kruskal-Szekeres coordinates. The latter are occasionally used in discussions of the "thermodynamics" associated with the de Sitter cosmological horizon, as they bring out most clearly the analogies with the Schwarzschild event horizon.

We start with (39.44), with the inclusion of the curvature radius ℓ, corresponding to a cosmological constant $\Lambda = 3/\ell^2$, and introduce the corresponding tortoise coordinate r^* in the standard way via

$$\begin{aligned} ds^2 &= -(1-r^2/\ell^2)dt^2 + (1-r^2/\ell^2)^{-1}dr^2 + r^2 d\Omega_2^2 \\ &= (1-r^2/\ell^2)[-dt^2 + (1-r^2/\ell^2)^{-2}dr^2] + r^2 d\Omega_2^2 \\ &= (1-r^2/\ell^2)[-dt^2 + (dr^*)^2] + r^2 d\Omega_2^2 \ . \end{aligned} \tag{39.50}$$

In this case the relation

$$dr^* = (1-r^2/\ell^2)^{-1}dr \quad \Rightarrow \quad r^* = \tfrac{1}{2}\ell \ \log\frac{\ell+r}{\ell-r} \tag{39.51}$$

can be explicitly inverted to give r as a function of r^*,

$$r = \ell \tanh r^*/\ell \ . \tag{39.52}$$

Note that the horizon at $r=\ell$ corresponds to $r^* \to +\infty$.

Introducing in the usual way also the retarded and advanced coordinates $u = t - r^*, v = t + r^*$, one can now write the metric in Eddington-Finkelstein coordinates (u, r, θ, ϕ) or (v, r, θ, ϕ), leading to

$$ds^2 = -(1-r^2/\ell^2)dv^2 + 2dv\ dr + r^2 d\Omega_2^2 \tag{39.53}$$

(and likewise for the retarded coordinates). Kruskal-Szekeres coordinates can now be introduced by starting with

$$ds^2 = -(1-r^2/\ell^2)du\ dv + r^2 d\Omega_2^2 \ , \tag{39.54}$$

and noting that

$$1 - r^2/\ell^2 = \frac{1}{\cosh^2 r^*/\ell} = \frac{4\mathrm{e}^{(u-v)/\ell}}{(1+\mathrm{e}^{(u-v)/\ell})^2} \tag{39.55}$$

and

$$r = \ell\frac{(1-\mathrm{e}^{(u-v)/\ell})}{(1+\mathrm{e}^{(u-v)/\ell})} \ . \tag{39.56}$$

Then the substitution (note the sign change relative to the Kruskal-Szekers coordinates for the Schwarzschild metric)

$$U = \ell \mathrm{e}^{u/\ell} \quad , \quad V = -\ell \mathrm{e}^{-v/\ell} \tag{39.57}$$

leads to the Kruskal-Szekeres (double null) form of the de Sitter metric,

$$ds^2 = -\frac{4\ell^4}{(\ell^2 - UV)^2}\, dU dV + \ell^2 \frac{(\ell^2 + UV)^2}{(\ell^2 - UV)^2}\, d\Omega_2^2 \ , \tag{39.58}$$

with the past / future cosmological horizon at $U = 0$ or $V = 0$ respectively.

39.2.8 Interlude on (A)dS Schwarzschild

The above form (39.44) of the de Sitter metric in static coordinates should be familiar from (what appeared to be) quite a different context in section 24, in particular the discussion of Birkhoff's theorem in section 24.6. In that context we had noted that the characteristic form (24.75),

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega^2 \tag{39.59}$$

of the metric is implied not just by the vacuum Einstein equations for spherical symmetry but, more generally, by the Einstein equations in spherical symmetry whenever $T^t_{\ t} = T^r_{\ r}$. This condition is, in particular, satisfied, when the matter content is that of a cosmological constant,

$$T_{\mu\nu} = -\frac{\Lambda}{8\pi G_N} g_{\mu\nu} \quad \Rightarrow \quad T^t_{\ t} = T^r_{\ r} = -\frac{\Lambda}{8\pi G_N} \ . \tag{39.60}$$

Then the Einstein equations (24.71) reduce to the simple equation

$$m'(r) = 4\pi G_N\, r^2(-T^t_{\ t}) = \Lambda r^2/2 \tag{39.61}$$

for the mass function $m(r)$, which is evidently solved by

$$m(r) = m_0 + \Lambda r^3/6 = m_0 \pm r^3/2\ell^2 \ , \tag{39.62}$$

leading to

$$f(r) = 1 - \frac{2m(r)}{r} = 1 - \frac{2m_0}{r} \mp r^2/\ell^2 \ . \tag{39.63}$$

<u>Remarks:</u>

1. In particular, for $m_0 = 0$ and a positive cosmological constant we recover the above static metric (39.44) (for $\ell = 1$, but we could have done all the contructions for any ℓ).

2. Likewise, we learn that for $\Lambda < 0$ there will be a static coordinate system in which the anti-de Sitter metric takes the form (39.44) with $-r^2 \to +r^2$. We will recover and reconfirm this below, see (39.76), when studying the AdS metrics more systematically.

3. Remarkably, by including the integration constant m_0 we have actually found a more general class of solution of the Einstein equations with a cosmological constant, which are interesting in their own right, namely the *(A)dS Schwarzschild solutions*

$$ds^2 = -\left(1 - \frac{2m_0}{r} \mp r^2/\ell^2\right) dt^2 + \left(1 - \frac{2m_0}{r} \mp r^2/\ell^2\right)^{-1} dr^2 + r^2 d\Omega_2^2 \ , \quad (39.64)$$

already mentioned briefly in section 30.2. These describe a black hole of mass m_0 immersed in an (A)dS background.

39.2.9 Painlevé-Gullstrand-like Coordinates

In section 27.2 we had seen how to construct, from a static spherically symmetric metric of the standard (and, as it subsequently turned out, ubiquitous) form (39.59) e.g. a metric with with flat constant-time slices by a coordinate transformation $T(t,r) = t + \psi(r)$. In the case at hand (39.44), with $f(r) = 1 - r^2$, the condition (27.20) leads to

$$1 - C(r)^2 = f(r) \quad \Rightarrow \quad C(r) = \pm r \quad \Rightarrow \quad \psi'(r) = \pm r/(1-r^2) \ , \qquad (39.65)$$

which is solved by

$$\psi(r) = \pm\tfrac{1}{2}\ln(1-r^2) \ . \qquad (39.66)$$

Performing the coordinate transformation form the static metric (39.44) with this choice of $\psi(r)$ leads to the metric

$$\begin{aligned} ds^2 &= -(1-r^2)dT^2 \pm 2r dT dr + (dr^2 + r^2 d\Omega^2) \\ &= -dT^2 + (dr \pm r dT)^2 + r^2 d\Omega^2 \ , \end{aligned} \qquad (39.67)$$

which is the de Sitter analog of Painlevé-Gullstrand coordinates for the Schwarzschild metric, and which now allows one to extend the metric beyond the past or future horizon at $r = 1$.[180]

<u>Remarks:</u>

1. The Killing vector $\partial_T = \partial_t$ becomes space-like for $r > 1$, so while this is an extension of the static metric of de Sitter space, the metric is no longer static outside the original static patch.

[180] See M. Parikh, *New Coordinates for de Sitter Space and de Sitter Radiation*, `arXiv:hep-th/0204107` for some applications of this PG-form of the de Sitter line-element.

2. It is evident, by imitating the argument in (27.16), that T is nothing other than the proper time along the family of $E = 1$ geodesics (39.46) of the static metric,

$$-\dot{T}^2 + (\dot{r} - r\dot{T})^2 = -1 \quad \text{and} \quad T = \tau \quad \Rightarrow \quad \dot{r} = r \quad \text{and} \quad E = 1 \ . \tag{39.68}$$

3. This metric can also be obtained from the metric (39.32) in planar coordinates (we give these a subscript p now)

$$ds^2 = -dt_p^2 + \mathrm{e}^{\pm 2t_p} d\vec{x}^2 = -dt_p^2 + \mathrm{e}^{\pm 2t_p}(dr_p^2 + r_p^2 d\Omega^2) \tag{39.69}$$

   by the simple coordinate transformation

$$(T, r) = (t_p, \mathrm{e}^{\pm t_p} r_p) \quad \Rightarrow \quad ds^2 = -(1 - r^2)dT^2 \mp 2r dr dT + (dr^2 + r^2 d\Omega^2) \ . \tag{39.70}$$

   This also identifies the metric (39.67) as the special $H = \pm 1$ case (H the Hubble parameter) of the general PG-like form of cosmological Robertson-Walker metrics given in (34.43).

4. In an analogous way, one can construct de Sitter analogues of Eddington-Finkelstein coordinates etc.

Clearly, there are many more possibilities, but this shall suffice. It should be clear from the above examples how to construct other coordinate systems for dS adapted to one's needs.[181]

39.3 Some Coordinate Systems for anti-de Sitter space

Coordinates for anti-de Sitter space can be described in precise analogy with the de Sitter case. Our starting point is the defining equation (39.5),

$$-(z^0)^2 + (z^1)^2 + (z^2)^2 + (z^3)^2 - (z^4)^2 = -1 \ . \tag{39.71}$$

This has the topology $S^1 \times \mathbb{R}^3$, as can be seen by writing the equation as

$$(z^0)^2 + (z^4)^2 = 1 + (z^1)^2 + (z^2)^2 + (z^3)^2 \ . \tag{39.72}$$

For z^k fixed, this describes a circles in the (z^0, z^4)-plane. As the metric is negative-definite on that plane, this show that the surface defined by (39.71) has closed timelike curves through every point. To avoid such a paradoxical and pathological situation, we will "pass to the covering space", which amounts to replacing $S^1 \to \mathbb{R}$. It is actually

[181] And for more information about the interesting physics of de Sitter space see e.g. M. Spradlin, A. Strominger, A. Volovich, *Les Houches Lectures on de Sitter Space*, `arXiv:hep-th/0110007`.

this resulting space, without closed timelike curves, that is usually referred to as anti-de Sitter space, and we will follow that convention.

Here is an overview of the coordinate systems and topics discussed in this section:

1. Global (and Static) Coordinates
2. Conformal Coordinates, Conformal Boundary and Penrose Diagrams
3. Isotropic (Spatially Conformally Flat) Coordinates
4. Cosmological (Hyperbolic Slicing) Coordinates
5. de Sitter Slicing Coordinates
6. anti-de Sitter Slicing Coordinates
7. Poincaré Coordinates
8. Plane Wave AdS Coordinates
9. Codimension-2 Hyperbolic Slicing Coordinates
10. Painlevé-Gullstrand-like Coordinates?

39.3.1 Global (and Static) Coordinates

Global coordinates are provided by writing the general solution of (39.72) as

$$(z^0)^2 + (z^4)^2 = \cosh^2\rho \quad , \quad (z^1)^2 + (z^2)^2 + (z^3)^2 = \sinh^2\rho \ , \tag{39.73}$$

and by, in turn, writing the general solution to these equations as

$$z^0 = \cosh\rho\sin\tau \quad , \quad z^k = n^k\sinh\rho \quad , \quad z^4 = \cosh\rho\cos\tau \ . \tag{39.74}$$

Then one finds the metric

$$ds^2 = -\cosh^2\rho d\tau^2 + d\rho^2 + \sinh^2\rho d\Omega_2^2 \tag{39.75}$$

with $0 \leq \rho < \infty$ and $-\infty < \tau < +\infty$.

Remarks:

1. These coordinates again make the periodic nature of the time-direction of AdS manifest. The embedding hyperboloid would be covered by choosing τ to be an angular variable with period 2π. In the universal covering space, however, without closed timelike curves, τ and $\tau \pm 2\pi$ are not identified, and the range of τ is $-\infty < \tau < +\infty$.

2. Alternatively, one can write $r = \sinh\rho$, with $0 \leq r < \infty$, and now one recognises the metric

$$ds^2 = -(1+r^2)d\tau^2 + (1+r^2)^{-1}dr^2 + r^2 d\Omega_2^2 \ . \tag{39.76}$$

as the negative curvature counterpart of the static de Sitter metric (39.44), already anticipated in connection with the general solution (39.63) of the spherically symmetric Einstein equations with a cosmological constant. Notice in particular that this is of the general $f - f^{-1}$ form

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_2^2 \quad , \quad f(r) = 1 + r^2 \ . \tag{39.77}$$

3. Note that, in contrast to the dS case, AdS has a global timelike Killing vector, namely ∂_τ. It follows from the parametrisation of the embedding coordinates that τ-translations are the same thing as a rotation in the (negative definite) (z^0, z^4)-plane,

$$\partial_\tau = (\partial_\tau z^A)\partial_A = z^4\partial_0 - z^0\partial_4 \ , \tag{39.78}$$

and thus ∂_τ is identified with the Killing vector J_{04} (39.8) of the embedding space.

39.3.2 Conformal Coordinates, Conformal Boundary and Penrose Diagrams

Starting with the global metric above, which we write as

$$ds^2 = \cosh^2\rho(-d\tau^2 + d\rho^2/\cosh^2\rho + \tanh^2\rho d\Omega_2^2) \ , \tag{39.79}$$

we can introduce a "conformal radial" coordinate θ via

$$d\theta = \frac{d\rho}{\cosh\rho} \ , \tag{39.80}$$

which (as in the case of the conformal time coordinate of de Sitter space, cf. section 39.2.2), has the trigonometric solution

$$\cosh\rho = \frac{1}{\cos\theta} \ . \tag{39.81}$$

The difference is that here the range of ρ is mapped to

$$\cosh\rho \in [1, +\infty) \quad \Rightarrow \quad \theta \in [0, \pi/2) \ . \tag{39.82}$$

Using

$$\cosh\rho = \frac{1}{\cos\theta} \quad \Rightarrow \quad \sinh\rho = \tan\theta \quad , \quad \tanh\rho = \sin\theta \ , \tag{39.83}$$

the metric takes the form

$$\begin{aligned} ds^2 &= \frac{1}{\cos^2\theta}\left(-d\tau^2 + d\theta^2 + \sin^2\theta d\Omega_2^2\right) \\ &= \frac{1}{\cos^2\theta}\left(-d\tau^2 + d\Omega_3^2\right) \ . \end{aligned} \tag{39.84}$$

Thus anti-de Sitter space is conformal to

$$d\tilde{s}^2 = -d\tau^2 + d\Omega_3^2 \qquad \text{with } \ 0 \leq \theta < \pi/2 \ \ , \tag{39.85}$$

i.e. AdS is conformal to one half of the Einstein static universe (which has the standard range $0 \leq \theta \leq \pi$). Surfaces of constant τ are thus half-spheres (discs) with boundary at $\theta = \pi/2$, and one can visualise AdS as a solid cylinder infinitely extended in the time direction. The points with $\theta = \pi/2$ correspond to $\rho = \infty$, and $\theta = \pi/2$ is the *conformal boundary* $\mathcal{I}$ of AdS. This conformal boundary is timelike,

$$d\tilde{s}^2|_{\theta=\pi/2} = -d\tau^2 + d\Omega_2^2 \tag{39.86}$$

with topology

$$\mathcal{I} \approx S^2 \times \mathbb{R} \ \ , \tag{39.87}$$

and unites future and past null infinity $\mathcal{I}^{\pm}$ as well as spatial infinity i^0, symbolically

$$\mathcal{I} = \mathcal{I}^+ \cup \mathcal{I}^- \cup i^0 \ \ . \tag{39.88}$$

Likewise, for the $(d+2)$-dimensional anti-de Sitter space, say, one has

$$AdS_{d+2}: \quad \mathcal{I} \approx \mathbb{R} \times S^d \ \ , \tag{39.89}$$

and this conformal boundary can be regarded as the (spatial) conformal compactification of $(d+1)$-dimensional Minkowski space,

$$\mathbb{R}^{1,d} \to \mathbb{R} \times S^d \ \ . \tag{39.90}$$

Suppressing, as usual, the transverse 2- (or d-) sphere, the metric of this conformal completion of anti-de Sitter space is

$$d\tilde{s}^2 = -d\tau^2 + d\theta^2 \qquad \tau \in (-\infty, +\infty) \quad , \quad \theta \in [0, \pi/2] \ \ . \tag{39.91}$$

Since the range of τ is infinite while that of θ is finite, there is no way to compress this into a finite range of coordinates for both while preserving the condition that lightrays are diagonal. In other words, any further conformal transformation of the metric with line element $d\tilde{s}^2$ that maps the τ-interval to a finite range will squeeze the θ-interval to a point, which is not particularly helpful for visualisation purposes. Thus the best one can do is think of AdS as an infinite strip (or, as mentioned above, as an infinite solid cylinder), as displayed in Figure 58.

This diagram may not appear to be particularly informative at first sight. However, it displays and highlights several characteristic and peculiar features of AdS, in particular that AdS has a timelike boundary, i.e. a boundary with Lorentzian signature (a lower-dimensional space-time in its own right, equipped with a conformal class of metrics). In particular, starting with 5-dimensional AdS, the boundary $\mathcal{I}$ is a 4-dimensional space-time (which can be viewed as Minkowski space, with the space compactifed to a sphere).

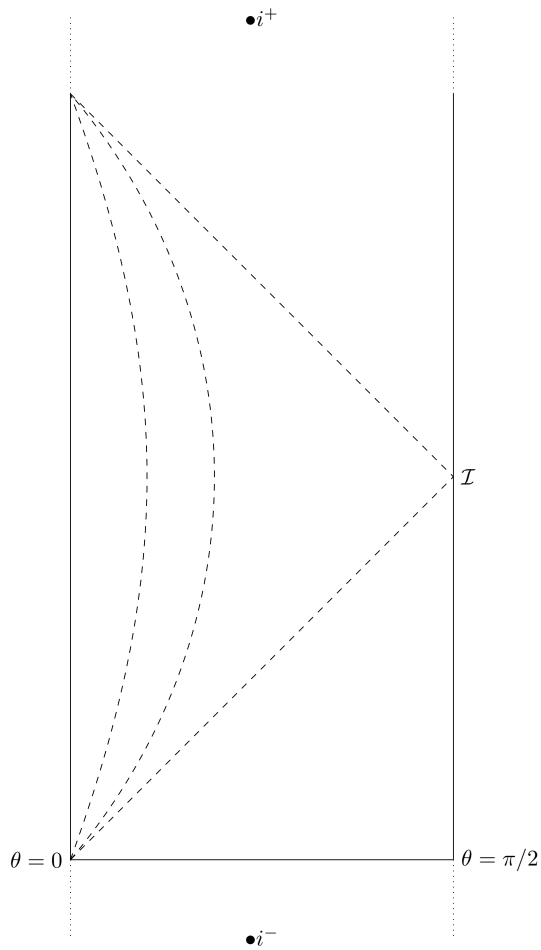


Figure 58: Penrose diagram of anti-de Sitter space. $\theta = 0$ represents the center (interior) of AdS, $\theta = \pi/2$ the timelike boundary $\mathcal{I}$. The diagram is infinitely extended to the future and past, with timelike future/past infinity $i^{\pm}$ residing there. Also indicated are a lightray "reflected" at $\mathcal{I}$ and some timelike geodesics.

Moreover, as is evident from the diagram (and we will confirm by a quick calculation below), lightrays can reach the boundary $\mathcal{I}$ ("infinity") in finite coordinate time. Indeed, radial lightrays are governed by the pair of equations

$$\cosh^2\rho\dot{\tau}^2 = \dot{\rho}^2 \quad , \quad \cosh^2\rho\dot{\tau} = E \ , \tag{39.92}$$

the first being the radial null condition and the second the conserved energy associated to τ-translation invariance. These can be combined into

$$\dot{\rho}^2 = E^2/\cosh^2\rho \quad \Rightarrow \quad \frac{d}{d\lambda}\sinh\rho = \pm E \ . \tag{39.93}$$

For outgoing lightrays (the plus-sign), one thus has

$$\sinh\rho(\lambda) = E(\lambda - \lambda_0) \ , \tag{39.94}$$

and thus $\rho \to \infty$ is reached for infinite values of the afffine parameter, $\lambda \to \infty$. For the coordinate time, on the other hand, one finds

$$\dot{\tau} = E/\cosh^2\rho \quad \Rightarrow \quad \tan\tau = \lambda - \lambda_0 \ . \tag{39.95}$$

Thus these null geodesics start at $\rho = 0$ at $\tau = 0$ and reach $\mathcal{I}$ ($\rho = \infty$) at $\tau = \pi/2$ (independently of E or λ_0).

This also means that for any spacelike hypersurface Σ (such as the horizontal line at the bottom of the diagram) there are points to the future of Σ which are such that there are past-directed causal (null) geodesics from that point that do not intersect that surface Σ (because they run into the boundary $\mathcal{I}$). This makes it plausible that specifying initial data for some fields (scalar fields, say) on some spacelike hypersurface alone is not enough to determine the future evolution of the field. Anti-de Sitter space is thus an example of a space-time which has no *Cauchy surfaces* which would lead to a well-defined Cauchy initial value problem (and one also says that such a space-time is not *globally hyperbolic*).

In embryonic form, this problem already arises for (null) geodesics, and in the diagram I have continued the lightray beyond $\mathcal{I}$ by adopting a particular prescription for evolving the lightray after it hits $\mathcal{I}$, namely "reflecting" boundary conditions at $\mathcal{I}$. It turns out that also for fields a well-defined evolution requires specifying not only initial data on some hypersurface but also boundary conditions on $\mathcal{I}$. In analysing this issue, the conformal relation between anti-de Sitter space and the Einstein static universe (which has Cauchy surfaces and a well-defined initial value problem) turns out to be valuable.[182]

Many of these things, in particular the existence of a timelike boundary on which fields "live" (namely the boundary values of the bulk fields), combined with the fact mentioned before that the isometry group of 5-dimensional AdS, $SO(4,2)$, coincides with the *conformal group* of the 4-dimensional (boundary) Minkowski space, are crucial basic ingredients in the celebrated AdS/CFT correspondence relating a gravitational (quantum) theory in the Anti-de Sitter "bulk" space-time to a conformal non-gravitational quantum field theory on the (conformal) boundary (see the references in footnote 179 for an introduction).

39.3.3 Isotropic (Spatially Conformally Flat) Coordinates

As in the case of the Schwarzschild metric (cf. (24.45)-(24.48)), it is straightforward to pass from the above standard static spherically symmetric form of the metric to the metric in isotropic (or spatially conformally flat) form. We again set $r = r(\rho)$, and write

$$f(r)^{-1}(dr^2 + f(r)r^2 d\Omega_2^2) = f(r)^{-1}(dr/d\rho)^2 \left(d\rho^2 + \left(r^2 f(r)/(dr/d\rho)^2\right) d\Omega^2\right) \ . \tag{39.96}$$

[182] S. Avis, C. Isham, D. Storey, *Quantum field theory in anti-de Sitter space-time*, Phys. Rev. D18 (1978) 3565-3576.

Requiring that the term in brackets equals the flat metric $d\rho^2 + \rho^2 d\Omega^2$ in polar coordinates, one finds the condition

$$\rho^2 (dr/d\rho)^2 = r^2 f(r) \quad , \tag{39.97}$$

which in the case at hand, $f(r) = 1 + r^2$, is solved by

$$r(\rho) = \frac{\rho}{1 - \rho^2/4} \quad , \tag{39.98}$$

with $0 \leq \rho \leq 2$. Thus

$$f(r) = 1 + r^2 = \frac{(1 + \rho^2/4)^2}{(1 - \rho^2/4)^2} \quad , \tag{39.99}$$

and

$$f(r)^{-1}(dr/d\rho)^2 = r^2/\rho^2 = (1 - \rho^2/4)^{-2} \quad , \tag{39.100}$$

so that in its full glory the anti-de Sitter metric in isotropic coordinates takes the form

$$\begin{aligned} ds^2 &= -\left(\frac{(1 + \rho^2/4)^2}{(1 - \rho^2/4)^2}\right) dt^2 + (1 - \rho^2/4)^{-2}(d\rho^2 + \rho^2 d\Omega^2) \\ &= -\left(\frac{(1 + \vec{x}^2/4)^2}{(1 - \vec{x}^2/4)^2}\right) dt^2 + (1 - \vec{x}^2/4)^{-2}\, d\vec{x}^2 \quad . \end{aligned} \tag{39.101}$$

Comparison with the standard metric on the hyperboloid in isotropic form (14.37), namely

$$ds^2 = (1 + k\vec{x}^2/4)^{-2}\, d\vec{x}^2 \quad . \tag{39.102}$$

for $k = -1$, shows that in these coordinates the slices of constant t are not just conformally flat (by construction) but actually hyperbolic maximally symmetric.

39.3.4 Cosmological (Hyperbolic Slicing) Coordinates

An alternative slicing of the anti-de Sitter metric by hyperboloids can be obtained by writing the defining equation (39.71) as

$$-(z^0)^2 + (z^1)^2 + (z^2)^2 + (z^3)^2 \equiv \eta_{\alpha\beta} z^\alpha z^\beta = (z^4)^2 - 1 \quad . \tag{39.103}$$

We see that the slices of constant z^4 are hyperboloids H^3 for $|z^4| < 1$ and de Sitter spaces dS_3 for $z^4 > 1$. In the fomer case, adapated coordinates to this slicing are

$$z^\alpha = \sin t\, n^\alpha \quad \text{with} \quad \eta_{\alpha\beta} n^\alpha n^\beta = -1 \quad , \quad z^4 = \cos t \quad , \tag{39.104}$$

leading to the metric

$$ds^2 = -dt^2 + \sin^2 t\, d\tilde{\Omega}_3^2 \tag{39.105}$$

which is precisely the $k = -1$, $\Lambda < 0$ solution (37.65) of the Friedmann equations. In contrast to the previous hyperbolic-slicing coordinates in (39.101), directly derived from the global coordinates, the present coordinates evidently do not cover the space-time globally (but only the patch $|z^4| < 1$).

39.3.5 de Sitter Slicing Coordinates

When $z^4 > 1$, the slices of constant z^4 are de Sitter spaces dS_3, and we obtain an analogue of the de Sitter Slicing coordinates (39.29) of de Sitter space. Corresponding adapted coordinates are

$$z^\alpha = \sinh\rho\, n^\alpha \quad \text{with} \quad \eta_{\alpha\beta}n^\alpha n^\beta = +1 \quad , \quad z^4 = \cosh\rho \ , \tag{39.106}$$

leading to the metric

$$ds^2 = d\rho^2 + \sinh^2\rho d\Omega_{1,2}^2 \ . \tag{39.107}$$

or, with $r = \sinh\rho$,

$$ds^2 = \frac{dr^2}{1+r^2} + r^2 d\Omega_{1,2}^2 \ . \tag{39.108}$$

We had seen in section 14, that the metrics on the three types of maximally symmetric Riemannian spaces $\mathbb{R}^3$, S^3 and H^3 could be written collectively as

$$\mathbb{R}^3, S^3, H^3: \qquad ds^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega_2^2 \ . \tag{39.109}$$

Analogously, we now see from (39.29), (39.108) and the form of the Minkowski metric (37.15) in Rindler-like coordinates, that the metrics on $\mathbb{R}^{1,3}$, dS_4 and AdS_4 can collectively be written as

$$\mathbb{R}^{1,3}, dS_4, AdS_4: \qquad ds^2 = \frac{dr^2}{1-kr^2} + r^2 d\Omega_{1,2}^2 \ , \tag{39.110}$$

39.3.6 anti-de Sitter Slicing Coordinates

Anti-de Sitter space also has a coordinate systems in which constant radial slices are themselves again anti-de Sitter spaces (of one dimension less, of course). This is obtained by a simple variant of the previous construction. Namely, instead of introducing n^α that parametrise a dS_3, split the defining equation as

$$-(z^0)^2 + (z^1)^2 + (z^2)^2 - (z^4)^2 = -1 - (z^3)^2 \tag{39.111}$$

and note that for fixed z^3 the left-hand side defines an AdS_3. Noting also that the right-hand side is ≤ -1, write

$$1 + (z^3)^2 = \cosh^2\rho \quad \Rightarrow \quad z^3 = \sinh\rho \tag{39.112}$$

and introduce the parametrisation

$$z^\alpha = \cosh\rho n^\alpha \quad (\alpha = 0,1,2,4) \quad \text{with} \quad \eta_{\alpha\beta}n^\alpha n^\beta = -1 \tag{39.113}$$

so that the n^α parametrise AdS_3, with

$$\eta_{\alpha\beta}dn^\alpha dn^\beta = d\tilde{\Omega}_{1,2}^2 \ . \tag{39.114}$$

Then one finds the metric

$$ds^2 = d\rho^2 + \cosh^2\rho d\tilde{\Omega}_{1,2}^2 \ . \tag{39.115}$$

39.3.7 Poincaré Coordinates

A somewhat unobvious but particularly interesting and useful way of parametrising the solution to (39.71) is to write (a certain amount of hindsight helps)

$$z^\alpha = rx^\alpha \;\; (\alpha = 0,1,2) \quad , \quad z^4 - z^3 = r \quad , \quad z^4 + z^3 = r^{-1} + r\eta_{\alpha\beta}x^\alpha x^\beta \tag{39.116}$$

Even though this is obscure, in these coordinates the metric takes the particularly simple and easy-to-use form

$$\begin{aligned} ds^2 &= \frac{dr^2}{r^2} + r^2\eta_{\alpha\beta}dx^\alpha dx^\beta \\ &= z^{-2}(\eta_{\alpha\beta}dx^\alpha dx^\beta + dz^2) \qquad (r = z^{-1}) \\ &= d\rho^2 + \mathrm{e}^{2\rho}\eta_{\alpha\beta}dx^\alpha dx^\beta \qquad (r = \mathrm{e}^\rho) \;\; . \end{aligned} \tag{39.117}$$

I have (somewhat redundantly) listed explicitly these 3 closely related parametrisations since all choices r, z and ρ are commonly found in the literature (with what I have here called z also frequently called r).

Remarks:

1. These are the AdS counterpart of the planar coordinates for dS space, and the space-time counterpart of the uppper-half-plane model of hyperbolic geometry discussed in section 11.3, see in particular (11.74). In these coordinates, a (2+1)-dimensional Poincaré symmetry is manifest, as well as a scaling symmetry $(x^\alpha, z) \to (\lambda x^\alpha, \lambda z)$. In this coordinate system it is also completely manifest that the metric is conformally flat, i.e. differs from the Minkowski metric only by an overall positive factor.

2. These coordinates do not cover all of AdS, as can for instance be seen by noting that radial null-geodesics can reach $r = 0$ (or $z = \infty$, say), at finite values of the affine parameter: null condition and conserved energy give (x^1 and x^2 are kept fixed)
$$\dot{r}^2 = r^4\dot{t}^2 \quad \text{and} \quad r^2\dot{t} = E \quad \Rightarrow \quad \dot{r}^2 = E^2 \;\; . \tag{39.118}$$
The solution for decreasing r is therefore
$$r(\lambda) = -E(\lambda - \lambda_0) \;\; , \tag{39.119}$$
which reaches $r = 0$ for $\lambda = \lambda_0$. Thus lightrays exit from the *Poincaré patch*, i.e. the region of the AdS spcae-time covered by these Poincaré coordinates, at finite values of the affine parameter. This boundary of the Poincaré patch at $r = 0$ or $z = \infty$ is also occasionally known as the *Poincaré horizon*.

3. Outgoing lightrays, with
$$r(\lambda) = r_0 + E\lambda \quad , \tag{39.120}$$
can reach $r = \infty$ only for infinite values of the affine parameter λ, $r(\lambda) = E(\lambda - \lambda_0)$. However, as in the calculation in global coordinates in section 39.3.2, $r = \infty$ is reached in finite coordinate time,
$$\begin{aligned} \dot{t} = E/r^2 = 1/(r_0 + E\lambda)^2 \quad &\Rightarrow \quad t(\lambda) = t_0 + E(1/r_0 - 1/(r_0 + E\lambda)) \\ &\Rightarrow \quad \lim_{\lambda\to\infty} t(\lambda) = t_0 \quad . \end{aligned} \tag{39.121}$$

4. The *conformal boundary* $\mathcal{I}$ resides at $r \to \infty$ or $z \to 0$. Up to an infinite conformal factor r^2 , the metric induced on $\mathcal{I}$ (rather, the part of $\mathcal{I}$ covered by these coordinates) is just the Minkowski metric with line-element $\eta_{\alpha\beta}dx^\alpha dx^\beta$. Thus, compared with the conformal boundary in global coordinates, the Poincaré patch just misses the "point at infinity" that compactifies the spatial directions to S^2 (cf. the discussion in section 39.3.2).

5. Writing
$$\eta_{\alpha\beta}dx^\alpha dx^\beta = -dt^2 + d\vec{x}^2 \quad , \tag{39.122}$$
the relation between this Poincaré time coordinate $x^0 = t$ and the global time coordinate τ in (39.75) is given by the (somewhat unobvious) relation
$$\tan\tau = \frac{2t}{1 + z^2 + \vec{x}^2 - t^2} \quad . \tag{39.123}$$

6. Occasionally it is useful to introduce null coordinates in the Poincaré plane spanned by the x^α, e.g. via
$$t = v + u/2 \quad , \quad x^2 = v - u/2 \tag{39.124}$$
(a more symmetric choice would of course have been possible). Relabelling the remaining coordinate $x^1 \to x$, the metric evidently takes the form
$$ds^2 = z^{-2}(-2du\, dv + dx^2 + dz^2) \quad . \tag{39.125}$$

7. We had already seen in section 37.5, in equations (37.60), (37.61) and (37.62), and then again in section 39.2 that the de Sitter metric could be written in such a way that the constant time slices are maximally symmetric spatial slices with either $k = 0$ (Planar Coordinates (39.32)), or $k = +1$ (Global Coordinates, (39.15)), or $k = -1$ (Hyperbolic Coordinates (39.26)).

   Analogously, the anti-de Sitter metric can be written in such a way that the metric on radial slices are maximally symmetric space-times with any sign of the curvature, $k = 0$ Minkowski space-time in Poincaré coordinates (39.117), $k = +1$ de Sitter slices in the coordinates (39.107), or $k = -1$ anti-de Sitter slices in the coordinates (39.115),
$$AdS_4: \qquad ds^2 = d\rho^2 + f_k(\rho)^2 d\Omega^2_{(k)} \quad , \tag{39.126}$$

where

$$f_k(\rho) = \begin{cases} \exp \pm\rho & \text{for } k=0 \\ \sinh\rho & \text{for } k=+1 \\ \cosh\rho & \text{for } k=-1 \end{cases} \quad \text{and} \quad d\Omega^2_{(k)} = \begin{cases} \eta_{\alpha\beta}dx^\alpha dx^\beta & \text{for } k=0 \\ d\Omega^2_{1,2} & \text{for } k=+1 \\ d\tilde{\Omega}^2_{1,2} & \text{for } k=-1 \end{cases} \tag{39.127}$$

39.3.8 Plane Wave AdS Coordinates

Starting with the null-form (39.125) of the AdS metric in Poincaré coordinates and performing the coordinate transformation

$$(u,v,x,z) = \big(\tan U, V + \tfrac{1}{2}(Z^2+X^2)\tan U, X/\cos U, Z/\cos U\big) \quad , \tag{39.128}$$

the AdS metric takes the form

$$ds^2 = Z^{-2}(-2dU\;dV - (X^2+Z^2)dU^2 + dX^2 + dZ^2) \quad . \tag{39.129}$$

This is the AdS metric in *plane wave coordinates* or the AdS plane wave metric (the reason for this nomenclature will be explained below). This metric can also be obtained directly from the embedding coordinates by solving (39.71) via the parametrisation

$$\begin{aligned} z^0 - z^1 &= \frac{\sin U}{Z} \quad , \quad z^0 + z^1 = \frac{2V\cos U + (Z^2+X^2)\sin U}{Z} \\ z^3 - z^4 &= \frac{\cos U}{Z} \quad , \quad z^3 + z^4 = \frac{2V\sin U - (Z^2+X^2)\cos U}{Z} \\ z^2 &= \frac{X}{Z} \quad . \end{aligned} \tag{39.130}$$

The metric in these coordinates has a number of remarkable properties:[183]

1. First of all, note that the plane wave AdS metric (39.129) differs from the null Poincaré metric (39.125) only by the 2nd term $\sim (X^2+Z^2)dU^2$. In spite of this, the global properties of this metric are very different from those of the metric in Poincaré coordinates. In particular, unlike the Poincaré coordinates, which only cover the Poincaré patch of the anti-de Sitter space-time, these plane wave coordinates provide a geodesically complete / global covering of the anti-de Sitter space-time. This can be seen
   - either by showing that all geodesics can be extended to infinite values of the affine parameter (essentially what happens is that the additional term in

[183] See e.g. M. Blau, J. Hartong,. B. Rollier, *Geometry of Schroedinger Space-Times, Global Coordinates, and Harmonic Trapping*, `arXiv:0904.3304 [hep-th]`, *Geometry of Schroedinger Space-Times II: Particle and Field Probes of the Causal Structure*, `arXiv:1005.0760 [hep-th]` (in these articles the coordinates denoted here by (U,Z) are called (T,R)) and references therein.

the metric acts like a harmonic oscillator potential for the geodesic equation which prevents geodesics from running off to infinity in finite affine parameter);

- or by convinicing oneself that the embedding (39.130) indeed covers all points of the hyperboloid (39.71).

2. The fact that this global metric is so similar to the Poincaré metric is in marked contrast to the relation between the Poincaré metric and the AdS metric in the usual global coordinates (39.76) which appears to bear no resemblance whatsoever to the Poincaré metric. It is also intriguing that the relation between Poincaré time t and the plane wave AdS time U given in (39.128) is so much simpler than the relation between Poincaré time and the usual global time coordinate τ (39.123),

$$\tan U = u \qquad \text{versus} \qquad \tan\tau = \frac{2t}{1+z^2+\vec{x}^2-t^2} \quad . \tag{39.131}$$

3. This issue can still be sharpened somewhat by introducing a parameter ω into the coordinate transformation (39.128) through

$$(u,v,x,z) = \left(\omega^{-1}\tan\omega U, V+\tfrac{1}{2}\omega(Z^2+X^2)\tan\omega U, X/\cos\omega U, Z/\cos\omega U\right) \quad , \tag{39.132}$$

leading to the 1-parameter family of metrics

$$ds^2 = Z^{-2}(-2dU\;dV - \omega^2(X^2+Z^2)dU^2 + dX^2 + dZ^2) \quad , \tag{39.133}$$

with a harmonic oscillator of frequency ω (thus the notation). This family of coordinate systems is geodesically complete for all $\omega > 0$ and interpolates between the Poincaré metric for $\omega = 0$ (the value for which also obligingly the coordinate transformation (39.132) reduces to the identity transformation) and the metric (39.129) for $\omega = 1$.

4. This raises the question if, despite their dissimilarity, a 1-parameter family of metrics can be found that interpolates between the AdS metric in Poincaré coordinates and the usual global coordinates. This is indeed possible (and not too hard once one knows that one should look for it). This metric can be found in the first reference in the preceding footnote 183.

5. Up to the overall factor of Z^{-2}, the metric in (39.129),

$$d\bar{s}^2 = -2dU\;dV - (X^2+Z^2)dU^2 + dX^2 + dZ^2 \quad . \tag{39.134}$$

is that of a *plane wave* in Brinkmann coordinates (43.19), whose general form is (with $v \to -v$ to match with the conventions used here)

$$d\bar{s}^2 = -2dudv + A_{ab}(u)x^a x^b du^2 + d\vec{x}^2 \quad . \tag{39.135}$$

Such metrics will be discussed in some detail in section 43. In this context it is well known that indeed the coeefficient matrix of du^2 acts as a harmonic oscillator potential. Moreover, the existence of the coordinate transformation (39.128) from or to Poincaré coordinates can be understood as an uplift to AdS of the coordinate transformation that exhibits the fact that isotropic plane waves (i.e. with the same frequencies in all directions) are conformally flat (43.46).

39.3.9 Codimension-2 Hyperbolic Slicing Coordinates

Yet another possibility is to split (39.71) as

$$-(z^0)^2 + (z^1)^2 + (z^2)^2 = -1 + (z^4)^2 - (z^3)^2 \quad . \tag{39.136}$$

Codimension-2 Slices of constant (z^3, z^4) are thus either 2-dimensional hyperboloids H^2 or 2-dimensional de Sitter space-times dS_2, depending on whether $(z^4)^2 - (z^3)^2 < 1$ (for H^2) or $(z^4)^2 - (z^3)^2 > 1$ (for dS_2). We will only consider the former case here (but it should by now be evident how to treat the latter, or other variations of this theme).

Even with the condition $(z^4)^2 - (z^3)^2 < 1$ it turns out that there are still two different cases to consider, namely either $(z^4)^2 - (z^3)^2 < 0$ or $0 < (z^4)^2 - (z^3)^2 < 1$.

- $(z^4)^2 - (z^3)^2 < 0$

  In this case we solve (39.136) in terms of a radial coordinate r by

  $$-(z^0)^2 + (z^1)^2 + (z^2)^2 = -r^2 \quad \text{and} \quad (z^3)^2 - (z^4)^2 = r^2 - 1 > 0 \quad . \tag{39.137}$$

  This is accomplished by writing $z^\alpha = rn^\alpha$ for $\alpha = 0, 1, 2$ where the n^α parametrise a hyperboloid H^2, i.e. $-(n^0)^2 + (n^1)^2 + (n^2)^2 = -1$, and identifying the time-coordinate with a boost in the (3,4)-plane, so that

  $$z^\alpha = rn^\alpha \quad , \quad z^3 = (r^2-1)^{1/2} \cosh t \quad , \quad z^4 = (r^2-1)^{1/2} \sinh t \quad . \tag{39.138}$$

  Then one finds the metric

  $$ds^2 = -(r^2-1)dt^2 + (r^2-1)^{-1}dr^2 + r^2 d\tilde{\Omega}_2^2 \quad . \tag{39.139}$$

  <u>Remarks:</u>

  1. Since $r \geq 1$, it is perhaps more natural to parametrise this as $r = \cosh \alpha$, upon which the metric takes the form

     $$ds^2 = -\sinh^2 \alpha dt^2 + d\alpha^2 + \cosh^2 \alpha \, d\tilde{\Omega}_2^2 \quad . \tag{39.140}$$

     This bears a superficial (and perhaps confusing) resemblance to global AdS coordinates (39.75), in which the metric was

     $$ds^2 = -\cosh^2 \rho d\tau^2 + d\rho^2 + \sinh^2 \rho \, d\Omega_2^2 \quad , \tag{39.141}$$

     (but note the exchange $\cosh \leftrightarrow \sinh$ and $S^2 \leftrightarrow H^2$).

2. For AdS, the metric in global coordinates (39.76), Poincaré coordinates (39.117) and the above hyperbolic coordinates (39.139) can be written collectively as

$$
\begin{aligned}
ds^2 &= -(r^2+k)dt^2 + (r^2+k)^{-1}dr^2 + r^2 d\Omega^2_{(k)} \\
d\Omega^2_{(k)} &= \begin{cases} d\Omega_2^2 & \text{for } k=+1 \\ d\vec{x}^2 & \text{for } k=0 \\ d\tilde{\Omega}_2^2 & \text{for } k=-1 \end{cases}
\end{aligned}
\tag{39.142}
$$

- $0 < (z^4)^2 - (z^3)^2 < 1$

  This corresponds to $0 < r^2 < 1$ in the above parametrisation, and the form (39.139) of the metric already suggests that r is really a time coordinate in this case (and what appeared as t above is a spatial coordinate). Indeed, with the parametrisation

  $$z^\alpha = Tn^\alpha \quad , \quad z^3 = (1-T^2)^{1/2}\sinh R \quad , \quad z^4 = (1-T^2)^{1/2}\cosh R \tag{39.143}$$

  (39.136) is also solved, and gives rise to the metric

  $$ds^2 = -(1-T^2)^{-1}dT^2 + (1-T^2)dR^2 + T^2 d\tilde{\Omega}_2^2 \ . \tag{39.144}$$

  Note that this metric is time-dependent but invariant under R-translations. It is the AdS analogue of the de Sitter metric (39.48).

39.3.10 Painlevé-Gullstrand-like Coordinates?

In section 39.2.9 we had introduced PG-like coordinates (39.67)

$$ds^2 = -(1-r^2)dT^2 \pm 2rdTdr + (dr^2 + r^2 d\Omega^2) \tag{39.145}$$

for de Sitter space by starting with the static spherically symmetric form (39.44)

$$ds^2 = -(1-r^2)dt^2 + (1-r^2)^{-1}dr^2 + r^2 d\Omega_2^2 \ . \tag{39.146}$$

of the metric (to which we could thus apply the general $t \to t + \psi(r)$ procedure outlined in section 27.2). We had also seen there that the PG coordinates could be thought of as interpolating between static and planar coordinates (39.69),

$$ds^2 = -dt^2 + \mathrm{e}^{2t}d\vec{x}^2 = -dt^2 + \mathrm{e}^{2t}(dr_p^2 + r_p^2 d\Omega_2^2) \ , \tag{39.147}$$

being related to them by the simple coordinate transformation $r = r_p \exp t, T = t$.

Is there a counterpart of these relations for anti-de Sitter? At first sight, the answer to this question seems to be a clear "no". Indeed, the counterpart of static coordinates for de Sitter space are the static spherically symmetric and global coordinates

$$ds^2 = -(1+r^2)dt^2 + (1+r^2)^{-1}dr^2 + r^2 d\Omega_2^2 \ . \tag{39.148}$$

of anti-de Sitter space, in which the metric takes the standard form (39.59), with $f(r) = 1 + r^2$. However, attempting to shift $t \to T(t,r) = t + \psi(r)$ in order to find a metric with flat constant T spatial slices, $g_{rr} = 1$, requires solving the condition (27.20), which in the present case reads

$$1 - C(r)^2 \stackrel{!}{=} f(r) = 1 + r^2 \quad . \tag{39.149}$$

This is evidently not possible, so in this strict sense there are no PG-like coordinates for anti-de Sitter space.

However, there is something analogous that one can do. Comparing the de Sitter planar coordinates (39.147) with the anti-de Sitter metric in Poincaré coordinates (39.117),

$$ds^2 = d\rho^2 + \mathrm{e}^{2\rho}(-dt^2 + d\vec{x}^2) \quad , \tag{39.150}$$

and introducing, in analogy with (39.147), polar Milne coordinates (section 37.1),

$$ds^2 = d\rho^2 + \mathrm{e}^{2\rho}(-d\tau^2 + \tau^2 d\tilde{\Omega}_2^2) \tag{39.151}$$

one sees that, roughly speaking (39.147) and (39.151) differ from each other by an exchange of a radial with a time coordinate. And indeed, taking this hint seriously, one can construct analogues of PG coordinates that are adapted to a suitable family of spacelike geodesics, and which restrict to the flat Minkowski metric on radial slices of constant R, with R being proper distance along this family of spacelike geodesics.

While this is a useful exercise, we can also turn the procedure around, i.e.

- start with the metric (39.151) in Poincaré / Milne coordinates and perform the coordinate transformation $T = \tau \exp\rho$ to obtain a PG-like metric (with roles of time and radius exchanged);
- find a new radial coordinate $R(\rho, T)$ through $\rho = R + \psi(T)$ such that the metric is again diagonal, say.

The resulting metric should then be the analogue of the static spherically symmetric de Sitter metric, and turns out to be the metric (39.144) (which is indeed the analogue of the continuation (39.48) of the static de Sitter metric (39.146) beyond the horizon).

Implementing the first step, from (39.151) one finds

$$\begin{aligned} \tau = \mathrm{e}^{-\rho}T \quad \Rightarrow \quad ds^2 &= (1-T^2)d\rho^2 + 2T\,dT\,d\rho + (-dT^2 + T^2 d\tilde{\Omega}_2^2) \\ &= d\rho^2 - (dT - T d\rho)^2 + T^2 d\tilde{\Omega}_2^2 \quad . \end{aligned} \tag{39.152}$$

In particular, in this PG-like metric the metric on slices of constant ρ is exactly the Minkowski metric (in Milne coordinates).

For the next step we consider the transformation

$$\rho = R + \psi(T) \quad . \tag{39.153}$$

Choosing $\psi(T)$ to satisfy

$$\dot{\psi} = -T/(1-T^2) \quad \Rightarrow \quad \psi(T) = -\tfrac{1}{2}\ln(1-T^2) \tag{39.154}$$

(note the analogy with (39.66)) one finds precisely the anti-de Sitter metric in the form (39.144),

$$ds^2 = -(1-T^2)^{-1}dT^2 + (1-T^2)dR^2 + T^2 d\tilde{\Omega}_2^2 \ . \tag{39.155}$$

Thus the PG-like AdS metric (39.152) interpolates between the Poincaré (planar) metric (39.151) and the metric (39.155).

To see the relation between this PG-like metric and spatial geodesics, we observe the following:

1. In the metric (39.155), radial spatial geodesics satisfy

   $$-(1-T^2)^{-1}(T')^2 + (1-T^2)(R')^2 = +1 \tag{39.156}$$

   a prime denoting a derivative with respect to proper distance σ.

2. There is a conserved momentum P conjugate to R,

   $$P = (1-T^2)R' \quad \Rightarrow \quad P^2 - (T')^2 = 1 - T^2 \ , \tag{39.157}$$

   and geodesics with $P = 1$ are characterised by $T' = \pm T$.

3. The coordinate ρ of the PG-like metric (39.152) is precisely the proper distance σ along the spacelike radial geodesics with $P = 1$ and $T' = +T$,

   $$(\rho')^2 - (T' - T\rho')^2 = 1 \quad \text{and} \quad \rho = \sigma \quad \Rightarrow \quad T' = T \quad \text{and} \quad P = 1 \ . \tag{39.158}$$

Clearly, there are many more possibilities, but this shall suffice. It should be clear from the above examples how to construct other coordinate systems for AdS adapted to one's needs.[184]

39.4 Warped Products, Cones, and Maximal Symmetry

We have seen in the previous sections that the metrics of maximally symmetric space-times can frequently be written in a way which exhibits their slicing by lower-dimensional maximally symmetric spaces or space-times.

[184] And for more erudite and advanced investigations of the AdS geometry see the 1998 lectures by G. Gibbons, *Anti-de-Sitter spacetime and its uses*, now available as `arXiv:1110.1206 [hep-th]`, and I. Bengtsson, *Anti-de Sitter Space*, 1998 Lecture Notes, available as `www.fysik.su.se/~ingemar/Kurs.pdf`.

Typically, in the codimension-1 case these metrics have the general form

$$dS_K^2 = \epsilon d\sigma^2 + f(\sigma)^2 ds_k^2 \ , \tag{39.159}$$

where

- dS_K^2 is the line-element of a maximally symmetric space(-time) with constant curvature K
- ds_k^2 is the line-element of a maximally symmetric space(-time) of one dimension less, with constant curvature k
- σ is a radial (space) or time coordinate, depending on whether $\epsilon = +1$ or $\epsilon = -1$
- $f(\sigma)$ is an elementary function, typically $f(\sigma) = \sigma$, or trigonometric, or hyperbolic.

For $\epsilon = +1$, thus $\sigma = r$ a radial coordinate, and $f(r) = r$, one obtains what is known as the metric on the cone over the space(-time) with line-element ds_k^2, in general given by

$$g_{ij}(x)dx^i dx^j \quad \rightarrow \quad \text{Cone Metric:} \quad ds^2 = dr^2 + r^2 g_{ij} dx^i dx^j \ . \tag{39.160}$$

A familiar example is the Euclidean metric on $\mathbb{R}^{n+1}$, which can be written in polar coordinates as the cone metric over S^n,

$$d\vec{x}_{n+1}^2 = dr^2 + r^2 d\Omega_n^2 \tag{39.161}$$

so in this case the "cone" has actually been flattened out to $\mathbb{R}^{n+1}$. However, if one were to replace S^n by a less symmetric space, there would be a (conical) singularity at the tip $r = 0$ of the cone.

Likewise, for $\epsilon = -1$, thus $\sigma = t$ a time coordinate, (39.159) reduces to the Robertson-Walker metrics (34.1)

$$ds^2 = -dt^2 + a(t)^2 ds_k^2 \tag{39.162}$$

of cosmology.

I will refer to the general class of metrics in (39.159) as (generalised) spacelike or timelike cone metrics. They can also be considered as special cases of so-called *warped product* metrics, which are metrics of the form

$$ds^2 = h_{ab}(y)dy^a dy^b + f(y)^2 g_{ij}(x) dx^i dx^j \ . \tag{39.163}$$

Thus this is the metric on a product

$$M = B \times_f F \tag{39.164}$$

of spaces or space-times, with h_{ab} a metric on the *base* B, g_{ij} a metric on the *fibre* F, the $\times_f$ indicating that M does not carry the direct product metric but that the metric on the fibre F is twisted or warped by the function $f(y)$ on the base B. It is a reasonably elementary exercise to work out the Riemann curvture tensor on M in terms of the curvature tensors of the metrics h_{ab} and g_{ij}, and the function f and its derivatives, but we will not consider the issue in this generality.

Rather, returning to the issue of writing maximally symmetric metrics in the form (39.159), we now want to address the question what determines in general what choice of $\epsilon, k, f(\sigma)$ is required or possible to realise a maximally symmetric space(-time) with a given K, say (or any variation of this question).

A quick way to answer this question is to make use of the fact that maximally symmetric spaces are characterised by the property of having constant curvature, in the sense of (14.10)

$$R_{\alpha\beta\gamma\delta} = k(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}) \ . \tag{39.165}$$

Using this result for the Riemann curvature tensor of ds^2_k, it is straightforward to calculate the curvature tensor of the generalised cone metric ds^2_K. This is just a minor generalisation of the calculation of the Riemann tensor of the Robertson-Walker metric

$$ds^2 = -dt^2 + a(t)^2 ds^2_k \tag{39.166}$$

(the $\epsilon = -1$ version of (39.159)) performed in section 35.1. Requiring that the cone metric with line-element ds^2_K be maximally symmetric, i.e. that its curvature tensor also has this form, one finds the constraint

$$(\text{K1}): \quad k = \epsilon(f')^2 + Kf^2 \tag{39.167}$$

between the various objects appearing in (39.159), as well as the condition

$$(\text{K2}): \quad f'' + (K/\epsilon)f = 0 \ . \tag{39.168}$$

A special (and especially boring) case that we will take care of (and dismiss) first is $f' = 0$, i.e. f constant, so that (39.159) describes a direct product metric. Then one finds $K = k = 0$, and this just corresponds to the possibilities

$$\begin{aligned} d\vec{x}^2_{n+1} &= +dr^2 + d\vec{x}^2_n \\ d\vec{x}^2_{1,n} &= +dr^2 + d\vec{x}^2_{1,n-1} \\ d\vec{x}^2_{1,n} &= -dt^2 + d\vec{x}^2_n \ . \end{aligned} \tag{39.169}$$

of trivially extending a flat space(-time) to a higher-dimensional flat space(-time) by adding a direction.

We now concentrate on $f' \neq 0$. Then (K1) implies (K2) (by differentiation) and conversely (K2) implies (K1), with k arising as an integration constant (in fact, (K1) is the energy/Hamiltonian corresponding to the equation of motion (K2)).

- $\epsilon = +1$: Spacelike Cones

  In this case, σ is a spatial coordinate (which we will call r or ψ or ρ) and we have

$$f'' + Kf = 0 \quad \text{and} \quad k = (f')^2 + Kf^2 \ . \tag{39.170}$$

  1. $K = 0$:

     The solution is

$$f = ar + b \quad \Rightarrow \quad k = a^2 \geq 0 \tag{39.171}$$

     for constants a and b, and dismissing the case of constant f we have already dealt with, we can without loss of generality choose $f(r) = r$, so that we are dealing with a standard cone metric. Then $k = +1$ and we find either the usual polar coordinate decomposition (39.161) of the Euclidean metric, or its Lorentzian signature counterpart

$$d\vec{x}^2_{1,n} = dr^2 + r^2 d\Omega^2_{1,n-1} \tag{39.172}$$

     with $d\Omega^2_{1,n-1}$ the de Sitter (Lorentzian positive curvature) line-element. This is the Minkowski metric in the Rindler-like coordinates (37.15).

  2. $K = +1$:

     The solution is

$$f = a\sin\psi + b\cos\psi \quad \Rightarrow \quad k = a^2 + b^2 > 0 \ . \tag{39.173}$$

     Without loss of generality we can choose $k = +1$ and (by a suitable shift of ψ) $a = 1, b = 0$. In Euclidean signature, this gives

$$d\Omega^2_{n+1} = d\psi^2 + \sin^2\psi \ d\Omega^2_n \tag{39.174}$$

     which is the standard nested form (14.31) of the line-element on a sphere, and its Lorentzian signature counterpart is

$$d\Omega^2_{1,n} = d\psi^2 + \sin^2\psi \ d\Omega^2_{1,n-1} \tag{39.175}$$

     which is the de Sitter metric in de Sitter slicing coordinates (39.29).

  3. $K = -1$:

     The solution is

$$f = ae^{\rho} + be^{-\rho} \quad \Rightarrow \quad k = -4ab \ , \tag{39.176}$$

     and in this case there are essentially 3 distinct choices of a and b, leading to the 3 different possible values of k.

     (a) $k = 0$: This arises for $a = 0$ or $b = 0$, thus $f(\rho) = \exp \pm\rho$, leading to

$$\begin{aligned} d\tilde{\Omega}^2_{n+1} &= d\rho^2 + \exp \pm 2\rho \ d\vec{x}^2_n \\ d\tilde{\Omega}^2_{1,n} &= d\rho^2 + \exp \pm 2\rho \ d\vec{x}^2_{1,n-1} \end{aligned} \tag{39.177}$$

     which is the metric on H^{n+1} or the AdS metric in Poincaré coordinates (39.117).

(b) $k = +1$: This arises for $a = 1/2$, $b = -1/2$ (say), thus $f(\rho) = \sinh\rho$, leading to

$$\begin{aligned} d\tilde{\Omega}^2_{n+1} &= d\rho^2 + \sinh^2\rho \, d\Omega^2_n \\ d\tilde{\Omega}^2_{1,n} &= d\rho^2 + \sinh^2\rho \, d\Omega^2_{1,n-1} \end{aligned} \tag{39.178}$$

which is the standard form (14.33) of the metric on H^{n+1} and the AdS metric in de Sitter slicing coordinates (39.107) respectively.

(c) $k = -1$: This arises for $a = b = 1/2$. Thus $f(\rho) = \cosh\rho$, and this gives rise to

$$\begin{aligned} d\tilde{\Omega}^2_{n+1} &= d\rho^2 + \cosh^2\rho \, d\tilde{\Omega}^2_n \\ d\tilde{\Omega}^2_{1,n} &= d\rho^2 + \cosh^2\rho \, d\tilde{\Omega}^2_{1,n-1} \end{aligned} \tag{39.179}$$

The former is the hyperbolic analogue of the nested coordinates for the sphere, the latter the AdS metric in AdS slicing coordinates (39.115) .

- $\epsilon = -1$: Timelike Cones

In this case, σ is a time coordinate (which we will call t), and in order for ds^2_K to have Lorentzian signature (and not two time directions), the metric to be warped (with line element ds^2_k) necessarily has Euclidean signature, so this reduces the number of possibilities somewhat compared to the case $\epsilon = +1$.

The equations governing this case are

$$f'' - Kf = 0 \quad \text{and} \quad k = -(f')^2 + Kf^2 \ . \tag{39.180}$$

1. $K = 0$:

   The solution is

$$f = at + b \quad \Rightarrow \quad k = -a^2 \leq 0 \tag{39.181}$$

   and discarding the case of constant f we are left with $f(t) = t$ and $k = -1$. The corresponding metric

$$d\vec{x}^2_{1,n} = -dt^2 + t^2 d\tilde{\Omega}^2_n \tag{39.182}$$

   is the Minkowski metric in Milne coordinates (37.4).

2. $K = +1$:

   The solution is

$$f = a\mathrm{e}^{t} + b\mathrm{e}^{-t} \quad \Rightarrow \quad k = 4ab \tag{39.183}$$

   and in this case there are 3 distinct choices of a and b, leading to the 3 different possible values of k.

   (a) $k = 0$: This arises for $a = 0$ or $b = 0$, thus $f(t) = \exp \pm t$, and one finds

$$d\Omega^2_{1,n} = -dt^2 + \mathrm{e}^{\pm 2t} d\vec{x}^2_n \tag{39.184}$$

   which is the de Sitter metric in planar coordinates (39.32), equivalently the $k = 0$ cosmological FLRW solution for a positive cosmological constant.

(b) $k=+1$: This arises for $a=b=1/2$, say, so the solution is $f(t)=\cosh t$, leading to the de Sitter metric in global coordinates (39.15),

$$d\Omega^2_{1,n} = -dt^2 + \cosh^2 t\, d\Omega^2_n \qquad (39.185)$$

or, equivalently, to the $k=+1$ cosmological FLRW solution for a positive cosmological constant.

(c) $k=-1$: This arises for $a=-b=1/2$, i.e. $f(t)=\sinh t$, leading to

$$d\Omega^2_{1,n} = -dt^2 + \sinh^2 t\, d\tilde{\Omega}^2_n \qquad (39.186)$$

which is the de Sitter metric in hyperbolic slicing coordinates (39.26) or, equivalently, the $k=-1$ cosmological FLRW solution for a positive cosmological constant.

3. $K=-1$:

The solution is

$$f = a\sin t + b\cos t \quad \Rightarrow \quad k = -(a^2+b^2) < 0 \ . \qquad (39.187)$$

Choosing $a=1, b=0$, with $k=-1$, one recovers the AdS metric in hyperbolic slicing coordinates (39.105),

$$d\tilde{\Omega}^2_{1,n} = -dt^2 + \sin^2 t d\tilde{\Omega}^2_n \ , \qquad (39.188)$$

equivalently the $k=-1$ cosmological FLRW solution for a negative cosmological constant.

We thus see that we have been able to reproduce many of the metrics found in sections 39.2 and 39.3 from this more general perspective, perhaps shedding some light on the zoo of coordinate systems found there. In particular, we have seen how the conditions (K1) (39.167) and (K2) (39.168) correlate the choice of curvatures k and K, the signature ϵ of the cone direction, and the choice of warping function f.

All of this can be straightforwardly generalised to multiple warpings / cones related to slicings of higher codimensions, but we will not pursue this here since it is not particularly enlightning in its own right and simply requires a steady hand.

40 Vaidya Metrics I: Bondi Gauge and Radiation Fields

Vaidya metrics are a (deceptively) simple generalisation of the Schwarzschild metric written in ingoing or outgoing Eddington-Finkelstein coordinates, in which the constant mass m is replaced by a mass function $m(v)$ or $m(u)$ depending on an advanced or retarded time coordinate, so that the metrics have the form (30.17)

$$\begin{aligned} ds^2 &= -f(v,r)dv^2 + 2dvdr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m(v)}{r} \\ ds^2 &= -f(u,r)du^2 - 2dudr + r^2 d\Omega^2 \quad , \quad f(u,r) = 1 - \frac{2m(u)}{r} \quad . \end{aligned} \tag{40.1}$$

The Vaidya metrics already made a brief appearance in the list of generalisations of the Schwarzschild metric in (30.17) in section 30.2. As already mentioned there, these metrics can be used as toy models to describe the formation of a black hole through ingoing null matter (for $m = m(v)$), or the radiation of a star for $m = m(u)$. They also provide a classical setting for (toy-)modelling an evaporating black hole. We also discussed some aspects of the Vaidya metric, related to trapped surfaces and horizons, in sections 32.8, 32.9 and 32.10.

As the Vaidya metrics are, for some reason, not discussed at any length (actually hardly mentioned) in any of the standard textbooks I am aware of, I will attempt to explain some of the elementary aspects of these metrics in some detail in this and subsequent sections, with occasional pointers to the literature for more detailed and advanced investigations.

40.1 Introduction: Ingoing and Outgoing Vaidya Metrics

To set the stage, we introduce the Vaidya metrics and list some of their basic properties.

1. The *Ingoing Vaidya Metric*:

   The metric has the form

   $$ds^2 = -f(v,r)dv^2 + 2dvdr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m(v)}{r} \quad . \tag{40.2}$$

   It has the following basic properties:

   (a) It reduces to the Schwarzschild metric in ingoing Eddington-Finkelstein coordinates for constant mass function $m(v) = m_0$.

   (b) It has the characteristic Kerr-Schild "flat + null" form (27.130) and (27.136)

   $$g_{\alpha\beta} = \eta_{\alpha\beta} + \frac{2m(v)}{r}\partial_\alpha v \partial_\beta v \tag{40.3}$$

   where $\eta_{\alpha\beta}$ is the Minkowski metric, written here in advanced coordinates,

   $$\eta_{\alpha\beta}dx^\alpha dx^\beta = -dv^2 + 2dvdr + r^2 d\Omega^2 \quad , \tag{40.4}$$

and

$$\eta^{\alpha\beta}\partial_\alpha v \partial_\beta v = \eta^{vv} = 0 \ . \tag{40.5}$$

This can also be made more explicit in terms of the Kerr-Schild (or Eddington) time coordinate $\tilde{t}$ defined by $v = \tilde{t} + r$, in terms of which the metric takes the form

$$ds^2 = -d\tilde{t}^2 + dr^2 + r^2 d\Omega^2 + \frac{2m}{r} d(\tilde{t} + r)^2 \ . \tag{40.6}$$

For the Schwarzschild metric, $\tilde{t}$ is simply related to the Schwarzschild coordinate t and the tortoise coordinate r^* by $v = t + r^* = \tilde{t} + r$.

(c) The curvature tensor simplifies accordingly, and the only non-vanishing component of the Einstein tensor $G_{\alpha\beta}$ is G_{vv}, with

$$G_{vv} = \frac{2m'(v)}{r^2} \ . \tag{40.7}$$

Equivalently, the only non-vanishing component of $G^\alpha_{\ \beta}$ is

$$G^r_{\ v} = \frac{2m'(v)}{r^2} \ . \tag{40.8}$$

Thus (40.2) solves the Einstein equations for an energy-momentum tensor of the form

$$T_{\alpha\beta} = \frac{m'(v)}{4\pi G_N r^2} \delta^v_\alpha \delta^v_\beta \ . \tag{40.9}$$

(d) By construction, this energy-momentum tensor is covariantly conserved, $\nabla_\alpha T^\alpha_{\ \beta} = 0$. Explicitly, the only non-trivial component of this equation is

$$\nabla_\alpha T^\alpha_{\ v} = 0 \quad \Leftrightarrow \quad r^{-2}\partial_r(r^2 T_{vv}) = \partial_r T_{vv} + \frac{2}{r} T_{vv} = 0 \ , \tag{40.10}$$

which is evidently satisfied for any energy-momentum tensor $\sim r^{-2}$.

(e) If $T_{\alpha\beta}$ is to satisfy some reasonable energy-condition (cf. section 22.1) like the null energy condition (NEC), which requires positivity of T_{vv}, one needs a non-decreasing mass, $m'(v) \geq 0$. The ingoing Vaidya metric thus describes the metric of a star or black hole with infalling *null dust* or *incoherent radiation*.

(f) The Kretschmann scalar (8.61) is

$$K \equiv R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} = 48 m(v)^2 / r^6 \ , \tag{40.11}$$

so that these metrics are singular at $r = 0$ for any non-trivial $m(v)$.

2. The *Outgoing Vaidya Metric*:

The metric has the form

$$ds^2 = -f(u,r)du^2 - 2du dr + r^2 d\Omega^2 \quad , \quad f(u,r) = 1 - \frac{2m(u)}{r} \ . \tag{40.12}$$

Its properties are, mutatis mutandis, largely analogous to those of the ingoing Vaidya metric, for example:

(a) It reduces to the Schwarzschild metric in outgoing Eddington-Finkelstein coordinates for constant mass function $m(u) = m_0$.

(b) The only non-vanishing component of the Einstein tensor is G_{uu}, with

$$G_{uu} = -\frac{2m'(u)}{r^2} \quad . \tag{40.13}$$

Thus (40.12) solves the Einstein equations for an energy tensor of the form

$$T_{\alpha\beta} = -\frac{m'(u)}{4\pi G_N r^2}\delta^u_\alpha \delta^u_\beta \quad . \tag{40.14}$$

Again this is conserved due to the r^{-2}-behaviour,

$$\partial_r T_{uu} + \frac{2}{r}T_{uu} = 0 \quad . \tag{40.15}$$

(c) If $T_{\alpha\beta}$ is to satisfy some reasonable energy-condition (like positivity of T_{uu}), one needs a non-increasing mass, $m'(u) \leq 0$. The outgoing Vaidya metric thus describes the metric of a radiating star (or, possibly, of an evaporating black hole).

A generalisation of the Vaidya metric can be obtained by allowing the mass function to be an arbitrary function of the coordinates (v, r) or (u, r):

3. The *Generalised Vaidya Metric*:

The metric has the form

$$ds^2 = -f(v,r)dv^2 + 2dvdr + r^2 d\Omega^2 \quad , \quad f(v,r) = 1 - \frac{2m(v,r)}{r} \tag{40.16}$$

(or its outgoing counterpart). In this case, one finds by a straightforward but uninspiring calculation that the non-vanishing components of the Einstein tensor $G^\alpha_{\ \beta}$ are

$$\begin{aligned} G^\theta_{\ \theta} = G^\phi_{\ \phi} &= -\frac{\partial_r^2 m(v,r)}{r} \\ G^r_{\ r} = G^v_{\ v} &= -\frac{2\partial_r m(v,r)}{r^2} \\ G^r_{\ v} &= +\frac{2m'(v,r)}{r^2} \end{aligned} \tag{40.17}$$

(a prime still denotes a v-derivative, partial r-derivatives are written explicitly).

Special cases of this generalised Vaidya metric include

(a) the Vaidya-Kottler (or Vaidya-(anti-)de Sitter) metrics with

$$f(v,r) = 1 - \frac{2m(v)}{r} - \frac{\Lambda r^2}{3} \quad , \tag{40.18}$$

which solve the Einstein equations for an energy-momentum tensor that is the sum of the Vaidya energy-momentum tensor (ingoing null dust) and a cosmological constant;

(b) the Vaidya-Reissner-Nordstrøm metrics with

$$f(v,r) = 1 - \frac{2m(v)}{r} + \frac{q^2}{r^2} \quad , \tag{40.19}$$

which solve the Einstein equations for an energy-momentum tensor that is the sum of the Vaidya energy-momentum tensor (ingoing null dust) and the electrostatic Maxwell energy-momentum tensor for a point charge with constant charge q;

(c) the Vaidya-Bonnor metrics with

$$f(v,r) = 1 - \frac{2m(v)}{r} + \frac{q(v)^2}{r^2} \quad , \tag{40.20}$$

which further generalise this to allow for an injection of charge in addition to mass into the star or black hole.[185]

By construction the energy-momentum tensor $T_{\alpha\beta} = G_{\alpha\beta}/8\pi G_N$ of the generalised Vaidya metric as defined above is conserved. However, only for very special choices of mass function $m(v,r)$ (as in the above examples) will one have a reasonable and reasonably physical matter content.[186]

40.2 Interpretation(s) of the Vaidya Mass Functions $m(v)$ and $m(u)$

The interpretation of the mass function $m(v)$ (or $m(u)$) of the Vaidya metric is brought out most clearly by noting that, as in (24.79),

$$g^{rr} = 1 - \frac{2m(v)}{r} \quad \text{or} \quad g^{rr} = 1 - \frac{2m(u)}{r} \quad , \tag{40.21}$$

so that $m(v)$ (or $m(u)$) is the invariantly defined Misner-Sharp mass (24.82) for spherical symmetry, measuring the amount of mass enclosed by the 2-sphere of constant v (or u) and r,

$$M_{MS}(v,r) = m(v) \quad , \quad M_{MS}(u,r) = m(u) \quad . \tag{40.22}$$

Introducing a time-coordinate $\tilde{t}$ by the relation

$$\tilde{t} = v - r \quad \text{or} \quad \tilde{t} = u + r \tag{40.23}$$

(this is modelled on and reduces to the Kerr-Schild (or Eddington) time-coordinate $\tilde{t}$,

$$v = t + r^* = \tilde{t} + r \quad \text{or} \quad u = t - r^* = \tilde{t} - r \tag{40.24}$$

[185] For a study of gravitational collapse respectively black hole evaporation in this setting see e.g. K. Lake, T. Zannias, *Strucutre of singularities in the spherical gravitational collapse of a null fluid*, Phys. Rev. D43 (1991) 1798-1802, and M. Parikh, F. Wilczek, *Global Structure of Evaporating Black Holes*, `arXiv:gr-qc/9807031`.

[186] For some other applications and appearances of the generalised Vaidya metric, see e.g. V. Husain, *Exact solutions for null fluid collapse*, `arXiv:gr-qc/9511011`, and A. Wang, Y. Wu, *Generalized Vaidya Solutions*, `arXiv:gr-qc/9803038`.

for the Schwarzschild metric introduced in (27.123) and (27.134)), the mass function $m = m(\tilde{t} \pm r)$ acquires the interpretation of the amount of mass enclosed by a 2-sphere of constant $\tilde{t}$ and r. Considering a fixed time-slice $\tilde{t} = \tilde{t}_0$ and taking $r \to \infty$ requires taking $v \to \infty$ respectively $u \to -\infty$. In the limit $r \to \infty$, the Misner-Sharp mass reduces to the ADM mass or energy (cf. sections 23.4 and 24.8) so that

$$E_{ADM} = m(v = +\infty) \quad \text{or} \quad E_{ADM} = m(u = -\infty) \tag{40.25}$$

(thus for physically meaningful space-times the mass function m should be bounded as $v \to \infty$ or $u \to -\infty$).

That this ADM mass limit of the mass function is indeed a conserved quantity (independent of the chosen time $\tilde{t}_0$) can be understood from the observation that at any given time a spacelike hypersurface will intersect all the constant v null worldlines along which null matter flows into the star or black hole. This is identical to the total mass of the black hole as $v \to \infty$. Likewise for the outgoing Vaidya metric a spatial slice extending to infinity will intercept all the outgoing null lines of constant u along which null matter escapes from the star, but the total energy (given by the initial mass $m(u = -\infty)$) will be conserved. At any finite r, thus finite u or v, the mass function can then be interpreted as the enclosed mass in a sphere of radius r at the time $\tilde{t}_0$.

If, instead of going to spatial infinity one goes to null infinity, instead of the ADM mass one has the so-called *Bondi-Sachs mass* $M_{BS}(u)$ at one's diposal (with u thought of as a coordinate at future null infinity $\mathcal{I}^+$ labelling the outgoing null geodesics of constant u). In particular, now keeping u fixed and taking $r \to \infty$ one finds that $m(u)$ agrees with the Bondi-Sachs mass at future null infinity,

$$M_{BS}(u) = m(u) \quad . \tag{40.26}$$

It keeps track of the mass decrease through the amount of radiation that escapes to infinity as the mass $m(u)$ decreases from its initial value $m(u = -\infty)$. This can be seen by writing the expression (40.14) for the energy-momentum tensor in terms of the (outgoing null) energy density ρ_{out} as

$$\rho_{out} = -\frac{m'(u)}{4\pi G_N r^2} \quad , \tag{40.27}$$

so that the total flux

$$\mathcal{F} = 4\pi r^2 \rho_{out} \tag{40.28}$$

through a sphere of radius r is independent of r and satisfies

$$\frac{d}{du} m(u) = -G_N \, \mathcal{F} \quad . \tag{40.29}$$

40.3 Einstein Equations in the Bondi Gauge (Radiative Coordinates)

All of the above are (of course) special cases of the general spherically symmetric metric, which can be conveniently parametrised in terms of 2 arbitrary functions $f(w,r)$ and $h(w,r)$ as

$$ds^2 = -e^{2h(w,r)}f(w,r)dw^2 + 2\epsilon\, e^{h(w,r)}dw\,dr + r^2 d\Omega^2 \quad , \tag{40.30}$$

where $\epsilon = \pm 1$, and we again parametrise $f(w,r)$ in terms of a mass function $m(w,r)$ as

$$f(w,r) = 1 - \frac{2m(w,r)}{r} \quad . \tag{40.31}$$

This is the general spherically symmetric metric written in radiative coordinates, or in the so-called *Bondi gauge*, the retarded / advanced Eddington-Finkelstein-like counterpart of the Schwarzschild-Birkhoff ansatz (24.68)

$$ds^2 = -e^{2h_s(t,r)}f_s(t,r)dt^2 + f_s(t,r)^{-1}dr^2 + r^2 d\Omega^2 \quad , \tag{40.32}$$

with

$$f_s(t,r) = 1 - \frac{2m_s(t,r)}{r} \tag{40.33}$$

(the subscript $_s$ on the functions indicating that these are a priori not the same functions as those appearing in the ansatz (40.30)).

<u>Remarks:</u>

1. The Bondi gauge is adapted to radial null geodesics in the sense that $w = const.$ are radial null rays. For $f(w,r) > 0 \leftrightarrow r > 2m(w,r)$, r decreases along future-directed null rays with $w \equiv v$ constant for $\epsilon = +1$ (ingoing coordinates), and r increases along future-directed null rays with $w \equiv u$ constant for $\epsilon = -1$ (outgoing coordinates).

2. Within the Bondi gauge, there is still the freedom of reparametrising w: any change of variables of the form

$$w \to \tilde{w}(w): \quad d\tilde{w} = e^{h(w)}dw \tag{40.34}$$

can be compensated (or induced) by a shift of $h(w,r)$,

$$h(w,r) \to h(w,r) + h(w) \quad . \tag{40.35}$$

3. Likewise, in the Schwarzschild gauge there is a corresponding freedom to reparametrise the time coordinate t by $t \to \tilde{t}(t)$. In the case of a static metric (timelike Killing vector), there is a preferred class of parametrisations, such that the Killing vector is $\sim \partial_t$ (equivalently that the metric is t-independent). This determines t up to affine transformations. If furthermore the metric is asymptotically flat, the

normalisation can be fixed by requiring e.g. that the norm of the Killing vector $||\partial_t||^2 \to -1$ asymptotically, and this then only leaves the (unavoidable because of time-translation invariance) ambiguity $t \to t + t_0$ for some constant t_0.

4. Starting from the metric in the Bondi gauge (40.30), one can go the Schwarzschild gauge (40.32) by introducing $w = w(t,r)$ through

$$\mathrm{e}^{h(w,r)} dw = \left(\mathrm{e}^{h_s(t,r)} dt + \epsilon f_s(t,r)^{-1} dr \right) \tag{40.36}$$

where $h(w,r)$ is the required integrating factor. Then (40.30) takes the form (40.32) with $f_s(t,r) = f(w,r)$, or more explicitly

$$f_s(t,r) = 1 - \frac{2m(w(t,r),r)}{r} = f(w(t,r),r) \quad \Leftrightarrow \quad m_s(t,r) = m(w(t,r),r) \ . \tag{40.37}$$

We will look at this in somewhat more detail in section 40.5.

If $h(w,r) = 0$ (or depends only on w) and f depends only on r, one can choose $h_s(t,r) = 0$ and (40.36) reduces to the standard relation

$$dw = dt + \epsilon f(r)^{-1} dr = dt + \epsilon dr^* \ . \tag{40.38}$$

between advanded or retarded and tortoise coordinates of a static black hole metric.

5. As in our discussion of the Einstein equations in the (Birkhoff-)Schwarzschild gauge in section 24.6 and of the interpretation of the mass function for Vaidya metrics in section 40.1, the mass function $m(w,r)$ is the invariantly and geometrically defined Hernandez-Misner or Misner-Sharp mass, for general spherically symmetric space-times of the form (24.80),

$$ds^2 = g_{ab}(z) dz^a dz^b + r(z)^2 d\Omega^2 \ , \tag{40.39}$$

given by the gradient-squared of the radius function $r(z)$,

$$m(z) = \frac{r(z)}{2} \left(1 - g^{ab}(z) \partial_a r(z) \partial_b r(z) \right) \ . \tag{40.40}$$

The observation above that in the transformation from the Bondi to the Schwarzschild gauge one has $m_s(t,r) = m(w,r)$ (40.37) is a particular manifestation of the fact that $m(z)$ is a scalar under coordinate transformations preserving the form (40.39) of the metric.

As we saw in section 24.6, solving the vacuum Einstein equations in the standard spherically symmetric gauge (40.32), one recovers Birkhoff's theorem and the Schwarzschild metric in the standard Schwarzschild coordinates. However, at least with the benenfit of hindsight, it is clear that a better gauge may be one which is adapted to radial lightrays rather than to static observers. Such an ansatz is provided by the general spherically symmetric metric written in the Bondi gauge (40.30).

For both signs, the independent (w, r)-components of the Einstein tensor take the simple form (the counterpart of (24.71))

$$\begin{aligned} G^w_{\ w} &= -\frac{2\partial_r m(w,r)}{r^2} \\ G^r_{\ w} &= +\frac{2\partial_w m(w,r)}{r^2} \\ G_{rr} &= +\frac{2\partial_r h(w,r)}{r} \quad . \end{aligned} \tag{40.41}$$

From these one can then also deduce the "missing" components, such as

$$\begin{aligned} G^w_{\ r} &= g^{w\alpha}G_{\alpha r} = g^{wr}G_{rr} = \epsilon \mathrm{e}^{-h}G_{rr} \\ G^r_{\ r} &= g^{r\alpha}G_{\alpha r} = g^{rr}G_{rr} + G^w_{\ w} = fG_{rr} + G^w_{\ w} \\ G_{ww} &= g_{w\alpha}G^\alpha_{\ w} = \epsilon \mathrm{e}^{h}G^r_{\ w} - f\mathrm{e}^{2h}G^w_{\ w} \quad . \end{aligned} \tag{40.42}$$

One sees that in the chosen (Bondi gauge) parametrisation, the components in (40.41) are the simplest complete set of independent components and building blocks of the Einstein tensor, and therefore a particularly convenient starting point for analysing and solving the Einstein equations. The angular components $G^\theta_{\ \theta} = G^\phi_{\ \phi}$ (that they are equal and that $G^\theta_{\ \phi} = 0$ is implied by spherical symmetry) are more involved, but are often not needed in practice, as they can be substituted by the Bianchi identities.

We now (re-)derive two simple but important implications of these formulae:

1. Birkhoff's Theorem (section 24.6)

   The vacuum Einstein equations now imply that $m(w, r) = m$ is a constant and that $h(w, r) = h(w)$ is independent of r. Thus the metric takes the form

   $$\begin{aligned} ds^2 &= -\left(1 - \frac{2m}{r}\right)\mathrm{e}^{2h(w)}dw^2 + 2\epsilon\, \mathrm{e}^{h(w)}dw\, dr + r^2 d\Omega^2 \\ &= -\left(1 - \frac{2m}{r}\right)d\tilde{w}^2 + 2\epsilon\, d\tilde{w}\, dr + r^2 d\Omega^2 \end{aligned} \tag{40.43}$$

   where the new coordinate $\tilde{w}$ labelling radial in or out geodesics is defined by the simple change of variables

   $$d\tilde{w} = \mathrm{e}^{h(w)}dw \quad , \tag{40.44}$$

   reflecting the gauge invariance (40.35). Setting $\tilde{w} = v$ for $\epsilon = +1$ and $\tilde{w} = u$ for $\epsilon = -1$, the result is therefore precisely the Schwarzschild metric in ingoing or outgoing Eddington-Finkelstein coordinates, and therefore ab initio regular at the future or past event-horizon.

2. Characterisation of Vaidya Metrics

   Starting from the general form (40.30) of the metric and its Einstein tensor (40.41), one can now deduce that the most general metric describing a purely ingoing ($\epsilon = +1$) or outgoing ($\epsilon = -1$) matter content with T_{ww} the only non-vanishing component of the Einstein tensor, is the ingoing or outgoing Vaidya metric. Indeed, from this assumption one deduces

$$G_{\alpha\beta} \sim \delta^w_\alpha \delta^w_\beta \quad \Rightarrow \quad \begin{array}{lll} G_{rr} = 0 & \Rightarrow & \partial_r h(w,r) = 0 \\ G^w_{\ w} = 0 & \Rightarrow & \partial_r m(w,r) = 0 \end{array} \tag{40.45}$$

   so that the metric can be put into the standard Vaidya form

$$ds^2 = -\left(1 - \frac{2m(\tilde{w})}{r}\right) d\tilde{w}^2 + 2\epsilon d\tilde{w} dr + r^2 d\Omega^2 \tag{40.46}$$

   by the same redefinition (40.44) of the null coordinate as in the vacuum case.

40.4 Description of In- and Outgoing Pure Radiation Fields

While we have seen above that Vaidya metrics are characterised by the fact that they have an energy-momentum tensor $T_{\alpha\beta} \sim \delta^w_\alpha \delta^w_\beta$, it is useful to rephrase this somewhat, and we will do this in an elementary fashion in this section. This can also be formulated in a geometrically somewhat more satisfactory (because less coordinate-dependent) form, in terms of geodesic congruences, and we will study these (for these and other reasons) in some detail later on.

In general, among the standard physically relevant types of energy-momentum tensors, in addition to the well-known cases of a Maxwell (electromagnetic) field,

$$T_{\alpha\beta} = -F_{\alpha\gamma} F_\beta^{\ \gamma} + \tfrac{1}{4} g_{\alpha\beta} F_{\gamma\delta} F^{\gamma\delta} \tag{40.47}$$

and a perfect fluid,

$$T_{\alpha\beta} = (\rho + p) u_\alpha u_\beta + p g_{\alpha\beta} \quad , \quad u_\alpha u^\alpha = -1 \ , \tag{40.48}$$

one also has what is known as a *pure radiation field* or *null dust*, characterised by an energy-momentum tensor of the form

$$T_{\alpha\beta} = \rho k_\alpha k_\beta \quad , \quad k_\alpha k^\alpha = 0 \ . \tag{40.49}$$

<u>Remarks:</u>

1. More generally, such an energy-momentum tensor could contain a sum over different species of massless particles moving at the speed of light,

$$T_{\alpha\beta} = \sum_a \rho^{(a)} k^{(a)}_\alpha k^{(a)}_\beta \quad , \quad k^{(a)}_\alpha k^{(a)\alpha} = 0 \ . \tag{40.50}$$

In particular, such an energy-momentum tensor is traceless,

$$T^\alpha_{\ \alpha} = \sum_a \rho^{(a)} k^{(a)\alpha} k^{(a)}_\alpha = 0 \quad . \tag{40.51}$$

2. Such an energy-momentum tensor can arise e.g. from null Maxwell fields or from massless scalar fields (in some geometric optics / eikonal approximation). For example a spherical outgoing scalar wave of the form $\phi(u,r) = \psi(u)/r$ with $u = t - r$ in Minkowski space gives rise to such an energy-momentum tensor when terms like $\psi'(u)/r$ (u-derivatives) dominate over terms like $\psi(u)/r^2$ (r-derivatives), leading to

$$T_{\alpha\beta} \approx \frac{\psi'(u)^2}{r^2}\delta_\alpha^{\ u}\delta_\beta^{\ u} \quad . \tag{40.52}$$

3. The absolute normalisation of the null energy density is not determined by this form of the energy-momentum tensor since it scales under a space-time dependent boost of k,

$$k \to \mathrm{e}^{-\alpha(x)}k \quad \Rightarrow \quad \rho(x) \to \mathrm{e}^{2\alpha(x)}\rho(x) \quad . \tag{40.53}$$

We now consider again a general spherically symmetric space-time and denote the tangent vectors to an ingoing (respectively outgoing) congruence of (not necessarily affinely parametrised) radial future oriented null geodesics by

$$\text{ingoing:} \quad n = n^\alpha\partial_\alpha \quad , \quad \text{outgoing:} \quad \ell = \ell^\alpha\partial_\alpha \quad , \tag{40.54}$$

cross-normalised to $n.\ell = g_{\alpha\beta}n^\alpha\ell^\beta = -1$ (for more on the properties of such vector fields in general see section 12.6).

Then an in- respectively outgoing pure radiation field is described by an energy-momentum tensor of the form

$$T_{\alpha\beta} = \rho_{in}n_\alpha n_\beta \quad \text{or} \quad T_{\alpha\beta} = \rho_{out}\ell_\alpha\ell_\beta \quad . \tag{40.55}$$

In the Bondi gauge (40.30), a natural (albeit asymmetric) choice for ℓ and n is

$$\begin{aligned} \epsilon = +1: &\quad n = -\partial_r \quad , \quad \ell = \mathrm{e}^{-h}\partial_v + \tfrac{1}{2}f\partial_r \\ \epsilon = -1: &\quad \ell = +\partial_r \quad , \quad n = \mathrm{e}^{-h}\partial_u - \tfrac{1}{2}f\partial_r \quad . \end{aligned} \tag{40.56}$$

<u>EXAMPLES:</u>

1. A purely ingoing radiation field has the energy-momentum tensor

$$T_{\alpha\beta} = \rho_{in}n_\alpha n_\beta \quad . \tag{40.57}$$

The general solution of the Einstein equations is given by the ingoing Vaidya metric (40.2). For the ingoing Vaidya metric one can choose

$$n = -\partial_r \quad , \quad \ell = \partial_v + \tfrac{1}{2}f(v,r)\partial_r \quad , \tag{40.58}$$

with covariant components

$$n_\alpha = -\delta_\alpha^{\ v} \quad , \quad \ell_v = -\tfrac{1}{2} f \ , \ \ell_r = 1 \ . \tag{40.59}$$

In these adapted coordinates (adapted to ingoing null geodesics) one has (suppressing the two angular dimensions)

$$n_\alpha n_\beta = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \ . \tag{40.60}$$

Therefore the energy-momentum tensor (40.9) of the ingoing Vaidya metric indeed has the purely ingoing form

$$T_{\alpha\beta} = \frac{m'(v)}{4\pi G_N r^2} n_\alpha n_\beta \ . \tag{40.61}$$

2. We can also easily deduce this specific form of the energy-momentum tensor from the geometry. Writing a general ansatz for the energy momentum tensor of the ingoing Vaidya metric as

$$T_{\alpha\beta} = \rho(v,r) n_\alpha n_\beta \ , \tag{40.62}$$

the covariant divergence of the energy-momentum tensor is, using the general formulae

$$n^\alpha \nabla_\alpha n^\beta = \kappa_n n^\beta \quad , \quad \nabla^\alpha n_\alpha = \kappa_n + \theta_n \tag{40.63}$$

from section 12.5 for the inaffinity and expansion of a null congruence,

$$\nabla^\alpha T_{\alpha\beta} = (n^\alpha \nabla_\alpha \rho + \rho \nabla^\alpha n_\alpha) n_\beta + \rho n^\alpha \nabla_\alpha n_\beta = (n^\alpha \nabla_\alpha \rho + (\theta_n + 2\kappa_n)\rho)\, n_\beta \ . \tag{40.64}$$

In the case at hand, with $n = -\partial_r$, we have $\kappa_n = 0$ (r is an affine parameter along ingoing radial null geodesics for the ingoing Vaidya metric), and in section 32.8 we determined the ingoing expansion (contraction) to be $\theta_n = -2/r$ (exactly as in Minkowski space).

Conservation of the energy-momentum tensor thus requires

$$\nabla^\alpha T_{\alpha\beta} = 0 \quad \Leftrightarrow \quad \partial_r \rho(v,r) + 2\rho(v,r)/r = 0 \quad \Leftrightarrow \quad \rho(v,r) = \frac{\rho(v)}{r^2} \ , \tag{40.65}$$

which is precisely the general form of the Vaidya energy-momentum tensor.

3. Likewise, a purely outgoing radiation field has the energy-momentum tensor

$$T_{\alpha\beta} = \rho_{out} \ell_\alpha \ell_\beta \ . \tag{40.66}$$

The general solution of the Einstein equations is given by the outgoing Vaidya metric (40.12). For the outgoing Vaidya metric one can choose

$$\ell = +\partial_r \quad , \quad n = \partial_u - \tfrac{1}{2} f(u,r) \partial_r \ , \tag{40.67}$$

with covariant components

$$\ell_\alpha = -\delta_\alpha^{\,u} \quad , \quad n_u = -\tfrac{1}{2} f \ \ , \ \ n_r = -1 \ \ . \tag{40.68}$$

In particular one has, as in the ingoing case,

$$\ell_\alpha \ell_\beta = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{40.69}$$

and the energy-momentum tensor (40.14) of the outgoing Vaidya metric indeed has the purely outgoing form

$$T_{\alpha\beta} = -\frac{m'(u)}{4\pi G_N r^2} \ell_\alpha \ell_\beta \ \ . \tag{40.70}$$

This makes it even more manifest that the ingoing Vaidya metric describes purely ingoing null matter (and the outgoing Vaidya metric purely outgoing null matter). Likewise, the generalised ingoing Vaidya metrics (40.16) cannot describe purely outgoing matter (and vice-versa), as can be seen from the Einstein tensor (40.17): outgoing matter should have an energy-momentum tensor proportional to $\ell_\alpha \ell_\beta$ which, in ingoing coordinates, has the form

$$\ell_\alpha \ell_\beta = \begin{pmatrix} \frac{1}{4} f^2 & -\frac{1}{2} f \\ -\frac{1}{2} f & 1 \end{pmatrix} \tag{40.71}$$

(the expressions for n and ℓ from (40.58) are still valid in this case, since the conditions $n^2 = \ell^2 = 0$ and $n.\ell = -1$ are purely algebraic and do not depend on whether or not the mass function depends on r). In particular, therefore, for outgoing pure radiation fields in ingoing coordinates one necessarily has $G_{rr} \neq 0$. The generalised ingoing Vaidya metric, on the other hand, has $G_{rr} = 0$ (and (40.41) shows that $G_{rr} \neq 0$ requires a non-trivial, i.e. r-dependent, $h(w,r)$).

40.5 Vaidya Metrics in the Schwarzschild Gauge

We have just seen that what characterises the Vaidya metric is a matter content consisting of purely in- or outgoing radiation, and that in the Bondi gauge this corresponds to an energy-momentum tensor of the form

$$T_{\alpha\beta} = \rho(w,r) k_\alpha k_\beta \quad , \quad \rho(w,r) = \frac{\epsilon m'(w)}{4\pi G_N r^2} \tag{40.72}$$

where $k_\alpha = -\delta_\alpha^{w}$. It is also of interest to understand what characterises these metrics in the usual Schwarzschild-Birkhoff gauge (40.32) and how to write the Vaidya metrics in this gauge. In principle, the answer to this question is provided by the coordinate transformation (40.36),

$$dw(t,r) = \mathrm{e}^{-h(w(t,r),r)} \left(\mathrm{e}^{h_s(t,r)} dt + \epsilon f_s(t,r)^{-1} dr \right) \tag{40.73}$$

between the Bondi and Schwarzschild gauges. If the metric in the Bondi gauge is of the Vaidya form, one has $h(w,r)=0$ (as well as $m(w,r)=m(w)$), and (40.73) reduces to

$$dw(t,r) = \left(\mathrm{e}^{h_s(t,r)}dt + \epsilon f_s(t,r)^{-1}dr\right) \tag{40.74}$$

where

$$f_s(t,r) = 1 - \frac{2m_s(t,r)}{r} = 1 - \frac{2m(w(t,r))}{r} \quad . \tag{40.75}$$

Here we have used the fact, noted in section 40.3, that the mass function transforms as a scalar under this coordinate transformation (40.37), and we will for notational simplicity set $m_s = m$ in the following.

Thus in practice one needs to solve the equations

$$\frac{\partial w(t,r)}{\partial r} = \epsilon\left(1 - \frac{2m(w(t,r))}{r}\right)^{-1} \tag{40.76}$$

and

$$\frac{\partial w(t,r)}{\partial t} = \mathrm{e}^{h_s(t,r)} \quad , \tag{40.77}$$

with $h_s(t,r)$ related to $f_s(t,r)$ by the integrability condition

$$\partial_t\partial_r w(t,r) = \partial_r\partial_t w(t,r) \quad , \tag{40.78}$$

This is an arduous task in general, even for simple Vaidya metrics. A quick way to determine the general form of the Vaidya metric in the Schwarzschild gauge, without having to explicitly solve these equations in order to determine the coordinate transformation $w = w(t,r)$, is to write

$$dm(w) = m'(w)dw = m'(w)\mathrm{e}^{h_s}dt + m'(w)\epsilon f_s^{-1}dr \quad , \tag{40.79}$$

so that

$$\partial_t m(t,r) = m'(w)\mathrm{e}^{h_s} \quad , \quad \partial_r m(t,r) = \epsilon m'(w) f_s^{-1} \quad . \tag{40.80}$$

In particular, this implies that

$$\frac{\partial_t m(t,r)}{\partial_r m(t,r)} = \epsilon f_s \mathrm{e}^{h_s} \quad , \tag{40.81}$$

which gives the desired relation between $h_s(t,r)$ and $f_s(t,r)$ or $m(t,r)$. In particular, the general Vaidya metric can now be written as

$$\begin{aligned} ds^2 &= -\mathrm{e}^{2h_s(t,r)}f_s(t,r)dt^2 + f_s(t,r)^{-1}dr^2 + r^2d\Omega^2 \\ &= \left(1 - \frac{2m(t,r)}{r}\right)^{-1}\left[-\left(\frac{\partial_t m(t,r)}{\partial_r m(t,r)}\right)^2 dt^2 + dr^2\right] + r^2d\Omega^2 \quad . \end{aligned} \tag{40.82}$$

This is the general form of the Vaidya metric in the Schwarzschild gauge.

While we have performed this coordinate transformation at the level of the metric, we can of course also do this at the level of the field equations and the energy-momentum tensor. This is instructive in its own right, as it will tell us what is the form of the energy-momentum tensor in the Schwarzschild gauge which will simply give rise to a Vaidya metric in (Schwarzschild-) disguise upon solving the Einstein equations.

On the one hand, in the gauge (40.32) the (t,r)-components of the Einstein tensor have the simple form (24.71)

$$\begin{aligned}\partial_r m(t,r) &= 4\pi G_N\, r^2(-T^t_{\ t})\\ \partial_t m(t,r) &= 4\pi G_N\, r^2(+T^r_{\ t})\\ \partial_r h_s(t,r) &= 4\pi G_N\, r f_s(t,r)^{-1}(-T^t_{\ t}+T^r_{\ r})\ .\end{aligned} \tag{40.83}$$

On the other hand transforming

$$T_{\alpha\beta} = \rho\delta^w_\alpha\delta^w_\beta \tag{40.84}$$

to the Schwarzschild gauge, using (40.74), one finds that the non-vanishing components are

$$T_{tt} = \mathrm{e}^{2h_s}\rho \quad , \quad T_{tr} = \epsilon\mathrm{e}^{h_s}f_s^{-1}\rho \quad , \quad T_{rr} = f_s^{-2}\rho \tag{40.85}$$

or

$$T^t_{\ t} = -f_s^{-1}\rho \quad , \quad T^t_{\ r} = -\epsilon\mathrm{e}^{-h_s}f_s^{-2}\rho \quad , \quad T^r_{\ t} = \epsilon\mathrm{e}^{h_s}\rho \quad , \quad T^r_{\ r} = f_s^{-1}\rho\ . \tag{40.86}$$

Inserting this into the Einstein equations (40.83) one (re)discovers the relations (40.80).

In particular, from (40.86) one has

$$T^\alpha_{\ \alpha} = T^t_{\ t} + T^r_{\ r} = 0 \quad \Leftrightarrow \quad T_{tt} = \mathrm{e}^{2h_s}f_s^2 T_{rr}\ , \tag{40.87}$$

which reflects the fact that the original Vaidya energy-momentum tensor was traceless. Equivalently, using terminology that is adapted to the coordinates (t,r) of the Schwarzschild gauge, we can rephrase this as the statement that the energy density and radial pressure are equal,

$$T^r_{\ r} + T^t_{\ t} = 0 \quad \Leftrightarrow \quad \rho = P_r\ . \tag{40.88}$$

Moreover, one has the relation

$$T^t_{\ t} = \epsilon\mathrm{e}^{h_s}f_s T^t_{\ r}\ , \tag{40.89}$$

which expresses the lightlike nature of the energy-momentum content. Indeed, in the Schwarzschild gauge an ingoing (for $\epsilon = +1$) respectively outgoing (for $\epsilon = -1$) radial null vector k^α is characterised by

$$k^\alpha k_\alpha = 0 \quad \Rightarrow \quad k^r = -\epsilon f_s\mathrm{e}^{h_s}k^t\ , \tag{40.90}$$

and the relation (40.89) can be rephrased as the statement that

$$T^t_{\ t} = \epsilon\mathrm{e}^{h_s}f_s T^t_{\ r} \quad \Leftrightarrow \quad T^t_{\ \alpha}k^\alpha = 0\ . \tag{40.91}$$

Together

- the tracelessness condition (40.88)
- the null condition (40.91)
- and the fact that the energy-momentum tensor is spherically symmetric and purely longitudinal (i.e. that its transverse angular components are zero)

characterise the (evidently highly constrained and idealised) Vaidya energy-momentum content in the Schwarzschild gauge.

41 Vaidya Metrics II: Radial Null and Timelike Geodesics

41.1 Radial Null Geodesics for Ingoing Vaidya

In order to improve our understanding of the physics of the Vaidya metrics, and to explore their causal structure, in this section we now look at radial and timelike null geodesics and their properties.

Null geodesics are characterised first of all by the null-condition

$$-f(v,r)dv^2 + 2dv\;dr = 0 \quad . \tag{41.1}$$

This needs to be supplemented by the null geodesic equations

$$\begin{aligned} \ddot{v} + \frac{m(v)}{r^2}\dot{v}^2 \quad &= 0 \\ \ddot{r} + \frac{m'(v)}{r}\dot{v}^2 \quad &= 0 \end{aligned} \tag{41.2}$$

(the null condition (41.1) has been used to put the r-equation into this simple form) or by appropriate first integrals of these equations arising from conserved charges. These are available only for special choices of $m(v)$. There are two cases (I am aware of), namely

1. a constant mass function $m(v) = m_0$ with its time-translation Killing vector $\partial_v = \partial_t$ (the Schwarzschild metric), and
2. a linear mass function $m(v) = \mu v$ with a scaling / dilatation symmetry.

We will discuss these separately below.

Continuing for now with the general (but not generalised) Vaidya metric, ingoing radial null geodesics satisfy $dv = 0$ or $\dot{v} = 0$, and thus from (41.2)

$$v(\tau) = v_0 \quad , \quad r(\tau) = r_0 + \dot{r}_0\,\tau \quad . \tag{41.3}$$

The tangent vector $\dot{x}^\alpha = dx^\alpha/d\tau$ is future-oriented, i.e. such that its scalar product with $\partial_t = \partial_v$ is negative, for $\dot{r}_0 < 0$,

$$g_{\alpha\beta}\dot{x}^\alpha(\partial_t)^\beta = g_{\alpha v}\dot{x}^\alpha = \dot{r} = \dot{r}_0 \stackrel{!}{<} 0 \quad , \tag{41.4}$$

so that indeed the radius decreases along future-oriented ingoing null geodesics. Moreover, since r is affinely related to τ, r is an affine parameter along these ingoing null geodesics. In particular, this means that the ingoing null vector field $n = -\partial_r$ introduced in (40.58) is affinely parametrised, i.e. satisfies

$$n^\alpha\nabla_\alpha n^\beta = 0 \quad . \tag{41.5}$$

One can also see this directly from the calculation of the Christoffel symbols,

$$n^\alpha \nabla_\alpha n^\beta = \Gamma^\beta_{rr} = 0 \tag{41.6}$$

(since the only non-trivial metric component $g_{r\beta}$ is the constant $g_{rv} = 1$).

By contrast, the outgoing null vector field ℓ of (40.58), chosen to satisfy $\ell^\alpha n_\alpha = -1$, is then not affinely parametrised. Indeed, using

$$\begin{aligned} & (\ell^\alpha \nabla_\alpha \ell^\beta)\ell_\beta = \tfrac{1}{2}\ell^\alpha \nabla_\alpha (\ell^\beta \ell_\beta) = 0 \\ \Rightarrow\quad & \ell^\alpha \nabla_\alpha \ell^r = \tfrac{1}{2} f \ell^\alpha \nabla_\alpha \ell^v \\ \Rightarrow\quad & \ell^\alpha \nabla_\alpha \ell^\beta = (\ell^\alpha \nabla_\alpha \ell^v)\ell^\beta \end{aligned} \tag{41.7}$$

and noting (e.g. from the v-equation in (41.2)) that the only non-trivial Christoffel symbol $\Gamma^v_{\alpha\beta}$ is $\Gamma^v_{vv} = m(v)/r^2$, one deduces the inaffinity of ℓ^α,

$$\ell^\alpha \nabla_\alpha \ell^\beta = \frac{m(v)}{r^2}\ell^\beta \equiv \kappa_\ell \ell^\beta \quad . \tag{41.8}$$

Affinely parametrised outgoing radial null geodesics, on the other hand, satisfy (41.2) and

$$-f(v,r)dv + 2dr = 0 \quad \Leftrightarrow \quad -f(v,r)\dot{v} + 2\dot{r} = 0 \quad \Leftrightarrow \quad 2\frac{dr}{dv} = f(v,r) \quad . \tag{41.9}$$

Remarkably, with the help of the null condition (41.9) the non-linear second order geodesic equations can be integrated to first-order differential equations.[187] As the derivation is not given in that reference, we provide it here. First of all, we write

$$m'(v)\dot{v} = \dot{m} \tag{41.10}$$

and use (41.9) to eliminate the remaining $\dot{v}$ from the r-equation in (41.2). Then one finds

$$\begin{aligned} & 0 = \ddot{r} + \dot{r}\frac{2\dot{m}}{r-2m} = \ddot{r} - \dot{r}\frac{d}{d\tau}\log(r-2m) + \dot{r}^2\frac{1}{r-2m} \\ \Leftrightarrow\quad & \frac{d}{d\tau}\log(\dot{r}/(r-2m)) = -\dot{r}/(r-2m) \\ \Leftrightarrow\quad & \dot{r}/(r-2m) = 1/(\tau-\tau_0) \end{aligned} \tag{41.11}$$

or

$$\frac{dr}{d\tau} = \frac{r(\tau) - 2m(v(\tau))}{\tau - \tau_0} \quad . \tag{41.12}$$

From (41.9) one then also deduces

$$\frac{dv}{d\tau} = \frac{2r(\tau)}{\tau-\tau_0} \quad , \tag{41.13}$$

[187] R. Waugh, K. Lake, *Backscattering radiation in the Vaidya metric near zero mass*, Phys. Lett. B116 (1986) 154-156.

so that these are future-directed curves (increasing v) for $\tau > \tau_0$. Equations (41.12) and (41.13) imply the outgoing null condition (41.9) and govern the behaviour of outgoing lightrays in the Vaidya metric. Without loss of generality we can set $\tau_0 = 0$.

These 1st order equations allow us to rewrite the null geodesic equations (41.2) for outgoing null geodesics in a way that will turn out to be useful later on: the geodesic equation for r in (41.2) together with (41.13) leads to

$$\ddot{r} + \frac{4m'(v)}{\tau^2} r = 0 \ ; \tag{41.14}$$

likewise, the geodesic equation for v in (41.2) together with (41.13) leads to

$$\ddot{v} + \frac{4m(v)}{\tau^2} = 0 \ . \tag{41.15}$$

These equations become particularly tractable (namely linear) in the case of a linear mass function $m(v) = \mu v$ that we will study in detail later.

As a warm-up exercise, a first application of the above results, and for comparison (and, later on, matching) purposes, we first rederive the equations for outgoing lightrays and for the horizon generators in the constant mass Schwarzschild case $m(v) = m_0$ in ingoing Eddington-Finkelstein coordinates before addressing the same problem in the dynamical Vaidya context.

For the Schwarzschild metric one has $m'(v) = 0$ and (41.2) implies that $\ddot{r} = 0$ so that

$$\dot{r} = E \tag{41.16}$$

is conserved. This constant of integration can equivalently be regarded as the conserved energy associated to the time-translation invariance of the Schwarzschild metric generated by $\partial_t = \partial_v$. Indeed, the corresponding conserved charge is

$$E = -g_{\alpha\beta}(\partial_v)^\alpha \dot{x}^\beta = f(r)\dot{v} - \dot{r} \overset{(41.9)}{=} \dot{r} \ . \tag{41.17}$$

Then (41.12) already gives the solution for $r = r(\tau)$, namely

$$r(\tau) = 2m_0 + E(\tau - \tau_0) \ , \tag{41.18}$$

(so that r is related to τ by an affine transformation, i.e. r is an affine parameter also along outgoing geodesics, unless $E = 0$), and (41.13) then gives

$$v(\tau) = 4m_0 \log(\tau - \tau_0) + 2E\tau + c \ . \tag{41.19}$$

For $\tau \to \tau_0$, these lightrays emerge from the past horizon $r = 2m_0$ and $v \to -\infty$. This illustrates the past geodesic incompleteness of the ingoing Eddington-Finkelstein coordinates.

A special case is the geodesic (or S^2-family of geodesics) for $E = 0$, for which

$$r(\tau) = 2m_0 \equiv r_0 \quad , \quad v(\tau) = 2r_0 \log\tau + v_1 \ , \tag{41.20}$$

with τ determined up to affine transformations $\tau \to a\tau + b$ and $v_1 = v(\tau = 1)$. These null geodesics lie on and generate the future event horizon of the Schwarzschild black hole.

<u>REMARKS:</u>

1. The explicit solution (41.20) shows that the Kruskal coordinate
$$V = \mathrm{e}^{v/4m_0} = \tau \tag{41.21}$$
is an affine parameter along the horizon.

2. Note also, for comparison with the Vaidya metric, that (41.9) implies that the 2nd derivative d^2r/dv^2 is
$$\frac{d^2r}{dv^2} = \frac{m_0}{2r^2} f(r) \quad . \tag{41.22}$$
For $r > 2m_0$ one has $f(r) > 0$. Thus $r(v)$ is convex and an initially outgoing lightray, $dr/dv > 0$, will remain outgoing at all times. For $r < 2m_0$, on the other hand, one has $f(r) < 0$ and therefore $dr/dv < 0$ and $d^2r/dv^2 < 0$. Thus $r(v)$ is concave, moving towards smaller values of r, and will ultimately reach $r = 0$ at a finite value of v.

41.2 Radial Null Geodesics for Outgoing Vaidya

One can go through the same exercises for the outgoing Vaidya metric (40.12)
$$ds^2 = -f(u,r)du^2 - 2dudr + r^2 d\Omega^2 \quad , \quad f(u,r) = 1 - \frac{2m(u)}{r} \quad . \tag{41.23}$$
Here we now assume that the mass function $m(u)$ is non-negative ($m(u) \geq 0$) and non-increasing ($m'(u) \leq 0$).

Outgoing null geodesics are given by $u = const.$, and r is an affine parameter along these outgoing null geodesics.

Ingoing (decreasing r) null geodesics are determined by the null-condition
$$f(u,r)\dot{u} + 2\dot{r} = 0 \tag{41.24}$$
and the null geodesic equations
$$\begin{aligned} \ddot{r} - \frac{m'(u)}{r}\dot{u}^2 &= 0 \\ \ddot{u} - \frac{m(u)}{r^2}\dot{u}^2 &= 0 \end{aligned} \tag{41.25}$$
(note the sign flip relative to (41.2) in both equations). Again, these can be integrated to first order equations, and the counterpart of (41.12) and (41.13) is
$$\dot{r} = \frac{r - 2m(u)}{\tau} \quad , \quad \dot{u} = -\frac{2r}{\tau} \tag{41.26}$$

(a sign flip only in the 2nd equation, and we have set an integration constant τ_0 to zero). We have $\dot{r} < 0$ and $\dot{u} > 0$ for $r > 2m$ and $\tau < 0$, the restriction on the range of τ already suggesting a potential future geodesic incompleteness of this coordinate system. We will come back to this issue below.

With the help of (41.26), the equations (41.25) can be put into the form

$$\ddot{r} - \frac{4m'(u)}{\tau^2} r = 0 \quad , \quad \ddot{u} - \frac{4m(u)}{\tau^2} = 0 \ , \tag{41.27}$$

the counterparts of (41.14) and (41.15), and again these will become linear decoupled harmonic oscillator equations in the case of a linear (and now decreasing) mass function $m(u)$.

41.3 Gravitational Redshift for Outgoing Vaidya

In order to improve our understanding of the Vaidya geometry, we will now relate the data given by the geometry and matter content (metric and energy-momentum tensor) to those measured by an observer. We will concentrate on the outgoing (radiating) Vaidya metric, but of course the ingoing case can be discussed in complete analogy.[188]

We denote by $(x^\alpha)^\bullet$ the timelike proper-time normalised 4-velocity of a radially moving observer, with τ_o the observer's proper time,

$$(x^\alpha)^\bullet = \frac{dx^\alpha(\tau_o)}{d\tau_o} = (u^\bullet, r^\bullet, 0, 0) \ . \tag{41.28}$$

so that

$$g_{\alpha\beta}(x^\alpha)^\bullet (x^\beta)^\bullet = -1 \quad \Leftrightarrow \quad 2u^\bullet r^\bullet + f(u,r)(u^\bullet)^2 = +1 \ . \tag{41.29}$$

This implies, in particular, that $u^\bullet \neq 0$, and we will take future-oriented paths to mean $u^\bullet > 0$. We will also mostly be interested in observers that travel (or at least start off) in the region $r > 2m(u)$ outside the (unphysical) past apparent horizon.

Introducing the quantity

$$E_u = f u^\bullet + r^\bullet \tag{41.30}$$

(the rationale for this notation will be explained in section 41.4), one can solve (41.29) for $u^\bullet$,

$$u^\bullet = (E_u + r^\bullet)^{-1} \tag{41.31}$$

and (41.29) can be written as

$$(r^\bullet)^2 + f(u,r) = (E_u)^2 \ . \tag{41.32}$$

For $f(u,r) = 1 - 2m(u)/r$ with $m(u)$ bounded one has an asymptotically flat metric in the (crude) sense that $f \to 1$ for $r \to \infty$. In that case, it follows from the Vaidya

[188] This discussion is adapted (and expanded) from R. Lindquist, R. Schwartz, C. Misner, *Vaidya's Radiating Schwarzschild Metric*, Phys. Rev. 137 (1965) 5B, B1364-B1368.

line element that the proper time for an observer at rest at infinity is simply $d\tau_\infty = du$. Thus (41.31) can be interpreted as the relation between the observer's proper time and the proper time at infinity,

$$d\tau_\infty = (E_u + r^\bullet)^{-1} d\tau_o \quad . \tag{41.33}$$

As usual, this formula will be related to that for gravitational redshift to be discussed below.

We will now apply these formulae to

1. determine the gravitational redshift of outgoing lightrays in the outgoing Vaidya geometry
2. determine the energy density of the background energy-momentum tensor as seen by the observer,

We will also use them in section 41.4 to take a more detailed look at, and interpret, the equations for timelike geodesics, and in section 41.6 to locate and detect potential surfaces of infinite redshift and discuss the issues that arise in relation to them.

1. The null wave vector k^α of outgoing lightrays, in particular of the outgoing radiation due to the energy-momentum tensor, will be proportional to the affinely parametrised outgoing null vector $\ell^\alpha = (\partial_r)^\alpha$ (40.67). The frequency ω_o of this lightray as determined by the timelike observer will essentially be given by the projection of the wave vector k^α onto the observer's rest-frame, namely

   $$\omega_o = -(x^\alpha)^\bullet k_\alpha \sim -(x^\alpha)^\bullet \ell_\alpha = u^\bullet = (E_u + r^\bullet)^{-1} \quad . \tag{41.34}$$

   We can also equivalently phrase this in terms of the observer emitting outgoing lightrays. Since outgoing lightrays travel along lines of constant u, signals with initial separation Δu are received at infinity with the same separation Δu. The different perceived frequencies are due to the differences in proper time, and this explains the equivalence between (41.33) and (41.34).

2. It follows from (40.66) and (40.70) that the (null) energy density of the Vaidya metric is

   $$\rho_{out} = -\frac{m'(u)}{4\pi G_N r^2} \quad , \tag{41.35}$$

   so that (as noted before, in section 40.2) the total flux through the sphere of radius r is independent of r and given by

   $$\mathcal{F} = 4\pi r^2 \rho_{out} = -m'(u)/G_N \quad , \tag{41.36}$$

   and likewise

   $$\mathcal{F} = m'(v)/G_N \tag{41.37}$$

for the ingoing Vaidya metric. Here the convention has been chosen that the flux $\mathcal{F} \sim \epsilon m'(w)$ is positive for both ingoing and outgoing radiation.

However, an observer will not necessarily detect this static energy density and flux. The same reasoning as above for the redshift leads to the conclusion that the energy-density of the outgoing radiation provided by the background energy-momentum tensor measured by this observer in his rest-frame is

$$\rho_o = T_{\alpha\beta}(x^\alpha)^\bullet (x^\beta)^\bullet = \rho_{out}(\ell_\alpha (x^\alpha)^\bullet)^2 = \rho_{out}(u^\bullet)^2 \tag{41.38}$$

or

$$\rho_o = -\frac{m'(u)}{4\pi G_N r^2}(u^\bullet)^2 \quad . \tag{41.39}$$

Thus ρ_o can be written in terms of the observer's radial velocity $r^\bullet$ as

$$\rho_o = -\frac{m'(u)}{4\pi G_N r^2}(E_u + r^\bullet)^{-2} \quad . \tag{41.40}$$

For an observer at rest at infinity one thus finds the total flux or luminosity

$$\mathcal{F}_\infty = \lim_{r^\bullet \to 0} \lim_{r \to \infty} 4\pi r^2 \rho_0 = -m'(u)/G_N = \mathcal{F} \tag{41.41}$$

It is related to the locally observed flux

$$\mathcal{F}_o = 4\pi r^2 \rho_o = (-m'(u)/G_N)(u^\bullet)^2 \tag{41.42}$$

by

$$\mathcal{F}_o = \mathcal{F}_\infty (u^\bullet)^2 = \mathcal{F}_\infty (E_u + r^\bullet)^{-2} \quad . \tag{41.43}$$

As for the Hubble distance - redshift relation (section 34.9), the double redshift factor is due to (a) the redshift of the energy (e.g. of each individual photon) as it moves outwards and (b) the dilation of the time-interval over which the energy is emitted (e.g. of the rate at which photons are emitted).

41.4 Radial Timelike Geodesics for Outgoing Vaidya

Up to now, we considered an observer with an arbitrary timelike radial worldline parametrised by $x^\alpha = x^\alpha(\tau_o)$, with 4-velocity $(x^\alpha)^\bullet$. We now specialise to timelike radial geodesics. The condition for a timelike proper time normalised 4-velocity is still (41.29),

$$2u^\bullet r^\bullet + f(u,r)(u^\bullet)^2 = +1 \quad . \tag{41.44}$$

but now the dynamics is governed by the radial Lagrangian

$$\mathcal{L} = -\tfrac{1}{2} f(u,r)(u^\bullet)^2 - r^\bullet u^\bullet \quad . \tag{41.45}$$

In particular the (lightcone) energy (per unit mass) conjugate to u is

$$E_u = -\frac{\partial \mathcal{L}}{\partial u^\bullet} = f u^\bullet + r^\bullet \quad , \tag{41.46}$$

thus justifying the notation / abbreviation introduced in (41.30). Equation (41.32),

$$(r^\bullet)^2 + f(u,r) = (E_u)^2 \quad , \tag{41.47}$$

then has the usual interpretation of a one-dimensional efffective potential equation for the radial dynamics. The main difference from the static case is that here E_u is not conserved.

For a solution of the Euler-Lagrange equations, one can determine the time-dependence of E_u from

$$E_u^\bullet = -\frac{\partial \mathcal{L}}{\partial u} = -\frac{m'(u)}{r}(u^\bullet)^2 \quad . \tag{41.48}$$

Using (41.42), this can also bewritten in the intuitively attractive form

$$E_u^\bullet = \frac{G_N \mathcal{F}_o}{r} \quad , \tag{41.49}$$

stating that the change in energy of the particle is due to the energy flux (luminosity) of the background energy-momentum tensor.

The radial equation of motion turns out to take the remarkably simple form

$$r^{\bullet\bullet} = -\frac{G_N \mathcal{F}_o}{r} - \frac{m(u)}{r^2} \quad . \tag{41.50}$$

It is best obtained not by differentiation of (41.47) (and division by $r^\bullet$ assuming that one is not dealing with circular paths), since it requires a bit of rearrangement to put the resulting equation into the form (41.50), but rather as the Euler-Lagrange equation for r, rewritten using (41.31). Indeed, the Euler-Lagrange equation reads

$$\frac{d}{d\tau_o}\frac{\partial \mathcal{L}}{\partial r^\bullet} = \frac{\partial \mathcal{L}}{\partial r} \quad \Leftrightarrow \quad u^{\bullet\bullet} = \frac{m(u)}{r^2}(u^\bullet)^2 \tag{41.51}$$

while

$$-\frac{u^{\bullet\bullet}}{(u^\bullet)^2} = \left(\frac{1}{u^\bullet}\right)^\bullet = (r^\bullet + E_u)^\bullet \quad . \tag{41.52}$$

Using (41.49) and (41.52), (41.51) can be written in the form (41.50).

<u>Remarks:</u>

1. This shows that the energy-flux gives a new non-Newtonian long-range contribution to the gravitational force, induced by the varying Newtonian term $m(u)/r^2$.

2. The new force term $-G_N \mathcal{F}_o/r$ potentially dominates at large distances. The critical radius at which the new term begins to dominate is at $r \sim m/\mathcal{F}_o$ (factors of c suppressed), which is estimated by Lindquist et al to be totally irrelevant for solar system dynamics, but might play a role in highly relativistic phases of gravitational collapse.

3. If one considers non-radial motion, with conserved angular momentum L, then the additional terms in the radial equation of motion are just the standard angular momentum barrier term $\sim L^2/r^3$ and the standard general relativistic correction term $\sim m(u)L^2/r^4$.

41.5 Future Incompletetness of Outgoing Eddington-Finkelstein Coordinates

We will now look at some issues related to the potential future incompleteness of the outgoing coordinate system at $u = +\infty$ for the general outgoing Vaidya metric. We could have analogously discussed the potential past-incompleteness of the ingoing coordinate system at $v = -\infty$, but past horizons are generally considered to be less physically relevant than future horizons, which can form in the process of gravitational collapse, and thus we focus on the outgoing case.

To set the stage for the subsequent discussions, and to remind ourselves of the basic properties of outgoing coordinates, we will briefly recall the solutions for ingoing null geodesics in these coordinates for the two special cases $m(u) = 0$ (Minkowski space) and $m(u) = m_0$ constant (Schwarzschild in outgoing Eddington-Finkelstein coordinates that cover the Schwarzschild patch as well as the past white hole region).

1. For $m(u) = 0$ one has Minkowski space-time and the coordinate u is related to the usual Minkowski coordinates (t, r) by $u = t - r$. The 1st order geodesic equations (41.26) are simply
$$\dot{r} = r/\tau \quad , \quad \dot{u} = -2r/\tau \ , \tag{41.53}$$
and therefore
$$r(\tau) = -E\tau \quad , \quad u(\tau) = 2E\tau + c \qquad (-\infty < \tau < 0) \ , \tag{41.54}$$
with $E > 0$ and c integration constants. As τ evolves from $-\infty$ to 0, r decreases from $+\infty$ to 0, and u increases from $-\infty$ to the finite value c. At $\tau = 0$, the lightray reaches the utterly harmless coordinate singularity at $r = 0$ and the lightray then bounces back to positive values of r (i.e. turns into an outgoing lightray) and can be extended all the way to $\tau \to +\infty$, where $u \to +\infty$ as well. Thus the complete solution can be described by
$$r(\tau) = E|\tau| \quad , \qquad u(\tau) = 2E\tau + c \qquad (-\infty < \tau < +\infty) \ . \tag{41.55}$$
The integration constant c can be identified with (and used to construct) the ingoing lightcone coordinate $v = t + r$ from the outgoing lightcone coordinate $u = t - r$. Indeed, prior to the extension to $\tau > 0$ (i.e. for geodesics that are ingoing) one has
$$c = u - 2E\tau = u + 2r = t + r = v \ , \tag{41.56}$$
so that ingoing null geodesics are lines of constant v.

2. For $m(u) = m_0 > 0$ a positive constant one has the Schwarzschild geometry, and the coordinate u is related to the usual Schwarzschild time coordinate t and the tortoise coordinate r^*,
$$r^* = r + 2m \log |r - 2m| \ , \tag{41.57}$$

by $u = t - r^*$. The 1st order geodesic equations (41.26) are simply

$$\dot{r} = (r - 2m_0)/\tau \quad , \quad \dot{u} = -2r/\tau \ , \tag{41.58}$$

which integrate to

$$r(\tau) = 2m_0 - E\tau \quad , \quad u(\tau) = -4m_0 \log|\tau| + 2E\tau + c \qquad (-\infty < \tau < 0) \ . \tag{41.59}$$

In this case, the situation is quite different. As $\tau \to 0_-$, one has $r \to 2m_0$ and $u \to +\infty$. This is an infinite redshift surface and the future event horizon (points on the past horizon correspond to $r = 2m$ and u finite). Thus the space-time and the coordinates need to be extended beyond $u = +\infty$.

In the present case this is easily done by noting that the integration constant c is (up to other constants) equal to the ingoing Eddington-Finkelstein coordinate $v = t + r^*$. Indeed,

$$c = u(\tau) + 4m_0 \log|\tau| - 2E\tau = u(\tau) + 2r^*(\tau) + \tilde{c} = v(\tau) + \tilde{c} \ . \tag{41.60}$$

Thus ingoing radial lightrays are lines of constant v and the ingoing Eddington-Finkelstein coordinates (v, r) cover the Schwarzschild patch as well as its future extension beyond the future horizon (now located at $r = 2m$ with v finite, $v = -\infty$ corresponding to the past horizon at which the newly constructed ingoing Eddington-Finkelstein coordinates (v, r) are incomplete).

41.6 Infinite Gravitational Redshift and Future Incompleteness

In section 41.3, we determined the gravitational redshift in the outgoing Vaidya geometry. We now try to determine and locate possible surfaces of infinite redshift. To that end, note that it follows from (41.33) that infinite time dilation or infinite gravitational redshift occurs when $E_u + r^\bullet \to 0$. It then follows from (41.32) that a necessary condition for this to occur is that $f(u, r) \to 0$ or $r \to 2m(u)$ (from "above"),

$$\text{infinite redshift:} \quad E_u + r^\bullet \to 0 \quad \Rightarrow \quad r \to 2m(u) \ . \tag{41.61}$$

Since $f > 0$ for $r > 2m(u)$, and $u^\bullet > 0$ for a future oriented path it then follows from (41.30) that

$$E_u + r^\bullet = f u^\bullet + 2r^\bullet \to 0 \quad \Rightarrow \quad r^\bullet < 0 \ . \tag{41.62}$$

It is easy to see that this cannot occur at finite u:

- If $u^\bullet$ remained bounded as $r \to 2m(u)$, then (41.29) would imply

$$f(u, r) \to 0 \quad \Rightarrow \quad 2u^\bullet r^\bullet \to +1 \ , \tag{41.63}$$

which rules out a negative $r^\bullet$. Indeed, $r = 2m(u)$ with u finite is exactly like the past event horizon for the Schwarzschild metric which can only be crossed along future-directed paths in the direction of increasing r (and in section 32.9 we already identifed this surface concretely as what is known as the past apparent or trapping horizon of the outgoing Vaidya metric).

- The fact that $u^\bullet$ cannot be bounded as $r \to 2m(u)$ is also implied by (41.31),

$$E_u + r^\bullet \to 0 \quad \Rightarrow \quad u^\bullet = (E_u + r^\bullet)^{-1} \to +\infty \ . \tag{41.64}$$

Thus there is an infinite redshift surface (of a freely falling relatively to a static observer) at $r = 2m(u = \infty)$.

We will now show that, as in the Schwarzschild case, this infinite redshift surface is at finite affine distance, i.e. that this surface can be reached in finite proper time (for timelike geodesic observers) or affine parameter (for ingoing lightrays). This means that the outgoing Vaidya coordinates are (like their Schwarzschild counterparts, the outgoing Eddington-Finkelstein coordinates) future-incomplete and that the space-time needs to be extended beyond this infinite redshift surface (an issue we will briefly turn to afterwards).

To that end we use the 2nd of the equations (41.26), which we integrate to

$$\dot{u} = -\frac{2r}{\tau} \quad \Rightarrow \quad \frac{d\tau}{\tau} = -\frac{du}{2r(u)} \quad \Rightarrow \quad \log|\tau| = -\int du/2r(u) \ , \tag{41.65}$$

or

$$|\tau| = \mathrm{e}^{-\int du/2r(u)} \ . \tag{41.66}$$

As $r(u) \to 2m(u)$, the leading term in this integral is

$$-4\log|\tau| = \int du/m(u) + \ldots \tag{41.67}$$

It follows from (41.67) that r reaches $2m(u)$ at the finite time $\tau = 0$ (selected by the choice of integration constant) iff the integral $\int du/m(u)$ diverges for large u, i.e. iff $m(u)$ grows slower than linearly at large u. Since we are only considering non-increasing $m(u)$ anyway, this shows that for any non-increasing function $m(u)$ that is not identically zero for some $u \geq u_0$ (then the previous reasoning leading to the conclusion that $r \to 2m(u)$ requires $u \to \infty$ does not apply) the surface of infinite redshift $r = 2m(\infty)$ is at finite affine distance.

This includes, in particular, the following cases:

1. Consider first the case when $m(u)$ is bounded away from zero, i.e. when one has

$$m(u) \geq m_0 > 0 \ . \tag{41.68}$$

In this case one can reduce the argument to that for the Schwarzschild metric with constant mass m_0. Indeed, in this case the integral (41.67) is

$$-4\log|\tau| = \int \frac{du}{m(u)} \leq \int \frac{du}{m_0} = u/m_0 \tag{41.69}$$

or

$$u \geq -4m_0 \log|\tau| \quad . \tag{41.70}$$

Thus one reproduces the Schwarzschild result (with inequality because the mass was at least as big as m_0), with the conclusion

$$u \to \infty \quad \text{for} \quad \tau \to 0_- \quad . \tag{41.71}$$

2. This does not yet show what happens in the case where $m(u) \to 0$ asymptotically for $u \to \infty$. A priori it is conceivable that whether one finds the Minkowski space behaviour ($u \to \infty$ only as $\tau \to \infty$, no need for completion) or the Schwarzschild behaviour ($u \to \infty$ for finite τ, completion required) depends on the rate at which $m(u) \to 0$.

   If one assumes that for large u the mass function $m(u)$ behaves like $m(u) \sim u^{-a}$ for some $a > 0$, say, or goes to zero exponentially, then for large u the integral in (41.67) gives

$$\begin{aligned} m(u) \sim u^{-a} \quad &\Rightarrow \quad -\log|\tau| \sim u^{a+1} \\ m(u) \sim \mathrm{e}^{-au} \quad &\Rightarrow \quad -\log|\tau| \sim \mathrm{e}^{au} \quad , \end{aligned} \tag{41.72}$$

   so this still implies $u \to \infty$ for $\tau \to 0_-$ for any $a > 0$ (actually for any $a > -1$ in the power-law case), in agreement with the general argument.

Thus we find that quite generally $r = 2m(u = \infty)$ behaves exactly like the future horizon of a static Schwarzschild black hole, which also "sits at" $t = +\infty$, where t is the Schwarzschild time, equivalently (up to some constant factor) the proper time of a non-geodesic static observer, or at $u = +\infty$, where u is the retarded Eddington-Finkelstein coordinate.

41.7 Some Comments on Future Extensions of Outgoing Vaidya

As we have seen above, the outgoing Vaidya coordinates are future-incomplete and one needs to future-extend the space-time beyond $u = +\infty$. This issue was first pointed out by Lindquist et al (footnote 188) who stated, however, that they were unable to find an extension. Indeed, this issue turns out to be far from trivial and is not resolved in general.

As regards this issue of future incompleteness and future extension, the two special cases recalled in section 41.5 above should be prototypical in the sense that for the non-increasing non-negative mass function $m(u)$ of the outgoing Vaidya metric one only has the following 4 options:

1. $m(u)$ decreases to 0 at some finite value u_0 of u (and then remains 0);

2. $m(u)$ decreases to a positive constant value $m_0 > 0$ at some finite value $u = u_0$ (and then remains constant);

3. $m(u)$ decreases to 0 asymptotically for $u \to \infty$;

4. $m(u)$ decreases to a positive constant value $m_0 > 0$ asymptotically for $u \to \infty$.

One then has the following situation:

1. In the first case, for $u \geq u_0$ the metric is just the Minkowski space in outgoing lightcone coordinates and the extension of the space-time should not be an issue. However, since this means that $r = 2m(u = \infty) \to 0$, the singularity of the metric at $r = 0$ is potentially dangerous. And indeed it is shown by Waugh and Lake (footnote 187) that backscattered light emitted towards smaller values of r is infinitely blueshifted for $r \to 0$. This implies that the backreaction of the backscattering cannot be ignored, and that classically the process of a black hole or star radiating away all its mass to leave behind Minkowski space is unstable to this backreaction.

2. In the second case, the metric is the Schwarzschild metric for $u \geq u_0$ and the future extension of the metric is well known (and can e.g. be described by ingoing Eddington-Finkelstein coordinates or by Kruskal-Szekeres coordinates).

3. In the third case, the infinite redshift surface recedes to $r \to 0$, and the space-time geometry is potentially singular there, either already prior to taking into account backreaction or, as above, once backscattering is taken into account.

4. This leaves the fourth case, with a surface of infinite redshift at the finite value $r = 2m(u = \infty)$ of the radius that can be reached in finite proper time or affine parameter and behaves very much like a future horizon as the potentially most interesting case to look at.

Finding a future-extension turns out to be not completely straightforward even in this case, and I will just close this section with some remarks:

- A future extension of the outgoing Vaidya metric in this case was first proposed by W. Israel in 1967.[189] It is based on a suitable generalisation of the remarkable Israel coordinates for the Schwarzschild metric discussed in section 27.11.

- However, one of the problems, already realised and discussed by W. Israel, is that, in order to extend the metric beyond $u = \infty$, one also has to extend the mass function $m(u)$ to that region, and it is not obvious (and in fact not true) that

[189] W. Israel, *Gravitational collapse of a radiating star*, Physics Letters A24 (1967) 184.

there is a unique way of doing this. This issue was further discussed by Fayos et al. who suggested a slight modification of the procedure proposed by Israel.[190]

- One could try to side-step this non-uniqueness issue by attempting to solve the Einstein equations directly in Kruskal-like double-null coordinates, but this generally leads to equations that cannot be solved analytically.[191]

[190] F. Fayos, M. Martin-Prats, J. Senovilla, *On the extension of Vaidya and Vaidya-Reissner-Nordström spacetimes*, Class. Quant. Grav. 12 (1995) 2565-2576.

[191] See e.g. B. Waugh, K. Lake, *Double-null coordinates for the Vaidya metric*, Phys. Rev. D34 (1986) 2978-2984, and F. Girotto, A. Saa, *Semi-analytical approach for the Vaidya metric in double-null coordinates*, `arXiv:gr-qc/0406067`.

42 Vaidya Metrics III: Linear Mass $m(v) = \mu v$ (a case study)

In the following, in order to illustrate some of the properties of Vaidya metrics discussed in the previous sections, we will focus mainly on the ingoing Vaidya metric (40.2), and in particular on the case where the mass function is a linear function of v, $m(v) = \mu v$. This tractable example already displays a rich and intricate structure (with a subtle dependence on the value of the mass parameter μ), and gives a good idea of the complexity of the properties of the general Vaidya metric.

We thus consider the Vaidya metric with a linear mass function

$$m(v) = \mu v \quad . \tag{42.1}$$

In order to avoid unphysical negative masses, we will only consider this space-time for $v \geq 0$ and glue it to empty Minkowski space with metric $ds^2 = -dv^2 + 2dv\, dr + r^2 d\Omega^2$ at $v = 0$. Thus we are considering the Vaidya metric

$$ds^2 = -f(v,r)dv^2 + 2dv\, dr + r^2 d\Omega^2 \tag{42.2}$$

with

$$f(v,r) = 1 - \frac{2m(v)}{r} \quad , \quad m(v) = \begin{cases} 0 & v \leq 0 \\ \mu v & v \geq 0 \end{cases} \tag{42.3}$$

Evidently this is a rather unphysical metric for $v \to \infty$ (the mass tending to infinity), and we will rectify this later on by glueing on a Schwarzschild metric at some time $v_0 > 0$ (with constant mass $m_0 = \mu v_0$).

42.1 Outgoing Lightrays for $m(v) = \mu v$: Derivation 1

In order to determine the event horizon, we will now first determine the outgoing lightrays in the linear mass Vaidya metric. To that end, we need to solve the 2nd order null geodesic equations (41.2) or the 1st order equations (41.12) and (41.13). Remarkably, both sets of equations simplify tremendously in the linear mass case $m(v) = \mu v$ (and we will see later on that this can be attributed to an additional symmetry of the problem arising from a homothety of the metric):

1. For a linear mass function, and written in terms of the new (non-affine) parameter t with $\tau = \exp t$, the 1st order equations (41.12) and (41.13) are simply a coupled system of linear homogeneous differential equations with constant coefficients, namely

$$\left.\begin{aligned} dr/dt &= r - 2\mu v \\ dv/dt &= 2r \end{aligned}\right\} \quad \Leftrightarrow \quad \frac{d}{dt}\begin{pmatrix} r \\ v \end{pmatrix} = \begin{pmatrix} 1 & -2\mu \\ 2 & 0 \end{pmatrix}\begin{pmatrix} r \\ v \end{pmatrix} \tag{42.4}$$

   This system of equations can be solved in standard ways, essentially by diagonalising and/or exponentiating the (2×2)-matrix appearing in the above equation.

2. Alternatively, one observes that for a linear mass function the non-linear coupled 2nd order null geodesic equations (41.2) reduce to two decoupled (and identical) linear harmonic oscillator equations for $r(\tau)$ and $v(\tau)$ respectively. Indeed, note that for a linear mass function $m(v) = \mu v$ (41.14) and (41.15) reduce to

$$\begin{aligned} \ddot{r}(\tau) + \frac{4\mu}{\tau^2}\, r(\tau) &= 0 \\ \ddot{v}(\tau) + \frac{4\mu}{\tau^2}\, v(\tau) &= 0 \ . \end{aligned} \tag{42.5}$$

This is simply the equation of a time-dependent harmonic oscillator with the special (scale-invariant) potential $\sim \tau^{-2}$ and with the same frequency 4μ for both r and v. It is straightforward to solve this equation for any μ.

Proceeding either way, one quickly learns that curiously the case $\mu = 1/16$ is special and, in fact, slightly more complicated (because of a resonance behaviour). In this section, we will determine the outgoing null geodesics for all μ using the equations (42.5). An alternative derivation based on the matrix equation (42.4) is given in the appendix (section 42.6).

To proceed, we first observe that a priori the solutions to the 2nd order equations (42.5) will involve 4 integration constants, but that (41.12) and (41.13) provide 2 relations among them. In practice, it is then most convenient to determine the general solution for $r(\tau)$, and to then determine $v(\tau)$ algebraically from $r(\tau)$ and $\dot{r}(\tau)$ using (41.12),

$$\tau\dot{r}(\tau) = r(\tau) - 2\mu v(\tau) \quad \Leftrightarrow \quad v(\tau) = \frac{1}{2\mu}\left(r(\tau) - \tau\dot{r}(\tau)\right) \ , \tag{42.6}$$

In order to solve (42.5), we write (reparametrise) the "frequency" 4μ in terms of a parameter ω as

$$4\mu = \omega(1-\omega) \quad \Leftrightarrow \quad \omega_\pm = \tfrac{1}{2}(1 \pm \sqrt{1-16\mu}) \ . \tag{42.7}$$

Then it is evident that (42.5) is solved by $r(\tau) \sim \tau^{\omega_\pm}$, since

$$\frac{d^2}{d\tau^2}\,\tau^\omega = \frac{\omega(\omega-1)}{\tau^2}\,\tau^\omega \ . \tag{42.8}$$

This already gives us the general solution unless $\omega_+ = \omega_-$, i.e. unless $\omega = 1/2$, $\mu = 1/16$, and the latter case requires a separate treatment.

1. For $\mu \neq 1/16$ we have

$$\mu \neq \tfrac{1}{16}: \quad \begin{aligned} r(\tau) &= c_+\tau^{\omega_+} + c_-\tau^{\omega_-} \\ v(\tau) &= 2(c_+/\omega_+)\tau^{\omega_+} + 2(c_-/\omega_-)\tau^{\omega_-} \ . \end{aligned} \tag{42.9}$$

Depending on the sign of

$$\Delta = 1 - 16\mu \ . \tag{42.10}$$

there are two different subcases:

(a) $m < 1/16, \Delta > 0$: the two roots are real and positive,

$$\omega_+ > \omega_- > 0 \quad , \tag{42.11}$$

and the integration constants $c_\pm$ are real.

(b) $\mu > 1/16, \Delta < 0$: one has two complex conjugate roots

$$\omega_\pm = \tfrac{1}{2} \pm i\Omega \quad , \quad \Omega^2 = (4\mu - 1/4) > 0 \quad . \tag{42.12}$$

In order for the solution to be real, the integration constants $c_\pm$ should then also be complex conjugates of each other,

$$c_\pm = c_1 \pm ic_2 \quad . \tag{42.13}$$

Noting that

$$\tau^{\omega_\pm} = \tau^{1/2} \mathrm{e}^{\pm i\Omega \log \tau} \tag{42.14}$$

the general solution can then, if desired, be recast into a manifestly real (but not necessarily more enlightning) form, now expressed in terms of real linear combinations of $\tau^{1/2} \cos(\Omega \log \tau)$ and $\tau^{1/2} \sin(\Omega \log \tau)$.

2. For $\mu = 1/16$, i.e. $\Delta = 0$, one has $\omega_+ = \omega_- = 1/2$. A second linearly independent solution of (42.5) (obtained e.g. by demanding a non-zero constant Wronskian) is $\tau^{1/2} \log \tau$. Thus the general solution (again making use of (42.6) for $v(\tau)$) has the form

$$\mu = \tfrac{1}{16} : \quad \begin{aligned} r(\tau) &= c\, \tau^{1/2} + d\, \tau^{1/2} \log \tau \\ v(\tau) &= 4r(\tau) - 8d\, \tau^{1/2} = (4c - 8d)\, \tau^{1/2} + 4d\, \tau^{1/2} \log \tau \quad . \end{aligned} \tag{42.15}$$

It is clear from the above solutions (42.9) and (42.15) that the qualitative behaviour of outgoing lightrays (and thus the lightcones) depends crucially on whether $\Delta > 0$, $\Delta = 0$ or $\Delta < 0$. We will make extensive use of these solutions below in order to determine the horizon structure in the eternal Vaidya geometry and the Vaidya-glued-to-Schwarzschild geometry, and I will just add some general qualitative comments here:

REMARKS:

1. One noteworthy feature of the solutions (42.9) and (42.15) is that a simultaneous scaling of (c_+, c_-) in (42.9) or (c, d) in (42.15) is simply equivalent to a scaling of the coordinates (v, r). This fact reflects a scaling symmetry of the metric that we will discuss below.

2. It turns out that, among all the above solutions, a special role will be played by those null geodesics that are *invariant* under this scaling, i.e. which are such that r and v are linearly related so that the geodesics are straight lines ($d^2r/dv^2 = 0$) in an (r, v)-diagram.

From (41.9), which in the present (linear mass) case reads

$$2\frac{dr}{dv} = f(v,r) \quad \Leftrightarrow \quad \frac{dr}{dv} + \frac{\mu v}{r} = \frac{1}{2} \quad , \tag{42.16}$$

one finds

$$\frac{d^2r}{dv^2} = \frac{\mu}{r}\left(\frac{v}{2r}\left(1 - \frac{2\mu v}{r}\right) - 1\right) \quad . \tag{42.17}$$

Thus $r''(v) = 0$ along

$$\begin{aligned} \frac{v}{2r}\left(1 - \frac{2\mu v}{r}\right) = 1 \quad &\Leftrightarrow \quad r^2 - vr/2 + \mu v^2 = 0 \\ &\Leftrightarrow \quad (r - v/4)^2 = (1 - 16\mu)(v/4)^2 \end{aligned} \tag{42.18}$$

with solution

$$r = \frac{v}{4}\left(1 \pm \sqrt{\Delta}\right) = \frac{\omega_\pm}{2}\, v \quad \Leftrightarrow \quad v = \frac{\omega_\mp}{2\mu}\, r \quad . \tag{42.19}$$

Again we see that the value $\mu = 1/16, \Delta = 0$ plays a special role, as these lines exist only for $\Delta \geq 0$.

3. For $\Delta > 0$ these are precisely the null geodesics with either $c_- = 0$ or $c_+ = 0$,

$$\Delta > 0 : \quad c_\mp = 0 \quad \Rightarrow \quad r(\tau) = \frac{\omega_\pm}{2}\, v(\tau) \quad , \tag{42.20}$$

while for $\Delta = 0$ these are the lines with $d = 0$,

$$\Delta = 0 : \quad d = 0 \quad \Rightarrow \quad r(\tau) = v(\tau)/4 \quad . \tag{42.21}$$

In these cases the remaining integration constant ($c_\pm$ or c) is redundant, since changing this integration constant is equivalent to a rescaling of the affine parameter τ and thus does not change the null geodesic. Moreover, since they are linear, these null geodesics are then invariant under a simultaneous scaling of r and v, as anticipated.

4. As the constant c in (42.15) can be shifted by a scaling of τ, one can set $c = 0$ without loss of generality if one permits oneself this scaling. Introducing the new (non-affine!) parameter λ through

$$\tau = \mathrm{e}^{-2\lambda} \tag{42.22}$$

(so that scaling τ corresponds to shifting λ) and defining $C = -2d$, the solution takes the form

$$r(\lambda) = C\lambda \mathrm{e}^{-\lambda} \quad , \quad v(\lambda) = 4C\lambda\mathrm{e}^{-\lambda} + 4C\mathrm{e}^{-\lambda} \quad . \tag{42.23}$$

This is the form of the solution given by Poisson.[192] It is more convenient, however, also for purposes of matching the Vaidya and Schwarzschild geodesics, to use the affinely parametrised solution given in (42.15).

[192] E. Poisson, *A Relativist's Toolkit: the Mathematics of Black Hole Mechanics*, section 5.7.2.

42.2 Some Comments on Homotheties, Geodesics and Wronskians

We have seen in a number of different ways that Vaidya metrics with a linear mass function have some special properties which hint at an underlying additional symmetry in this case. In this section I collect some (not indispensable) comments related to this symmetry, to homotheties, Wronskians and the like.

To discover and describe this additional symmetry, note that from the null geodesic equations (41.2) one finds for a general mass function $m(v)$

$$\frac{d}{d\tau}(r\dot{v} - v\dot{r}) = (vm'(v) - m(v))\frac{\dot{v}^2}{r} \quad . \tag{42.24}$$

Thus

$$D := r\dot{v} - v\dot{r} \tag{42.25}$$

is a constant of motion for outgoing null geodesics iff $m(v) = v\partial_v m(v)$, i.e. iff $m(v) = \mu v$ is a linear function of v,

$$\frac{d}{d\tau}D = 0 \quad \Leftrightarrow \quad m(v) = \mu v \quad . \tag{42.26}$$

This constant of motion can be understood and interpreted as the conserved charge associated to the dilatation symmetry (homothety) of the Vaidya metric with a linear mass function,

$$(v,r) \to (\lambda v, \lambda r) \quad \Rightarrow \quad ds^2 = -(1-2\mu v/r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \to \lambda^2\,ds^2 \quad . \tag{42.27}$$

Indeed, this dilatation symmetry is generated by the conformal Killing vector

$$C = v\partial_v + r\partial_r \quad , \quad \nabla_\alpha C_\beta + \nabla_\beta C_\alpha = 2g_{\alpha\beta} \tag{42.28}$$

and therefore (see the discussion in section 10.2, in particular equation (10.8)) leads to the conserved charge

$$Q_C = g_{\alpha\beta}C^\alpha \dot{x}^\beta \tag{42.29}$$

for any null geodesic,

$$\frac{d}{d\tau}Q_C = (\nabla_\alpha K_\beta)\dot{x}^\alpha \dot{x}^\beta = g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta = 0 \quad . \tag{42.30}$$

To see that D as defined in (42.25) is the same as the Q_C in (42.29), note that from (42.29) one has

$$Q_C = vg_{v\beta}\dot{x}^\beta + rg_{r\beta}\dot{x}^\beta = -f(v,r)v\dot{v} + r\dot{v} + v\dot{r} \quad . \tag{42.31}$$

Using the null condition (41.9) for radial null geodesics, $f(v,r)\dot{v} = 2\dot{r}$, one then finds (42.25),

$$f(v,r)\dot{v} = 2\dot{r} \quad \Rightarrow \quad Q_C = D \quad . \tag{42.32}$$

Remarks:

1. The scaling symmetry (42.27) implies that if $(r(\tau), v(\tau))$ is a solution to the geodesic equations, then so is $(\lambda r(\tau), \lambda v(\tau))$, something that we had already observed post facto based on the explicit solutions (42.9) and (42.15) of the null geodesic equations. We see e.g. from (42.25) that this scaling is essentially equivalent to changing the integration constant D by $D \to \lambda^2 D$, and therefore the freedom to choose the value of D reflects this scale invariance.

2. E.g. by calculating explicitly $r\dot{v} - \dot{r}v$ from the solutions (42.9) and (42.15), one finds that the constant of motion D is related to the integration constants $c_\pm$ or (c, d) by
$$\begin{aligned} \mu \neq \tfrac{1}{16}: &\quad D = -2\frac{c_+ c_-}{\omega_+\omega_-}\Delta \\ \mu = \tfrac{1}{16}: &\quad D = 8d^2 \quad . \end{aligned} \tag{42.33}$$
From this one sees that the scale-invariant geodesics (42.20) (with $c_+ = 0$ or $c_- = 0$) and (42.21) (with $d = 0$) are precisely the geodesics with $D = 0$ (which is the only scale-invariant value under $D \to \lambda^2 D$).

3. An obvious alternative interpretation of $D = r\dot{v} - v\dot{r}$ is as that of the Wronskian of the two solutions of the same harmonic oscillator equation (42.5).
$$\left.\begin{aligned} \ddot{r}(\tau) + \tfrac{4\mu}{\tau^2}\, r(\tau) = 0 \\ \ddot{v}(\tau) + \tfrac{4\mu}{\tau^2}\, v(\tau) = 0 \end{aligned}\right\} \quad \Rightarrow \quad \frac{d}{d\tau}(r\dot{v} - v\dot{r}) = 0 \quad . \tag{42.34}$$
The Wronskian $D = 0$ precisely when $v(\tau)$ and $r(\tau)$ are linearly dependent, i.e. when they describe a straight line in an (r, v)-diagram, as we have indeed seen in (42.20) and (42.21).

4. Inserting (41.12) and (41.13) into (42.25), one finds an algebraic relation between $r(\tau)$ and $v(\tau)$, namely
$$\tau\, D = \tau(r\dot{v} - v\dot{r}) = 2r^2 - vr + 2m(v)v = 2r^2 - vr + 2\mu v^2 \quad , \tag{42.35}$$
which can be written as
$$\frac{\tau D}{2} = (r - v/4)^2 + (\mu - 1/16)v^2 \tag{42.36}$$
or
$$4\sqrt{\mu}\,\tau\, D = (4\sqrt{\mu} + 1)(r - \sqrt{\mu}v)^2 + (4\sqrt{\mu} - 1)(r + \sqrt{\mu}v)^2 \quad . \tag{42.37}$$
This can be useful for visualising these null geodesics.

5. Because of their analytic tractability, these linear mass Vaidya metrics (also known as Vaidya metrics describing "self-similar gravitational collapse" due to the existence of the homothety) have been much studied in the literature. Other aspects and consequences of the homothety, in particular in relation to the characterisation and properties of the singularity, have been explored by Lake and Zannias.[193]

[193] K. Lake, T. Zannias, *Naked singularities in self-similar gravitational collapse*, Phys. Rev. D41 (1990) 3866-3868.

42.3 Event vs Apparent Horizons for $m(v) = \mu v$: Overview

With all this detailed knowledge of null geodesics in the linear mass Vaidya space-time at our disposal, it is now straightforward to determine the causal properties of this space-time (and subsequently of the space-time obtained by glueing Schwarzschild to Vaidya). As in our discussion above, the sign of $\Delta = 1 - 16\mu$ and the value $\mu = 1/16$ will turn out to play a special role.

The first and simplest exercise is to determine the apparent horizons and their geometry. The apparent horizon (32.122) is the hypersurface

$$f(v,r) = 0 \quad \Leftrightarrow \quad r = 2\mu v \qquad (v \geq 0) \ . \tag{42.38}$$

Thus in the linear mass case this is a straight line in an (r,v)-diagram. The induced metric (32.150) on the apparent horizon is

$$ds^2|_{f(v,r)=0} = 4m'(v)dv^2 + (2m(v))^2 d\Omega^2 = 4\mu dv^2 + 4\mu^2 v^2 d\Omega^2 \ , \tag{42.39}$$

or, in terms of r,

$$ds^2|_{f(v,r)=0} = \mu^{-1} dr^2 + r^2 d\Omega^2 \ . \tag{42.40}$$

Therefore the apparent horizon is spacelike for $\mu > 0$. Radial outgoing null geodesics reaching $r = 2\mu v$ will attain their maximal radius there and then turn around to smaller values of r.

The intrinsic geometry of the apparent horizon is thus manifestly flat for $\mu = 1$, and while the factor μ^{-1} may look harmless it actually leads to a non-trivial curvature tensor for $\mu \neq 1$, with a curvature singularity at $r = 0$. Explicitly, one finds (see section 8.7, in particular equations (8.90) and (8.92)) that e.g. the non-trivial components of the Riemann tensor and the Ricci scalar are

$$R^{\phi}{}_{\theta\phi\theta} = 1 - \mu \quad , \quad R = 2(1-\mu) r^{-2} \ . \tag{42.41}$$

Further information is obtained from looking at the 2nd derivative d^2r/dv^2, i.e. the 1st derivative of $f(v,r)$. From (42.17) one has

$$\frac{d^2 r}{dv^2} = -\frac{\mu}{r^3}\left[(r - v/4)^2 + (\mu - 1/16)v^2\right] \ , \tag{42.42}$$

and from (42.19) we know that $r''(v) = 0$ along the scale-invariant null geodesics (42.20) and (42.21) given by

$$r = r_{\pm} \equiv \frac{\omega_{\pm}}{2} v \ . \tag{42.43}$$

There are now three different cases to consider, depending on the sign of Δ.

1. $\mu > 1/16$ or $\Delta < 0$

There are no real roots for $\mu > 1/16$ and from (42.42) one sees that $r''(v) < 0$ is strictly negative everywhere. It turns out that in this case no lightray will be able to escape to infinity, i.e. there is no future null infinity $\mathcal{I}^+$ at all, and every lightray ends up in the spacelike singularity at $r = 0$ in the future.

This does not follow just from the fact that $d^2r/dv^2 < 0$. One has to show that all outgoing lightrays eventually reach and subsequently cross the apparent horizon at which $dr/dv = 0$. This can be established by using the explicit real form of the solution, with its characteristic $\tau^{1/2}$-modulated trigonometric dependence on $\Omega \log \tau$.

Thus in this case there is no event horizon. In fact, the entire space-time for $v \geq 0$ looks somewhat like the interior of region II of the Kruskal extension of the Schwarzschild space-time. This is clearly an artefact of having a mass that tends sufficiently rapidly to infinity. This is illustrated in the Penrose diagram 59.[194]

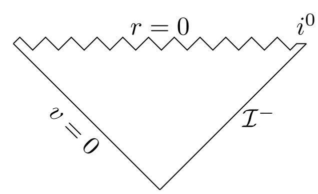


Figure 59: Penrose diagram of the linear mass $m(v) = \mu v$ Vaidya space-time for $v \geq 0$ and $\mu > 1/16$.

Note that it follows from (42.36) and (42.33) that for $\Delta < 0$ outgoing null geodesics satisfy

$$(r - v/4)^2 + (\mu - 1/16)v^2 = |c/\omega|^2(-\Delta)\tau \geq 0 \ , \tag{42.44}$$

where $|c|^2 = c_+c_- = c_+c_+^*$ and $|\omega|^2 = \omega_+\omega_- = \omega_+\omega_+^*$ denote the squares of the absolute value. In particular if $\tau_i > 0$ denotes the (initial) time at which $v(\tau_i) = 0$ (we are discarding $\tau_i = 0$ because for $\tau \to 0$ the functions $\tau^{1/2}\cos(\Omega \log \tau)$ and $\tau^{1/2}\sin(\Omega \log \tau)$ are badly behaved), one has

$$r(\tau_i)^2 = |c/\omega|^2(-\Delta)\tau_i > 0 \ . \tag{42.45}$$

This means that outgoing radial null geodesics start off at $v = 0$ at some positive value of r. As one of these outgoing (families of) null geodesics will turn out to become the event horizon for the metric obtained by combining Vaidya for $0 \leq v \leq v_0$ with Schwarzschild for $v \geq v_0$, this will lead to the conclusion that in the case $\Delta < 0$ the singularity at $r = 0$ is hidden behind an event horizon (no null geodesic emerging from $r = 0$ can escape to infinity).

[194] For this and the subsequent Penrose diagrams, see e.g. W. Hiscock, L. Williams, D. Eerdley, *Creation of particles by shell-focusing singularities*, Phys. Rev. D26 (1982) 751-760, and B. Waugh, K. Lake, *Double-null coordinates for the Vaidya metric*, Phys. Rev. D34 (1986) 2978-2984.

2. $\mu = 1/16$ or $\Delta = 0$

   For $\mu = 1/16$ one has a double root at $r_+ = r_- = v/4$. One has $d^2r/dv^2 \leq 0$, with equality only for the line (null hypersurface) $r = v/4$,

$$\frac{d^2r}{dv^2} = 0 \quad \Rightarrow \quad r = v/4 \ . \tag{42.46}$$

   This hypersurface is generated by the outgoing null geodesics with $D = 0$ (or $d = 0$ in (42.15)).

   It turns out that the behaviour of the other outgoing lightrays depends strongly on whether they are in the region $r > v/4$ or in the region $r < v/4$. If they are in one of those regions initially, they always remain there, and for $r > v/4$ the lightrays escape to infinity while for $r < v/4$ they reach a maximal radius at the apparent horizon $r = v/8$ and then return to smaller values of r. These results follow from the general solution (42.15) which we will analyse in a bit more detail below.

   Therefore the event horizon of the $\mu = 1/16$ linear mass Vaidya black hole is the scale-invariant null hypersurface $r = v/4$, outside the apparent horizon at $r = v/8$. See Figure 60.

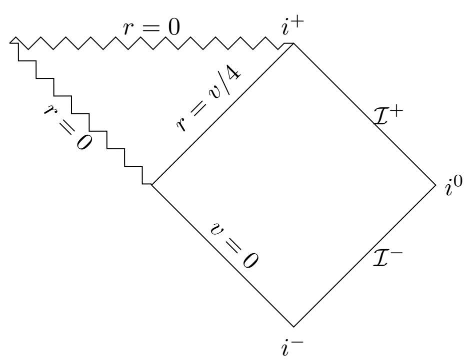


Figure 60: Penrose diagram of the linear mass $m(v) = \mu v$ Vaidya space-time for $v \geq 0$ and $\mu = 1/16$.

3. For $0 < \mu < 1/16$ there are two roots $r_\pm$ and a correspondingly richer phase diagram for the behaviour of outgoing lightrays in the (r, v)-plane. The two hypersurfaces $r = r(v)$ along which $r''(v) = 0$ are generated by the $D = 0$ null geodesics (42.43). They divide the $(r \geq 0, v \geq 0)$ quadrant into 3 wedges:

   (a) For $r > r_+ = (\omega_+/2)v$ one has $r''(v) < 0$, and $c_+ > 0$, $c_- < 0$ (thus $D > 0$).

   (b) For $r = r_+$ one has $r''(v) = 0$, and $c_+ > 0$, $c_- = 0$ (thus $D = 0$).

   (c) For $(\omega_-/2)v = r_- < r < r_+$ one has $r''(v) > 0$, and $c_+ > 0$, $c_- > 0$ (thus $D < 0$).

(d) For $r = r_-$ one has $r''(v) = 0$, and $c_+ = 0$, $c_- > 0$ (thus $D = 0$).

(e) For $r < r_-$ one has $r''(v) < 0$, and $c_+ < 0$, $c_- > 0$ (thus $D > 0$).

Since $2\mu < \omega_-/2$,

$$8\mu < 2\omega_- = 1 - \sqrt{1 - 16\mu} \quad \Leftrightarrow \quad \sqrt{1 - 16\mu} < 1 - 8\mu \tag{42.47}$$

(which is always satisfied in the given range of μ as can be seen by squaring the two positive sides of this inequality), the apparent horizon at $r = 2\mu v$ lies in the lowest wedge. This has the following implications for the behaviour of null geodesics (that can also be checked from the explicit solution (42.9)):

- Null geodesics that start off in the lowest wedge cross the apparent horizon at $r = 2\mu v$ and then return to smaller values of r (and ultimately to $r = 0$ at finite v), regardless of whether they were initially above the apparent horizon (truly outgoing at that time) or below it.
- Null geodesics in the central region, including the two lines $r = (\omega_\pm/2)v$ start off at $r = 0$ for $v = 0$ and reach $r = \infty$ for $v = \infty$, with $r''(v) \geq 0$.
- Null geodesics in the region $r > (\omega_+/2)v$ have $r(v = 0) > 0$ and escape to infinity with $r''(v) < 0$.

It follows that the lower line $r = r_- = (\omega_-/2)v$ is the event horizon of the $\Delta > 0$ linear mass Vaidya black hole. It again lies outside the apparent horizon. See Figure 61.

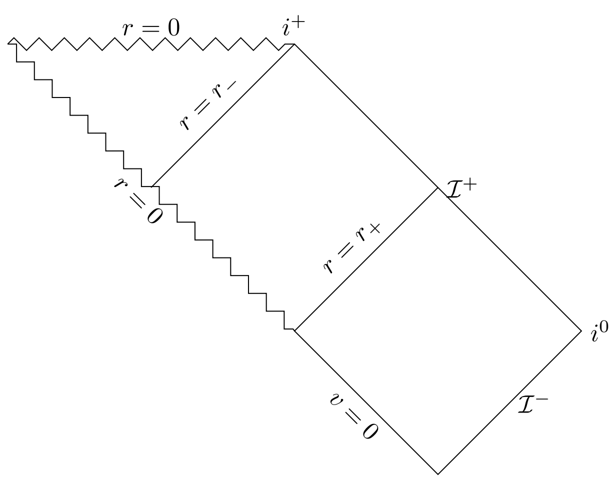


Figure 61: Penrose diagram of the linear mass $m(v) = \mu v$ Vaidya space-time for $v \geq 0$ and $\mu < 1/16$.

42.4 Null Geodesics, Horizons and Singularities for $\mu = 1/16$

In this section we take a slightly more detailed look at the special case $\mu = 1/16$. The solution for outgoing radial null geodesics is given in (42.15), which we now compactly write as

$$r(\tau) = c\,\tau^{1/2} + d\,\tau^{1/2}\log\tau = v(\tau)/4 + 2d\tau^{1/2} \quad . \tag{42.48}$$

The behaviour of these geodesics depends crucially on the sign of the constant d.

1. The solution with $d = 0$ is the special scale-invariant null geodesic $r(\tau) = v(\tau)/4$ (42.21) with

   $$r(\tau) = v(\tau)/4 = c\tau^{1/2} \quad . \tag{42.49}$$

   This requires $c > 0$, and changing c by a constant positive factor is equivalent to scaling τ. This thus describes the same geodesic, just with a different affine parameter. This null geodesic emerges from $r = 0$ at $v = 0$ and $\tau = 0$ and evidently escapes to infinity as $v \to \infty$ or $\tau \to \infty$.

2. For $d \neq 0$ there are two distinct types of geodesics, depending on the sign of d, namely those with $r > v/4$ at all times ($d > 0$) and those with $r < v/4$ at all times ($d < 0$):

   (a) For $d > 0$ one has $r > v/4$. The requirement $v(\tau) \geq 0$ (implying $r(\tau) \geq 0$ in this case) leads to the condition that

   $$\tau \geq \tau_{min} = \mathrm{e}^{(2 - c/d)} : \quad v(\tau_{min}) = 0 \quad , \quad r(\tau_{min}) = 2d\tau_{min}^{1/2} \quad . \tag{42.50}$$

   Regarding the long-term behaviour, one sees that

   $$\dot{r}(\tau) = \frac{r(\tau)}{2\tau} + d\tau^{-1/2} > 0 \tag{42.51}$$

   and $\ddot{r}(\tau) < 0$ at all (finite) times, with $\dot{r}(\tau) \to 0$ for $\tau \to \infty$ (so that qualitatively the behaviour of these lightrays is like that of the cosmological scale factor in the spatially flat $k = 0$ FLRW models). Thus these lightrays escape to infinity.

   (b) For $d < 0$ one has $r < v/4$. The requirement $r(\tau) \geq 0$ (implying $v(\tau) \geq 0$ in this case) leads to the condition that

   $$\tau \leq \tau_{max} = \mathrm{e}^{(c/|d|)} : \quad r(\tau_{max}) = 0 \quad , \quad v(\tau_{max}) = 8d\tau_{max}^{1/2} \quad . \tag{42.52}$$

   In this case one has

   $$\dot{r}(\tau) = \frac{r(\tau)}{2\tau} - |d|\tau^{-1/2} \quad . \tag{42.53}$$

   Thus $\dot{r}(\tau_a) = 0$ at the time τ_a when $r(\tau)$ crosses the apparent horizon,

   $$\dot{r}(\tau_a) = 0 \quad \Leftrightarrow \quad \tau_a = \mathrm{e}^{(c/|d| - 2)} < \tau_{max} \quad \Leftrightarrow \quad r(\tau_a) = v(\tau_a)/8 \quad . \tag{42.54}$$

Monotonicity of the mass function (and the explicit solution) imply that these will then not turn around again (and reach $r = 0$ at $\tau = \tau_{max}$).

Thus the null geodesic $r = v/4$ is the last null geodesic to (barely) escape to infinity, and the event horizon is the null hypersurface,

$$r = v/4 \quad \Rightarrow \quad -(1 - v/8r)dv^2 + 2dv\, dr = 0 \ , \tag{42.55}$$

generated by these null geodesics.

<u>Remarks:</u>

1. The event horizon at $r = v/4$ manifestly lies outside the apparent horizon $r = v/8$. In particular, while the outgoing expansion (32.132)

$$\theta_\ell = \frac{r - 2m(v)}{r^2} = \frac{r - v/8}{r^2} \tag{42.56}$$

is zero on the apparent horizon (by our informal definition of the apparent horizon in section 32.8), it is strictly positive on the event horizon,

$$\theta_\ell|_{v=4r} = \frac{1}{2r} > 0 \ . \tag{42.57}$$

2. The event horizon satisfies $r(v = 0) = 0$, so that it emerges from $r = 0$ at the time $v = 0$. At this time $v = 0$ the "singularity" at $r = 0$ is massless, $m(v = 0) = 0$, and one might be tempted to conclude, e.g. from the Kretschmann scalar (40.11), which now has the form,

$$R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta} = 48\mu^2v^2/r^6 \ , \tag{42.58}$$

that perhaps there is therefore no singularity at all at $v = 0$. However, this is not correct. That one really has a singularity at $r = v = 0$ even though $m(v = 0) = 0$ can be seen by noting that the above expression clearly blows up e.g. when one approaches $v = 0$ along lines with constant v/r,

$$v \to 0, r \to 0, v/r = \text{const} \quad \Rightarrow \quad R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta} \propto r^{-4} \to \infty \ . \tag{42.59}$$

This becomes more manifest when one introduces, instead of r, the scale-invariant coordinate $x = v/r$, say. Then one evidently has

$$R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta} = 48\mu^2x^6v^{-4} \tag{42.60}$$

which is clearly singular for $v \to 0$, x fixed.

3. In terms of this new coordinate x, what appears to be a critical "point" $r = v = 0$, the "origin" of the Vaidya-Bondi Eddington-Finkelstein-like coordinates, is mapped to the entire x-axis (or positive half-line for $v \geq 0$), and thus a coordinate transformation to the coordinates (v, x, θ, ϕ) or (r, x, θ, ϕ) is able to resolve / magnify the singularity at $r = v = 0$ and may also be useful for other purposes.

4. The fact that there are lightrays (albeit only a single S^2-family of lightrays) that can escape from the singularity to infinity means that the singularity is not completely hidden behind the event horizon, and is barely / marginally naked. Such massless singularities are considered to be relatively harmless, and therefore they are not considered to be genuine counterexamples to the spirit of the (or an appropriately formulated) cosmic censorship conjecture.

42.5 Linear $\mu = 1/16$ mass Vaidya glued to Schwarzschild

We now consider a slightly more realistic and interesting scenario in which there is an ingoing shell of radiation (null dust) during a finite interval of time v leaving behind a Schwarzschild black hole. In equations this means that we consider the Vaidya metric

$$ds^2 = -f(v,r)dv^2 + 2dv\,dr + r^2 d\Omega^2 \tag{42.61}$$

with

$$f(v,r) = 1 - \frac{2m(v)}{r} \quad , \quad m(v) = \begin{cases} 0 & v \leq 0 \\ v/16 & 0 \leq v \leq v_0 \\ m_0 = v_0/16 & v \geq v_0 \end{cases} \tag{42.62}$$

Note that the mass function is continuous, but that its derivative $m'(v)$ has a jump,

$$m'(v) = \tfrac{1}{16}[\theta(v) - \theta(v - v_0)] \ , \tag{42.63}$$

which translates into a corresponding discontinuity of the energy-momentum tensor $\sim m'(v)$. There is nothing particularly singular or unphysical about such a discontinuty of the energy-momentum tensor, which also arises when one describes the gravitational field of a star: inside there is matter, outside there is not, so the energy-momentum tensor jumps. Here instead we have influx of radiation between $v = 0$ and $v = v_0$ but none before of after that interval, so the energy-momentum tensor jumps twice. This should be contrasted with the situation of a collapsing "thin shell" of radiation that we considered in section 29.1, for which the mass function itself has a jump, leading to a δ-function localised energy-momentum tensor on the world volume of the shell.

To return to the current setting, for $v \geq v_0$ the metric with $f(v,r)$ given by (42.62) is the Schwarzschild metric, with apparent horizon = event horizon at

$$v \geq v_0 : \quad r_0 = 2m_0 = v_0/8 \ , \tag{42.64}$$

described parametrically by (41.20)

$$r(\tau) = 2m_0 \equiv r_0 \quad , \quad v(\tau) = 2r_0 \log \tau + v_0 \ , \tag{42.65}$$

where we have chosen the freedom to scale τ so that $v(\tau = 1) = v_0$. Thus this describes the location of the horizon for $\tau \geq 1$.

For $0 \le v \le v_0$, the apparent horizon is at $r = v/8$. This matches onto the apparent = event horizon of the Schwarzschild black hole for $v \ge v_0$.

In order to determine the event horizon of the total geometry, we now need to determine the lightray for $0 \le v \le v_0$ that matches onto the above event horizon of the Schwarzschild geometry. Since $r_0 = v_0/8 < v_0/4$, it is clear that this is to be sought among the geodesics given in (42.48) with $d < 0$,

$$r(\tau) = c\,\tau^{1/2} - |d|\,\tau^{1/2}\log\tau = v(\tau)/4 - 2|d|\tau^{1/2} \quad . \tag{42.66}$$

Since

$$r(\tau = 1) = c \quad , \quad v(\tau = 1) = 4c + 8|d| \quad , \tag{42.67}$$

the requirements $r(1) = r_0$ and $v(1) = v_0$ determine

$$r(\tau = 1) = r_0 \quad , \quad v(\tau = 1) = v_0 = 8r_0 \quad \Rightarrow \quad c = r_0 \quad , \quad d = -r_0/2 \ . \tag{42.68}$$

Therefore the event horizon is given by

$$\begin{aligned} r(\tau) &= \begin{cases} -\frac{1}{2}r_0\tau^{1/2}\log\tau + r_0\tau^{1/2} & 0 \le \tau \le 1 \\ r_0 & \tau \ge 1 \end{cases} \\ v(\tau) &= \begin{cases} -2r_0\tau^{1/2}\log\tau + 8r_0\tau^{1/2} & 0 \le \tau \le 1 \\ 2r_0\log\tau + 8r_0 & \tau \ge 1 \end{cases} \end{aligned} \tag{42.69}$$

REMARKS:

1. Both $r(\tau)$ and $v(\tau)$ and their first derivatives are continuous at $\tau = 1$. Note that it is automatic that the null geodesic reaches its maximal radial value at $\tau = 1$, $\dot{r}(\tau = 1) = 0$, because it crosses the apparent horizon $r = v/8$ there, $r(1) = r_0, v(1) = v_0 = 8r_0$. As a consequence, the first derivatives of the two pieces of $r(\tau)$ indeed match at $\tau = 1$, both parts of the event horizon give $\dot{r}(1) = 0$. The same is true for $\dot{v}(1)$; in both cases one has $\dot{v}(1) = 2r_0$.

2. The explicit expression for the event horizon confirms that in the Vaidya region it lies in the region between $r = v/4$ and the apparent horizon at $r = v/8$, thus outside the apparent horizon. Explicitly, for the difference between the event horizon and the apparent horizon (given at each time by $r = v/8$) one has

$$r(\tau) - v(\tau)/8 = -\tfrac{1}{4}r_0\tau^{1/2}\log\tau \ge 0 \quad \text{for} \quad 0 < \tau \le 1 \tag{42.70}$$

(with equality at $\tau = 1$).[195]

[195]For more sophisticated and general studies of the relation and distance between trapping and event horizons see e.g. A. Nielsen, *The spatial relation between the event horizon and trapping horizon*, `arXiv:1006.2448 [gr-qc]`, and I. Booth, J. Martin, *On the proximity of black hole horizons: lessons from Vaidya*, `arXiv:1007.1642 [gr-qc]`.

3. This last equation also gives us information about the rate of expansion of the event horizon, given by (41.12) by

$$\dot{r}(\tau) = (r(\tau) - v(\tau)/8)/\tau = -\tfrac{1}{4} r_0 \tau^{-1/2} \log \tau \geq 0 \quad \text{for} \quad 0 < \tau \leq 1 \ . \tag{42.71}$$

One might perhaps naively have expected the expansion rate of the horizon to *increase* during a period of infalling matter, but this equation shows that quite the opposite is true: during the period that matter falls in, the event horizon of course grows, but its expansion rate *decreases*, the expansion stopping when matter stops falling in at $\tau = 1$. This is a general and somewhat counterintuitive feature of event horizons, reflecting the non-locality of the event horizon. We discussed this in general terms with the help of the Raychaudhuri equation applied to the generating null congruence of an event horizon in section 32.7, and we also observed this in the, in other respects quite different, example of event horizons in the Oppenheimer-Snyder geometry of a collapsing star in section 32.12.

4. The Vaidya horizon generators emerge from $r = 0$ at $v = 0$ (at $\tau = 0$) where, as in the case of the eternal $\mu = 1/16$ Vaidya metric discussed before, one has a massless singularity. However, in contrast to the (future) eternal linear mass case (which can be regarded as being marginally naked), here there are many more null geodesics that escape from $r = 0$ to infinity, not just those along the event horizon. Indeed,

   (a) all geodesics that satisfy $r(v_0) > r_0$ are outside the Schwarzschild horizon, and thus describe outgoing null geodesics in the Schwarzschild geometry that escape to infinity; and

   (b) all geodesics that satisfy $r(v_0) < v_0/4$ emerged from $r = 0$ at time $v = 0$.

   It follows that all the null geodesics that satisfy

$$r_0 < r(v_0) \leq v_0/4 \tag{42.72}$$

   escape from $r = 0$ at $v = 0$ to infinity. In this sense this singularity is naked, and more naked than its (marginally naked) counterpart in the eternal Vaidya metric.

To conclude this discussion, I just mention that the situation for $\mu \neq 1/16$, i.e. for the Vaidya metric with

$$f(v,r) = 1 - \frac{2m(v)}{r} \quad , \quad m(v) = \begin{cases} 0 & v \leq 0 \\ \mu v & 0 \leq v \leq v_0 \\ m_0 = \mu v_0 & v \geq v_0 \end{cases} \tag{42.73}$$

is the following:

1. For $\mu > 1/16$ the entire singularity at $r = 0$ (the v-axis) is hidden behind the event horizon and geodesics that reach infinity emerged from some finite value of r at $v = 0$ (and for $v < 0$ there was just Minkowski space).

2. For $\mu < 1/16$ there are many lightlike geodesics that emerge from $r = 0$ and that reach infinity. Thus this is like the case $\mu = 1/16$, because the geodesics that escape from $r = 0$ ro $r \to \infty$ emerge from $r = 0$ at $v = 0$ (so that again this is a massless singularity). It would have been worse if they had emerged from $r = 0$ at some time $v > 0$, where the singularity is massive, but this does not happen.

It is straightforward to establish all this analytically by using the explicit solution (42.9) for the outgoing lightrays. It is also an instructive (and fairly elementary) exercise to generalise the above discussion to the case where one initially has a constant mass Schwarzschild black hole instead of Minkowski space, i.e. one has a mass function

$$m(v) = \begin{cases} m_0 & v \leq v_0 \\ m_0 + \mu(v - v_0) & v_0 \leq v \leq v_1 \\ m_1 = m_0 + \mu(v_1 - v_0) & v \geq v_1 \end{cases} \tag{42.74}$$

and the corresponding apparent horizon at $r = 2m(v)$.

42.6 Appendix: Outgoing Lightrays for $m(v) = \mu v$: Derivation 2

Here is an alternative derivation of the solutions (42.9) and (42.15) for outgoing null geodesics of the linear mass $m(v) = \mu v$ ingoing Vaidya metric for any μ by determining the solution of the coupled set of linear homogeneous differential equations (42.4) with constant coefficients,

$$\frac{d}{dt}\begin{pmatrix} r \\ v \end{pmatrix} = \begin{pmatrix} 1 & -2\mu \\ 2 & 0 \end{pmatrix}\begin{pmatrix} r \\ v \end{pmatrix} \quad . \tag{42.75}$$

We write this as

$$d\vec{x}/dt = L\vec{x} \tag{42.76}$$

with $\vec{x} = (r, v)^t$ $((.)^t$ denoting the transpose column vector) and

$$L = \begin{pmatrix} 1 & -2\mu \\ 2 & 0 \end{pmatrix} \quad . \tag{42.77}$$

One standard (but in general cumbersome) way to solve this equation is to exponentiate the matrix L,

$$\vec{x}(t) = \mathrm{e}^{tL}\vec{x}(0) \quad . \tag{42.78}$$

Another possibility is to diagonalise L either explicitly or implicitly. We follow the latter approach by making the ansatz

$$\vec{x}(t) = \sum_J \tilde{c}_J \vec{a}_J \mathrm{e}^{\omega_J t} \equiv \sum_J \tilde{c}_J \vec{x}_J(t) \quad , \tag{42.79}$$

where for a $(d \times d)$-matrix L the index $J = 1, 2, \ldots d$ labels the d linearly independent solutions, the frequencies ω_J and their eigendirections $\vec{a}_J$ are to be determined, and the $\tilde{c}_J$ are integration constants. Plugging this ansatz into (42.76) one finds

$$d\vec{x}_J/dt = 0 \quad \Leftrightarrow \quad (L - \omega_J \mathbb{1})\vec{a}_J = 0 \ . \tag{42.80}$$

Thus the ω_J are the eigenvalues of L and the $\vec{a}_J$ are the corresponding eigenvectors. The ω_J are the roots of the degree d polynomial equation

$$\det\left(M(\omega)\right) = 0 \quad \text{with} \quad M(\omega) = L - \omega\mathbb{1} \ . \tag{42.81}$$

There are now several possibilities:

1. If all the eigenvalues / roots are distinct, then (42.79) already gives the general solution.

2. If an eigenvalue is degenerate, but with linearly independent eigenvectors, then (42.79) also already gives the general solution.

3. If eigenvalues and eigenvectors are degenerate, then (42.79) does not provide d linearly independent solutions. A resonance phenomenon (familiar from forced oscillations) arises, resulting in the appearance of a term proportional to $t \exp \omega_J t$. Indeed, in such a case the general solution associated with the degenerate root ω_J (let us assume that it is twice degenerate) is of the form

$$\vec{x}_J = \tilde{c}_J \vec{a}_J e^{\omega_J t} + \tilde{d}_J \left(\vec{a}_J t\, e^{\omega_J t} + \vec{b}_J e^{\omega_J t}\right) \tag{42.82}$$

where $\vec{b}_J$ is the "secondary null vector" of $M(\omega_J)$, characterised by

$$M(\omega_J)\vec{b}_J = \vec{a}_J \quad \Leftrightarrow \quad L\vec{b}_J = \vec{a}_J + \omega_J \vec{b}_J \tag{42.83}$$

and evidently only defined modulo (i.e. up to addition of a multiple of) $\vec{a}_J$. Using (42.83), it is easy to see that indeed the 2nd part of (42.82) is a solution of (42.76),

$$(\frac{d}{dt} - L)\left(\vec{a}_J t\, e^{\omega_J t} + \vec{b}_J e^{\omega_J t}\right) = 0 \ . \tag{42.84}$$

The eigenvectors $\vec{a}_J$ can either be readily constructed by hand (for small d) or, in general, from the minors (cofactors) of the matrix M.

For (2×2)-matrices, the general situation evidently simplifies somewhat:

- First of all, the null eigenvectors of $M(\omega_J)$ can be chosen to be of the form

$$\vec{a}_J = (-M(\omega_J)_{22}, M(\omega_J)_{21})^t \quad \Rightarrow \quad M(\omega_J)\vec{a}_J = 0 \ . \tag{42.85}$$

  If $M_{22} = M_{21} = 0$, one can alternatively construct $\vec{a}_J$ from the components of the 1st row. This only fails if $M(\omega_J) = 0$ identically, in which case the construction of two linearly independent eigenvectors is trivial anyway: one can e.g. choose $(1, 0)^t$ and $(0, 1)^t$.

- Moreover, for (2×2)-matrices, case (2) in the above list can only arise if $M(\omega_J)=0$ identically. Otherwise there will be a resonance iff one has a degenerate root.

Specifically and concretely, in the case at hand, with L the (2×2)-matrix given in (42.77), one has

$$L=\begin{pmatrix}1 & -2\mu\\ 2 & 0\end{pmatrix} \quad\Rightarrow\quad M(\omega)=\begin{pmatrix}1-\omega & -2\mu\\ 2 & -\omega\end{pmatrix} \quad , \tag{42.86}$$

leading to the following intermediate results:

1. The characteristic equation is

$$\omega^2-\omega+4\mu=0 \quad\Rightarrow\quad \omega_J=\tfrac{1}{2}(1\pm\sqrt{1-16\mu})\equiv\omega_\pm \quad , \tag{42.87}$$

   and we see (rediscover from the present point of view) that there are 3 distinct cases, depending on the sign of $\Delta=1-16\mu$ (42.10):

   (a) $0<\mu<1/16$ or $\Delta>0$:
   In this case one has two distinct real roots $\omega_\pm$ with

$$0<\omega_-<\omega_+ \quad . \tag{42.88}$$

   (b) $\mu>1/16$ or $\Delta<0$:
   In this case one has two distinct complex conjugate roots

$$\omega_+=\tfrac{1}{2}(1+i\omega)=\omega_-^* \quad , \quad \omega^2=16\mu-1>0 \quad . \tag{42.89}$$

   (c) $\mu=1/16$ or $\Delta=0$:
   In this case one has one degenerate root $\omega=1/2$, the matrix

$$M(\omega=1/2)=\begin{pmatrix}1/2 & -1/8\\ 2 & -1/2\end{pmatrix} \tag{42.90}$$

   is not identically zero, and therefore in this case the solution will involve an additional

$$te^{t/2}=\tau^{1/2}\log\tau \tag{42.91}$$

   term, as we already saw in the explicit solution (42.48).

2. The eigenvectors $\vec{a}_J=\vec{a}_\pm$ can be chosen to have the simple form (42.85),

$$\vec{a}_\pm=\begin{pmatrix}-M(\omega_\pm)_{22}\\ M_{21}(\omega_\pm)\end{pmatrix}=\begin{pmatrix}\omega_\pm\\ 2\end{pmatrix} \quad . \tag{42.92}$$

3. For $\mu=1/16$, $\omega_\pm=1/2$, and the secondary null vector can be chosen to be

$$\vec{b}=\begin{pmatrix}1\\ 0\end{pmatrix} \quad\Rightarrow\quad M\vec{b}=\begin{pmatrix}1/2 & -1/8\\ 2 & -1/2\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}=\begin{pmatrix}1/2\\ 2\end{pmatrix}=\vec{a} \quad . \tag{42.93}$$

Putting everything together, we can now find the general solution in the 3 cases. We will immediately write things again in terms of proper time τ, related to the parameter t by $\tau = \exp t$.

1. For $\Delta \neq 0$ the general solution is

$$\begin{aligned} r(\tau) &= \tilde{c}_+\omega_+\tau^{\omega_+} + \tilde{c}_-\omega_-\tau^{\omega_-} \\ v(\tau) &= 2\tilde{c}_+\tau^{\omega_+} + 2\tilde{c}_-\tau^{\omega_-} \end{aligned} \tag{42.94}$$

   with $\tilde{c}_\pm$ real constants for $\Delta > 0$ and complex conjugates of each other for $\Delta < 0$. This agrees with the solution (42.9) with the identification $\tilde{c}_\pm\omega_\pm = c_\pm$.

2. For $\Delta = 0$ the general solution has the form (42.82),

$$\begin{pmatrix} r(\tau) \\ v(\tau) \end{pmatrix} = \tilde{c}\begin{pmatrix} 1/2 \\ 2 \end{pmatrix}\tau^{1/2} + \tilde{d}\left(\begin{pmatrix} 1/2 \\ 2 \end{pmatrix}\tau^{1/2}\log\tau + \begin{pmatrix} 1 \\ 0 \end{pmatrix}\tau^{1/2}\right) \quad . \tag{42.95}$$

   This agrees with the solution (42.15) with the identification $\tilde{c} = 2c - 4d$, $\tilde{d} = 2d$.

43 Exact Wave-like Solutions of the Einstein Equations

In section 23 we discussed wave-like solutions to the linearised Einstein equations. In this section, we will briefly discuss a class of solutions to the full non-linear Einstein equations which are also wave-like and thus generalise the solutions of that section (to which they reduce in the weak-field limit). These solutions are called *plane-fronted waves with parallel rays* or *pp-waves* for short. A special subset of these solutions are the so-called *exact gravitational plane wave metrics* or simply *plane waves.*

Such wave-metrics have been studied in the context of four-dimensional general relativity for a long time even though they are not (and were never meant to be) phenomenologically realistic models of gravitational plane waves. The reason for this is that in the far-field gravitational waves are so weak that the linearised Einstein equations and their solutions are adequate to describe the physics, whereas the near-field strong gravitational effects responsible for the production of gravitational waves, for which the linearised equations are indeed insufficient, correspond to much more complicated solutions of the Einstein equations (describing e.g. two very massive stars orbiting around their common center of mass).

However, pp-waves have been useful and of interest as a theoretical play-ground since they are in some sense the simplest essentially Lorentzian metrics with no non-trivial Riemannian counterparts. As such they also provide a wealth of counterexamples to conjectures that one might like to make about Lorentzian geometry by naive extrapolation from the Riemannian case. They have also enjoyed some popularity in the string theory literature as potentially exact and exactly solvable string theory “backgrounds”. However, they seem to have made it into very few textbook accounts of general relativity, and the purpose of this section is to at least partially fill this gap by providing a brief introduction to this topic.

43.1 Plane Waves in Rosen Coordinates: Heuristics

We have seen in section 23 that in the linearised approximation a metric describing the propagation of a plane wave in the x^3-direction (23.86) can be written as

$$ds^2 = -dt^2 + (\delta_{ab} + h_{ab})dx^a dx^b + (dx^3)^2 \quad , \tag{43.1}$$

with $h_{ab} = h_{ab}(t - x^3)$ “small”.

In terms of lightcone coordinates $U = z - t$, $V = (z + t)/2$ this can be written as

$$ds^2 = 2dUdV + (\delta_{ij} + h_{ij}(U))dy^i dy^j \quad . \tag{43.2}$$

We will now simply define a plane wave metric in general relativity to be a metric of the above form, dropping the assumption that h_{ij} be “small”,

$$d\bar{s}^2 = 2dUdV + \bar{g}_{ij}(U)dy^i dy^j \quad . \tag{43.3}$$

We will say that this is a plane wave metric in *Rosen coordinates*. This is not the coordinate system in which plane waves are usually discussed, among other reasons because typically in Rosen coordinates the metric exhibits spurious coordinate singularities. This led to the mistaken belief in the past that there are no non-singular plane wave solutions of the non-linear Einstein equations. We will establish the relation to the more common and much more useful *Brinkmann coordinates* below.

Plane wave metrics are characterised by a single matrix-valued function of U, but two metrics with quite different $\bar{g}_{ij}$ may well be isometric. For example,

$$d\bar{s}^2 = 2dUdV + U^2 d\vec{y}^2 \tag{43.4}$$

is isometric to the flat Minkowski metric whose natural presentation in Rosen coordinates is simply the Minkowski metric in lightcone coordinates,

$$d\bar{s}^2 = 2dUdV + d\vec{y}^2 \quad . \tag{43.5}$$

This is not too difficult to see, and we will establish this as a consequence of a more general result later on (but if you want to try this now, try scaling $\vec{y}$ by U and do something to V ...).

That (43.4) is indeed flat should in any case not be too surprising. It is the "null" counterpart of the "spacelike" fact that $ds^2 = dr^2 + r^2 d\Omega^2$, with $d\Omega^2$ the unit line element on the sphere, is just the flat Euclidean metric in polar coordinates, and the "timelike" statement that

$$ds^2 = -dt^2 + t^2 d\tilde{\Omega}^2 \quad , \tag{43.6}$$

with $d\tilde{\Omega}^2$ the unit line element on the hyperboloid, is just (a wedge of) the flat Minkowski metric. In cosmology this is known as the *Milne Universe* discussed in section 37.1, a rather trivial solution of the Friedmann equations with $k = -1$, $a(t) = t$ and $\rho = p = 0$.

It is somewhat less obvious, but still true, that for example the two metrics

$$\begin{aligned} d\bar{s}^2 &= 2dUdV + \sinh^2 U d\vec{y}^2 \\ d\bar{s}^2 &= 2dUdV + \mathrm{e}^{2U} d\vec{y}^2 \end{aligned} \tag{43.7}$$

are also isometric.

43.2 From pp-waves to plane waves in Brinkmann coordinates

In the remainder of this section we will study gravitational plane waves in a more systematic way. One of the characteristic features of the above plane wave metrics is the existence of a nowhere vanishing covariantly constant null vector field, namely ∂_V. We thus begin by deriving the general metric (line element) for a space-time admitting such a covariantly constant null vector field. We will from now on consider general $(d+2)$-dimensional space-times, where d is the number of transverse dimensions.

Thus, let Z be a parallel (i.e. covariantly constant) null vector of the $(d+2)$-dimensional Lorentzian metric $g_{\mu\nu}$, $\nabla_\mu Z^\nu = 0$. This condition is equivalent to the pair of conditions

$$\nabla_\mu Z_\nu + \nabla_\nu Z_\mu = 0 \tag{43.8}$$
$$\nabla_\mu Z_\nu - \nabla_\nu Z_\mu = 0 \ . \tag{43.9}$$

The first of these says that Z is a Killing vector field, and the second that Z is also a gradient vector field. If Z is nowhere zero, without loss of generality we can assume that

$$Z = \partial_v \tag{43.10}$$

for some coordinate v since this simply means that we are using a parameter along the integral curves of Z as our coordinate v. In terms of components this means that $Z^\mu = \delta^\mu_v$, or

$$Z_\mu = g_{\mu v} \ . \tag{43.11}$$

The fact that Z is null means that

$$Z_v = g_{vv} = 0 \ . \tag{43.12}$$

The Killing equation now implies that all the components of the metric are v-independent,

$$\partial_v g_{\mu\nu} = 0 \ . \tag{43.13}$$

The second condition (43.9) is identical to

$$\nabla_\mu Z_\nu - \nabla_\nu Z_\mu = 0 \ \Leftrightarrow \ \partial_\mu Z_\nu - \partial_\nu Z_\mu = 0 \ , \tag{43.14}$$

which implies that locally we can find a function $u = u(x^\mu)$ such that

$$Z_\mu = g_{v\mu} = \partial_\mu u \ . \tag{43.15}$$

There are no further constraints, and thus the general form of a metric admitting a parallel null vector is, changing from the x^μ-coordinates to $\{u, v, x^a\}$, $a = 1, \ldots, d$,

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= 2dudv + g_{uu}(u, x^c) du^2 + 2g_{au}(u, x^c) dx^a du + g_{ab}(u, x^c) dx^a dx^b \\ &\equiv 2dudv + K(u, x^c) du^2 + 2A_a(u, x^c) dx^a du + g_{ab}(u, x^c) dx^a dx^b \ . \end{aligned} \tag{43.16}$$

Note that if we had considered a metric with a covariantly constant timelike or spacelike vector, then we would have obtained the above metric with an additional term of the form $\mp dv^2$. In that case, the cross-term $2dudv$ could have been eliminated by shifting $v \to v' = v \mp u$, and the metric would have factorised into $\mp dv'^2$ plus a v'-independent metric. Such a factorisation does in general not occur for a covariantly constant null vector, which makes metrics with such a vector potentially more interesting than their timelike or spacelike counterparts.

There are still residual coordinate transformations which leave the above form of the metric invariant. For example, both K and A_a can be eliminated in favour of g_{ab}. We will not pursue this here, as we are primarily interested in a special class of metrics which are characterised by the fact that $g_{ab} = \delta_{ab}$,

$$ds^2 = 2dudv + K(u, x^b)du^2 + 2A_a(u, x^b)dx^a du + d\vec{x}^2 \quad . \tag{43.17}$$

Such metrics are called *plane-fronted waves with parallel rays*, or *pp-waves* for short. "plane-fronted" refers to the fact that the wave fronts $u = const.$ are planar (flat), and "parallel rays" refers to the existence of a parallel null vector. Once again, there are residual coordinate transformations which leave this form of the metric invariant. Among them are shifts of v, $v \to v + \Lambda(u, x^a)$, under which the coefficients K and A_a transform as

$$\begin{aligned} K &\to K + 2\partial_u \Lambda \\ A_a &\to A_a + \partial_a \Lambda \quad . \end{aligned} \tag{43.18}$$

Note in particular the "gauge transformation" of the (Kaluza-Klein) gauge field A_a, here associated with the null isometry generated by $Z = \partial_v$.

Plane waves are a very special kind of pp-waves. By definition, a plane wave metric is a pp-wave with $A_a = 0$ and $K(u, x^a)$ quadratic in the x^a (zero'th and first order terms in x^a can be eliminated by a coordinate transformation),

$$d\bar{s}^2 = 2dudv + A_{ab}(u)x^a x^b du^2 + d\vec{x}^2 \quad . \tag{43.19}$$

We will say that this is the metric of a plane wave in *Brinkmann coordinates*. The relation between the expressions for a plane wave in Brinkmann coordinates and Rosen coordinates will be explained in section 43.5. From now on barred quantities will refer to plane wave metrics.

Note that the metric of a plane wave in Brinkamnn coordinates is of the characteristic "flat + double-null" Kerr-Shild form (27.136) already briefly discussed in the context of the Schwarzschild metric in section 27.6. Indeed, we can write the components of the metric (43.19) as

$$\bar{g}_{\alpha\beta} = \eta_{\alpha\beta} + f(x) Z_\alpha Z_\beta \tag{43.20}$$

where

$$\eta_{\alpha\beta} dx^\alpha dx^\beta = 2dudv + d\vec{x}^2 \tag{43.21}$$

is the Minkowski metric in null coordinates, $f(x) = A_{ab}(u)x^a x^b$, and

$$Z_\alpha = \partial_\alpha u \tag{43.22}$$

is the characteristic (covariantly constant) null covector with respect to both the Minkowski metric and the plane wave metric.

In Brinkmann coordinates a plane wave metric is thus characterised by a single symmetric matrix-valued function $A_{ab}(u)$. Generically there is very little redundancy in the description of plane waves in Brinkmann coordinates, i.e. there are very few residual coordinate transformations that leave the form of the metric invariant, and the metric is specified almost uniquely by $A_{ab}(u)$. In particular, as we will see below, a plane wave metric is flat if and only if $A_{ab}(u) = 0$ identically. Contrast this with the non-uniqueness of the flat metric in Rosen coordinates. This uniqueness of the Brinkmann coordinates is one of the features that makes them convenient to work with in concrete applications.

43.3 Geodesics, Light-Cone Gauge and Harmonic Oscillators

We now take a look at geodesics of a plane wave metric in Brinkmann coordinates,

$$d\bar{s}^2 = 2dudv + A_{ab}(u)x^a x^b du^2 + d\vec{x}^2 \ , \tag{43.23}$$

i.e. the solutions $x^\mu(\tau)$ to the geodesic equations

$$\ddot{x}^\mu(\tau) + \bar{\Gamma}^\mu_{\nu\lambda}(x(\tau))\dot{x}^\nu(\tau)\dot{x}^\lambda(\tau) = 0 \ , \tag{43.24}$$

where an overdot denotes a derivative with respect to the affine parameter τ.

Rather than determining the geodesic equations by first calculating all the non-zero Christoffel symbols, we make use of the fact that the geodesic equations can be obtained more efficiently, and in a way that allows us to directly make use of the symmetries of the problem, as the Euler-Lagrange equations of the Lagrangian

$$\begin{aligned} \mathcal{L} &= \tfrac{1}{2}\bar{g}_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \\ &= \dot{u}\dot{v} + \tfrac{1}{2}A_{ab}(u)x^a x^b \dot{u}^2 + \tfrac{1}{2}\dot{\vec{x}}^2 \ , \end{aligned} \tag{43.25}$$

supplemented by the constraint

$$2\mathcal{L} = \epsilon \ , \tag{43.26}$$

where $\epsilon = 0$ ($\epsilon = -1$) for massless (massive) particles.

Since nothing depends on v, the *lightcone momentum*

$$p_v = \frac{\partial \mathcal{L}}{\partial \dot{v}} = \dot{u} \tag{43.27}$$

is conserved. For $p_v = 0$ the particle obviously does not feel the curvature and the geodesics are straight lines. A special case of this are the orbits of the parallel (and thus geodesic) Killing vector ∂_v. These are the rays of the gravitational wave and can also be regarded as the image of the Minkowski lightrays generated by $Z_\alpha = \partial_\alpha u$ under the Kerr-Schild map $\eta_{\alpha\beta} \to \bar{g}_{\alpha\beta}$ (43.20).

When $p_v \neq 0$, the case we will concentrate on henceforth, we choose the *lightcone gauge*

$$u = p_v \tau \ . \tag{43.28}$$

Then the geodesic equations for the transverse coordinates are the Euler-Lagrange equations

$$\ddot{x}^a(\tau) = A_{ab}(p_v\tau)x^b(\tau)p_v^2 \tag{43.29}$$

These are the equation of motion of a non-relativistic *harmonic oscillator*,

$$\ddot{x}^a(\tau) = -\omega^2_{ab}(\tau)x^b(\tau) \tag{43.30}$$

with (possibly time-dependent) frequency matrix

$$\omega^2_{ab}(\tau) = -p_v^2 A_{ab}(p_v\tau) \;\; , \tag{43.31}$$

The constraint

$$p_v\dot{v}(\tau) + \tfrac{1}{2}A_{ab}(p_v\tau)x^a(\tau)x^b(\tau)p_v^2 + \tfrac{1}{2}\dot{x}^a(\tau)\dot{x}^a(\tau) = 0 \tag{43.32}$$

for null geodesics (the case $\epsilon \neq 0$ can be dealt with in the same way) implies, and thus provides a first integral for, the v-equation of motion. Multiplying the oscillator equation by x^a and inserting this into the constraint, one finds that this can be further integrated to

$$p_v v(\tau) = -\tfrac{1}{2}x^a(\tau)\dot{x}^a(\tau) + p_v v_0 \;\; . \tag{43.33}$$

Note that a particular solution of the null geodesic equation is the purely "longitudinal" null geodesic

$$x^\mu(\tau) = (u = p_v\tau, v = v_0, x^a = 0) \;\; . \tag{43.34}$$

Along this null geodesic, all the Christoffel symbols of the metric (in Brinkmann coordinates) are zero. Hence Brinkmann coordinates can be regarded as a special case of *Fermi coordinates* (briefly mentioned at the beginning of section 3.6).

By definition the lightcone Hamiltonian is

$$H_{lc} = -p_u \;\; , \tag{43.35}$$

where p_u is the momentum conjugate to u in the gauge $u = p_v\tau$. With the above normalisation of the Lagrangian one has

$$\begin{aligned} p_u &= \bar{g}_{u\mu}\dot{x}^\mu = \dot{v} + A_{ab}(p_v\tau)x^a x^b p_v \\ &= -p_v^{-1}H_{ho}(\tau) \;\; , \end{aligned} \tag{43.36}$$

where $H_{ho}(\tau)$ is the (possibly time-dependent) harmonic oscillator Hamiltonian

$$H_{ho}(\tau) = \tfrac{1}{2}(\dot{x}^a\dot{x}^a - p_v^2 A_{ab}(p_v\tau)x^a x^b) \;\; . \tag{43.37}$$

Thus for the lightcone Hamiltonian one has

$$H_{lc} = \tfrac{1}{p_v}H_{ho} \;\; . \tag{43.38}$$

In summary, we note that in the lightcone gauge the equation of motion for a relativistic particle becomes that of a non-relativistic harmonic oscillator. This harmonic oscillator equation appears in various different contexts when discussing plane waves, and will therefore also reappear several times later on in this section.

43.4 Curvature and Singularities of Plane Waves

Due to the Kerr-Schild form (43.20) of the metric, curvature calculations greatly simplify, and it is easy to see that there is essentially only one non-vanishing component of the Riemann curvature tensor of a plane wave metric, namely

$$\bar{R}_{uaub} = -A_{ab} \quad . \tag{43.39}$$

In particular, therefore, because of the null (or chiral) structure of the metric, there is only one non-trivial component of the Ricci tensor,

$$\bar{R}_{uu} = -\delta^{ab}A_{ab} \equiv -\operatorname{Tr} A \quad , \tag{43.40}$$

the Ricci scalar is zero,

$$\bar{R} = 0 \quad , \tag{43.41}$$

and the only non-zero component of the Einstein tensor (8.108) is

$$\bar{G}_{uu} = \bar{R}_{uu} \quad . \tag{43.42}$$

Thus, as claimed above, the metric is flat iff $A_{ab} = 0$. Moreover, we see that in Brinkmann coordinates the vacuum Einstein equations reduce to a simple algebraic condition on A_{ab} (regardless of its u-dependence), namely that it be traceless.

A simple example of a vacuum plane wave metric in four dimensions is

$$d\bar{s}^2 = 2dudv + (x^2 - y^2)du^2 + dx^2 + dy^2 \quad , \tag{43.43}$$

or, more generally,

$$d\bar{s}^2 = 2dudv + [A(u)(x^2 - y^2) + 2B(u)xy]du^2 + dx^2 + dy^2 \tag{43.44}$$

for arbitrary functions $A(u)$ and $B(u)$. This reflects the two polarisation states or degrees of freedom of a four-dimensional graviton. Evidently, this generalises to arbitrary dimensions: the number of degrees of freedom of the traceless matrix $A_{ab}(u)$ correspond precisely to those of a transverse traceless symmetric tensor (a.k.a. a graviton).

The Weyl tensor is the traceless part of the Riemann tensor,

$$\bar{C}_{uaub} = -(A_{ab} - \frac{1}{d}\delta_{ab}\operatorname{Tr} A) \quad . \tag{43.45}$$

Thus the Weyl tensor vanishes (and, for $d > 1$, the plane wave metric is conformally flat) iff A_{ab} is pure trace (or isotropic in the transverse directions),

$$A_{ab}(u) = A(u)\delta_{ab} \quad . \tag{43.46}$$

For $d = 1$, every plane wave is conformally flat, as is most readily seen in Rosen coordinates.

When the Ricci tensor is non-zero (A_{ab} has non-vanishing trace), then plane waves solve the Einstein equations with null matter or null fluxes, i.e. with an energy-momentum tensor $\bar{T}_{\mu\nu}$ whose only non-vanishing component is $\bar{T}_{uu}$,

$$\bar{T}_{\mu\nu} = \rho(u)\delta_{\mu u}\delta_{\nu u} \quad . \tag{43.47}$$

Examples are e.g. null Maxwell fields $A_\mu(u)$ with field strength

$$F_{u\mu} = -F_{\mu u} = \partial_u A_\mu \quad . \tag{43.48}$$

Physical matter (with positive energy density) corresponds to $\bar{R}_{uu} > 0$ or $\mathrm{Tr}\, A < 0$.

It is pretty obvious by inspection that not just the scalar curvature but all the scalar curvature invariants of a plane wave, i.e. scalars built from the curvature tensor and its covariant derivatives, vanish since there is simply no way to soak up the u-indices.

Usually, an unambiguous way to ascertain that what appears to be a singularity of a metric is a true curvature singularity rather than just a singularity in the choice of coordinates is to exhibit a curvature invariant that is singular at that point. For example, for the Schwarzschild metric one has the Kretschmann scalar (27.163) $K = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} \sim m^2/r^6$, which shows that the singularity at $r = 0$ is a true singularity.

Now for plane waves all curvature invariants are zero. Does this mean that plane waves are non-singular? Or, if not, how does one detect the presence of a curvature singularity? One way to do this is to study the tidal forces acting on extended objects or families of freely falling particles. Indeed, in a certain sense the main effect of curvature (or gravity) is that initially parallel trajectories of freely falling non-interacting particles (dust, pebbles,...) do not remain parallel, i.e. that gravity has the tendency to focus (or defocus) matter. This statement find its mathematically precise formulation in the geodesic deviation equation (8.45),

$$(D_\tau)^2 \delta x^\mu = R^\mu_{\ \nu\lambda\rho}\dot{x}^\nu \dot{x}^\lambda \delta x^\rho \quad . \tag{43.49}$$

Here δx^μ is the separation vector between nearby geodesics. We can apply this equation to some family of geodesics of plane waves discussed in section 43.3. We will choose δx^μ to connect points on nearby geodesics with the same value of $\tau = u$. Thus $\delta u = 0$, and the geodesic deviation equation for the transverse separations δx^a reduces to

$$\frac{d^2}{du^2}\delta x^a = -\bar{R}^a_{\ ubu}\delta x^b = A_{ab}\delta x^b \quad . \tag{43.50}$$

This is (once again!) the harmonic oscillator equation, and generalises the corresponding equation (23.96) of the linearised theory to the present case.

We could have also obtained this directly by varying the harmonic oscillator (geodesic) equation for x^a, using $\delta u = 0$. We see that for negative eigenvalues of A_{ab} (physical matter) this tidal force is attractive, leading to a focussing of the geodesics. For vacuum plane waves, on the other hand, the tidal force is attractive in some directions and repulsive in the other (reflecting the quadrupole nature of gravitational waves).

What is of interest to us here is the fact that the above equation shows that A_{ab} itself contains direct physical information. In particular, these tidal forces become infinite where $A_{ab}(u)$ diverges. This is a true physical effect and hence the plane wave space-time is genuinely singular at such points.

Let us assume that such a singularity occurs at $u = u_0$. Since $u = p_v \tau$ is an affine parameter along the geodesic, this shows that any geodesic starting off at a finite value u_1 of u will reach the singularity in the finite "time" $u_0 - u_1$. Thus the space-time is geodesically incomplete and ends at $u = u_0$.

Since, on the other hand, the plane wave metric is clearly smooth for non-singular $A_{ab}(u)$, we can thus summarise this discussion by the statement that a plane wave is singular if and only if $A_{ab}(u)$ is singular somewhere.

43.5 From Rosen to Brinkmann coordinates (and back)

I still owe you an explanation of what the heuristic considerations of section 43.1 have to do with the rest of this section. To that end I will now describe the relation between the plane wave metric in Brinkmann coordinates,

$$d\bar{s}^2 = 2dudv + A_{ab}(u)x^a x^b du^2 + d\vec{x}^2 \quad , \tag{43.51}$$

and in Rosen coordinates,

$$d\bar{s}^2 = 2dUdV + \bar{g}_{ij}(U)dy^i dy^j \quad . \tag{43.52}$$

It is clear that, in order to transform the non-flat transverse metric in Rosen coordinates to the flat transverse metric in Brinkmann coordinates, one should change variables as

$$x^a = \bar{E}^a_i y^i \quad , \tag{43.53}$$

where $\bar{E}^a_i$ is a "vielbein" for $\bar{g}_{ij}$, i.e. it is a matrix which satisfies

$$\bar{g}_{ij} = \bar{E}^a_i \bar{E}^b_j \delta_{ab} \quad . \tag{43.54}$$

Denoting the inverse vielbein by $\bar{E}^i{}_a$, one has

$$\bar{g}_{ij}dy^i dy^j = (dx^a - \dot{\bar{E}}^a_i \bar{E}^i{}_c x^c dU)(dx^b - \dot{\bar{E}}^b{}_j \bar{E}^j{}_d x^d dU)\delta_{ab} \quad . \tag{43.55}$$

This generates the flat transverse metric as well as dU^2-term quadratic in the x^a, as desired, but there are also unwanted $dUdx^a$ cross-terms. Provided that $\bar{E}$ satisfies the symmetry condition

$$\dot{\bar{E}}_{ai}\bar{E}^i{}_b = \dot{\bar{E}}_{bi}\bar{E}^i{}_a \tag{43.56}$$

(such an $\bar{E}$ can always be found and is unique up to U-independent orthogonal transformations), these terms can be cancelled by a shift in V,

$$V \to V - \tfrac{1}{2}\dot{\bar{E}}_{ai}\bar{E}^i_{\ b}x^a x^b \quad . \tag{43.57}$$

Apart from eliminating the $dUdx^a$-terms, this shift will also have the effect of generating other dU^2-terms. Thanks to the symmetry condition, the term quadratic in first derivatives of $\bar{E}$ cancels that arising from $\bar{g}_{ij}dy^i dy^j$, and only a second-derivative part remains. The upshot of this is that after the change of variables

$$\begin{aligned} U &= u \\ V &= v + \tfrac{1}{2}\dot{\bar{E}}_{ai}\bar{E}^i_{\ b}x^a x^b \\ y^i &= \bar{E}^i_{\ a}x^a \quad , \end{aligned} \tag{43.58}$$

the metric (43.52) takes the Brinkmann form (43.51), with

$$A_{ab} = \ddot{\bar{E}}_{ai}\bar{E}^i_{\ b} \quad . \tag{43.59}$$

This can also be written as the *harmonic oscillator equation*

$$\ddot{\bar{E}}_{ai} = A_{ab}\bar{E}_{bi} \tag{43.60}$$

we had already encountered in the context of the geodesic (and geodesic deviation) equation.

Note that from this point of view the Rosen coordinates are labelled by d out of $2d$ linearly independent solutions of the oscillator equation, and the symmetry condition can now be read as the constraint that the *Wronskian* among these solutions be zero. Thus, given the metric in Brinkmann coordinates, one can construct the metric in Rosen coordinates by solving the oscillator equation, choosing a maximally commuting set of solutions to construct $\bar{E}_{ai}$, and then determining $\bar{g}_{ij}$ algebraically from the $\bar{E}_{ai}$.

In practice, once one knows that Rosen and Brinkmann coordinates are indeed just two distinct ways of describing the same class of metrics, one does not need to perform explicitly the coordinate transformation mapping one to the other. All one is interested in is the above relation between $\bar{g}_{ij}(U)$ and $A_{ab}(u)$, which essentially says that A_{ab} is the curvature of $\bar{g}_{ij}$,

$$A_{ab} = -\bar{E}^i_{\ a}\bar{E}^j_{\ b}\bar{R}_{UiUj} \quad . \tag{43.61}$$

The equations simplify somewhat when the metric $\bar{g}_{ij}(u)$ is diagonal,

$$\bar{g}_{ij}(u) = \bar{e}_i(u)^2\delta_{ij} \quad . \tag{43.62}$$

In that case one can choose $\bar{E}^a_{\ i} = \bar{e}_i\delta^a_i$. The symmetry condition is automatically satisfied because a diagonal matrix is symmetric, and one finds that A_{ab} is also diagonal,

$$A_{ab} = (\ddot{\bar{e}}_a/\bar{e}_a)\delta_{ab} \quad . \tag{43.63}$$

Conversely, therefore, given a diagonal plane wave in Brinkmann coordinates, to obtain the metric in Rosen coordinates one needs to solve the harmonic oscillator equations

$$\ddot{\bar{e}}_i(u) = A_{ii}(u)\bar{e}_i(u) \quad . \tag{43.64}$$

Thus the Rosen metric determined by $\bar{g}_{ij}(U)$ is flat iff $\bar{e}_i(u) = a_i U + b_i$ for some constants a_i, b_i. In particular, we recover the fact that the metric (43.4),

$$d\bar{s}^2 = 2dUdV + U^2 d\vec{y}^2 \tag{43.65}$$

is flat. We see that the non-uniqueness of the metric in Rosen coordinates is due to the integration 'constants' arising when trying to integrate a curvature tensor to a corresponding metric.

As another example, consider the four-dimensional vacuum plane wave (43.43). Evidently, one way of writing this metric in Rosen coordinates is

$$d\bar{s}^2 = 2dUdV + \sinh^2 U dX^2 + \sin^2 U dY^2 \quad , \tag{43.66}$$

and more generally any plane wave with constant A_{ab} can be chosen to be of this trigonometric form in Rosen coordinates.

43.6 More on Rosen Coordinates

Collecting the results of the previous sections, we can now gain a better understanding of the geometric significance (and shortcomings) of Rosen coordinates for plane waves.

First of all we observe that the metric

$$d\bar{s}^2 = 2dUdV + \bar{g}_{ij}(U)dy^i dy^j \tag{43.67}$$

defines a preferred family (congruence) of null geodesics, namely the integral curves of the null vector field ∂_U, i.e. the curves

$$(U(\tau), V(\tau), y^k(\tau)) = (\tau, V, y^k) \tag{43.68}$$

with affine parameter $\tau = U$ and parametrised by the constant values of the coordinates (V, y^k). In particular, the "origin" $V = y^k = 0$ of this congruence is the longitudinal null geodesic (43.34) with $v_0 = 0$ in Brinkmann coordinates.

In the region of validity of this coordinate system, there is a unique null geodesic of this congruence passing through any point, and one can therefore label (coordinatise) these points by specifying the geodesic (V, y^k) and the affine parameter U along that geodesic, i.e. by Rosen coordinates.

We can now also understand the reasons for the failure of Rosen coordinates: they cease to be well-defined (and give rise to spurious coordinate singularities) e.g. when geodesics in the family (congruence) of null geodesics interesect: in that case there is no longer a unique value of the coordinates (U, V, y^k) that one can associate to that intersection point.

To illustrate this point, consider simply $\mathbb{R}^2$ with its standard metric $ds^2 = dx^2 + dy^2$. An example of a "good" congruence of geodesics is the straight lines parallel to the x-axis. The corresponding "Rosen" coordinates ("Rosen" in quotes because we are not talking about null geodesics) are simply the globally well-defined Cartesian coordinates, x playing the role of the affine parameter U and y that of the transverse coordinates y^k labelling the geodesics. An example of a "bad" family of godesics is the straight lines through the origin. The corresponding "Rosen" coordinates are essentially just polar coordinates. Away from the origin there is again a unique geodesic passing through any point but, as is well known, this coordinate system breaks down at the origin.

With this in mind, we can now reconsider the "bad" Rosen coordinates

$$d\bar{s}^2 = 2dUdV + U^2 d\vec{y}^2 \tag{43.69}$$

for flat space. As we have seen above, in Brinkmann coordinates the metric is manifestly flat,

$$d\bar{s}^2 = 2dudv + d\vec{x}^2 \quad . \tag{43.70}$$

Using the coordinate transformation (43.58) from Rosen to Brinkmann coordinates, we see that the geodesic lines $y^k = c^k$, $V = c$ of the congruence defined by the metric (43.69) correspond to the lines $x^k = c^k u$ in Brinkmann (Minkowski) coordinates, but these are precisely the straight lines through the origin. This explains the coordinate singularity at $U = 0$ and further strengthens the analogy with polar coordinates mentioned at the end of section 43.1.

More generally, we see from (43.58) that the relation between the Brinkmann coordinates x^a and the Rosen coordinates y^k,

$$x^a = \bar{E}^a_k(U) y^k \quad , \tag{43.71}$$

and hence the expression for the geodesic lines $y^k = c^k$, becomes degenerate when $\bar{E}^a_k$ becomes degenerate, i.e. precisely when $\bar{g}_{ij}$ becomes degenerate. Brinkmann coordinates, on the other hand, provide a global coordinate chart for plane wave metrics.

The (almost) inevitablity of (coordinate) singularities in Rosen coordinates can be seen from the following argument.[196] Namely, it follows from the oscillator equation (43.60) that the determinant

$$\mathsf{E} = \det(\bar{E}^a_k) \tag{43.72}$$

satisfies

$$\ddot{\mathsf{E}}/\mathsf{E} = \mathrm{Tr}\, A + \left((\mathrm{Tr}\, M)^2 - \mathrm{Tr}(M^2)\right) \leq \mathrm{Tr}\, A = -\bar{R}_{uu} \quad , \tag{43.73}$$

[196]This is adapted from G. Gibbons, *Quantized Fields Propagating in Plane-Wave Spacetimes*, Commun. Math. Phys. 45 (1975) 191-202.

where use has been made of the expression $\bar{R}_{uu} = -\operatorname{Tr} A$ (43.40) for the Ricci tensor, and where M_{ab} is the symmetric matrix (43.56)

$$M_{ab} = \dot{\bar{E}}_{ai}\bar{E}^i_{\ b} \quad . \tag{43.74}$$

In particular, therefore, if $\bar{R}_{uu} > 0$, then $\mathsf{E}(u)$ is strictly concave downwards but positive at a non-degenerate point, so that necessarily $\mathsf{e}(u_0) = 0$ for some finite value of u_0, and the Rosen coordinate system breaks down there. By (43.47), $\bar{R}_{uu} > 0$ is equivalent to positivity of the lightcone energy density, a very reasonable requirement on the matter content.

43.7 Heisenberg Isometry Algebra of a Generic Plane Wave

We now study the isometries of a generic plane wave metric. In Brinkmann coordinates, because of the explicit dependence of the metric on u and the transverse coordinates, only one isometry is manifest, namely that generated by the parallel null vector $Z = \partial_v$. In Rosen coordinates, the metric depends neither on V nor on the transverse coordinates y^k, and one sees that in addition to $Z = \partial_V$ there are at least d more Killing vectors, namely the ∂_{y^k}. Together these form an Abelian translation algebra acting transitively on the null hypersurfaces of constant U.

However, this is not the whole story. Indeed, one particularly interesting and peculiar feature of plane wave space-times is the fact that they generically possess a *solvable* (rather than semi-simple) isometry algebra, namely a Heisenberg algebra, only part of which we have already seen above.

All Killing vectors V can be found in a systematic way by solving the Killing equations

$$L_V g_{\mu\nu} = \nabla_\mu V_\nu + \nabla_\nu V_\mu = 0 \quad . \tag{43.75}$$

I will not do this here but simply present the results of this analysis in Brinkmann coordinates. The upshot is that a generic $(2+d)$-dimensional plane wave metric has a $(2d+1)$-dimensional isometry algebra generated by the Killing vector $Z = \partial_v$ and the $2d$ Killing vectors

$$X(f_{(K)}) \equiv X_{(K)} = f_{(K)a}\partial_a - \dot{f}_{(K)a}x^a\partial_v \quad . \tag{43.76}$$

Here the $f_{(K)a}$, $K = 1, \ldots, 2d$ are the $2d$ linearly independent solutions of the *harmonic oscillator equation* (again!)

$$\ddot{f}_a(u) = A_{ab}(u)f_b(u) \quad . \tag{43.77}$$

These Killing vectors satisfy the algebra

$$[X_{(J)}, X_{(K)}] = W(f_{(J)}, f_{(K)})Z \tag{43.78}$$

$$[X_{(J)}, Z] = 0 \quad . \tag{43.79}$$

Here $W(f_{(J)}, f_{(K)})$, the *Wronskian* of the two solutions, is defined by

$$W(f_{(J)}, f_{(K)}) = \sum_a (\dot{f}_{(J)a} f_{(K)a} - \dot{f}_{(K)a} f_{(J)a}) \ . \tag{43.80}$$

It is constant (independent of u) as a consequence of the harmonic oscillator equation. Thus $W(f_{(J)}, f_{(K)})$ is a constant, non-degenerate, even-dimensional anti-symmetric matrix (non-degeneracy is implied by the linear independence of the solutions $f_{(J)}$.) Hence it can be put into standard (Darboux) form. Explicitly, a convenient choice of basis for the solutions $f_{(J)}$ is obtained by splitting the $f_{(J)}$ into two sets of solutions

$$\{f_{(J)}\} \to \{p_{(a)}, q_{(a)}\} \tag{43.81}$$

characterised by the initial conditions

$$\begin{aligned} p_{(a)b}(u_0) &= \delta_{ab} \quad \dot{p}_{(a)b}(u_0) = 0 \\ q_{(a)b}(u_0) &= 0 \quad\ \ \dot{q}_{(a)b}(u_0) = \delta_{ab} \ . \end{aligned} \tag{43.82}$$

Since the Wronskian of these functions is independent of u, it can be determined by evaluating it at $u = u_0$. Then one can immediately read off that

$$\begin{aligned} W(q_{(a)}, q_{(b)}) &= W(p_{(a)}, p_{(b)}) = 0 \\ W(q_{(a)}, p_{(b)}) &= \delta_{ab} \ . \end{aligned} \tag{43.83}$$

Therefore the corresponding Killing vectors

$$Q_{(a)} = X(q_{(a)}) \ , \qquad P_{(a)} = X(p_{(a)}) \tag{43.84}$$

and Z satisfy the canonically normalised Heisenberg algebra

$$\begin{aligned} &[Q_{(a)}, Z] = [P_{(a)}, Z] = 0 \\ &[Q_{(a)}, Q_{(b)}] = [P_{(a)}, P_{(b)}] = 0 \\ &[Q_{(a)}, P_{(b)}] = \delta_{ab} Z \ . \end{aligned} \tag{43.85}$$

43.8 Plane Waves with more Isometries

Generically, a plane wave metric has just this Heisenberg algebra of isometries. It acts transitively on the null hyperplanes $u = const.$, with a simply transitive Abelian subalgebra. However, for special choices of $A_{ab}(u)$, there may of course be more Killing vectors. These could arise from internal symmetries of A_{ab}, giving more Killing vectors in the transverse directions. For example, the conformally flat plane waves (43.46) have an additional $SO(d)$ symmetry (and conversely $SO(d)$-invariance implies conformal flatness).

Of more interest to us is the fact that for particular $A_{ab}(u)$ there may be Killing vectors with a ∂_u-component. The existence of such a Killing vector renders the plane wave homogeneous (away form the fixed points of this extra Killing vector). The obvious examples are plane waves with a u-independent profile A_{ab},

$$ds^2 = 2dudv + A_{ab}x^a x^b du^2 + d\vec{x}^2 \quad , \tag{43.86}$$

which have the extra Killing vector $X = \partial_u$. Since A_{ab} is u-independent, it can be diagonalised by a u-independent orthogonal transformation acting on the x^a. Moreover, the overall scale of A_{ab} can be changed, $A_{ab} \to \mu^2 A_{ab}$, by the coordinate transformation (boost)

$$(u, v, x^a) \to (\mu u, \mu^{-1} v, x^a) \quad . \tag{43.87}$$

Thus these metrics are classified by the eigenvalues of A_{ab} up to an overall scale and permutations of the eigenvalues.

Since A_{ab} is constant, the Riemann curvature tensor is covariantly constant,

$$\bar{\nabla}_\mu \bar{R}_{\lambda\nu\rho\sigma} = 0 \Leftrightarrow \partial_u A_{ab} = 0 \quad . \tag{43.88}$$

Thus a plane wave with constant wave profile A_{ab} is what is known as a locally symmetric space.

The existence of the additional Killing vector $X = \partial_u$ extends the Heisenberg algebra to the harmonic oscillator algebra, with X playing the role of the number operator or harmonic oscillator Hamiltonian. Indeed, X and $Z = \partial_v$ obviously commute, and the commutator of X with one of the Killing vectors $X(f)$ is

$$[X, X(f)] = X(\dot{f}) \quad . \tag{43.89}$$

Note that this is consistent, i.e. the right-hand side is again a Killing vector, because when A_{ab} is constant and f satisfies the harmonic oscillator equation then so does its u-derivative $\dot{f}$. In terms of the basis (43.84), we have

$$\begin{aligned} [X, Q_{(a)}] &= P_{(a)} \\ [X, P_{(a)}] &= A_{ab} Q_{(b)} \quad , \end{aligned} \tag{43.90}$$

which is the harmonic oscillator algebra.

Another way of understanding the relation between $X = \partial_u$ and the harmonic oscillator Hamiltonian is to look at the conserved charge associated with X for particles moving along geodesics. As we have seen in section 10.1, given any Killing vector X, the quantity

$$Q_X = X_\mu \dot{x}^\mu \tag{43.91}$$

is constant along the trajectory of the geodesic $x^\mu(\tau)$. For $X = \partial_u$ one finds

$$Q_X = p_u = g_{u\mu}\dot{x}^\mu \tag{43.92}$$

which we had already identified (up to a constant for non-null geodesics) as minus the harmonic oscillator Hamiltonian in section 43.3. This is indeed a conserved charge iff the Hamiltonian is time-independent i.e. iff A_{ab} is constant.

We thus see that the dynamics of particles in a symmetric plane wave background is intimately related to the geometry of the background itself.

Another class of examples of plane waves with an interesting additional Killing vector are plane waves with the non-trivial profile

$$A_{ab}(u) = u^{-2}B_{ab} \tag{43.93}$$

for some constant matrix $B_{ab} = A_{ab}(1)$. Without loss of generality one can then assume that B_{ab} and A_{ab} are diagonal, with eigenvalues the oscillator frequency squares $-\omega_a^2$,

$$A_{ab} = -\omega_a^2 \delta_{ab} u^{-2} \quad . \tag{43.94}$$

The corresponding plane wave metric

$$d\bar{s}^2 = 2dudv + B_{ab}x^a x^b \frac{du^2}{u^2} + d\vec{x}^2 \tag{43.95}$$

is invariant under the boost/scaling (43.87), corresponding to the extra Killing vector

$$X = u\partial_u - v\partial_v \quad . \tag{43.96}$$

Note that in this case the Killing vector $Z = \partial_v$ is no longer a central element of the isometry algebra, since it has a non-trivial commutator with X,

$$[X, Z] = Z \quad . \tag{43.97}$$

Moreover, one finds that the commutator of X with a Heisenberg algebra Killing vector $X(f)$, f_a a solution to the harmonic oscillator equation, is the Heisenberg algebra Killing vector

$$[X, X(f)] = X(u\dot{f}) \quad , \tag{43.98}$$

corresponding to the solution $u\dot{f}_a = u\partial_u f_a$ of the harmonic oscillator equation.

This concludes our brief discussion of plane wave metrics even though much more can and perhaps should be said about plane wave and pp-wave metrics, in particular in the context of the so-called *Penrose Limit* construction. For more on this see my lecture notes[197] (from which I also took the material in this chapter).

[197] M. Blau, *Lecture Notes on Plane Waves and Penrose Limits*, available from `http://www.blau.itp.unibe.ch/Lecturenotes.html`

44 Kaluza-Klein Theory

[Note: I have not taught, and hence not updated / corrected / improved, the material and the presentation of the material in this section in a long time and am not particularly happy with its current appearance.]

44.1 Motivation: Gravity and Gauge Theory

Looking at the Einstein equations and the variational principle, we see that gravity is nicely geometrised while the matter part has to be added by hand and is completely non-geometric. This may be perfectly acceptable for phenomenological Lagrangians (like that for a perfect fluid in cosmology), but it would clearly be desirable to have a unified description of all the fundamental forces of nature.

Today, the fundamental forces of nature are described by two very different concepts. On the one hand, we have - as we have seen - gravity, in which forces are replaced by geometry, and on the other hand there are the gauge theories of the electroweak and strong interactions (the standard model) or their (grand unified, ...) generalisations.

Thus, if one wants to unify these forces with gravity, there are two possibilities:

1. One can try to realise gravity as a gauge theory (and thus geometry as a consequence of the gauge principle).
2. Or one can try to realise gauge theories as gravity (and hence make them purely geometric).

The first is certainly an attractive idea and has attracted a lot of attention. It is also quite natural since, in a broad sense, gravity is already a gauge theory in the sense that it has a local invariance (under general coordinate transformations or, actively, diffeomorphisms). Also, the behaviour of Christoffel symbols under general coordinate transformations is analogous to the transformation behaviour of non-Abelian gauge fields under gauge transformations, and the whole formalism of covariant derivatives and curvatures is reminiscent of that of non-Abelian gauge theories.

At first sight, equating the Christoffel symbols with gauge fields (potentials) may appear to be a bit puzzling because we originally introduced the metric as the potential of the gravitational field and the Christoffel symbol as the corresponding field strength (representing the gravitational force). However, as we know, the concept of 'force' is itself a gauge (coordinate) dependent concept in General Relativity, and therefore these 'field strengths' behave more like gauge potentials themselves, with their curvature, the Riemann curvature tensor, encoding the gauge covariant information about the gravitational field. This fact, which reflects deep properties of gravity not shared by other forces, is just one of many which suggest that an honest gauge theory interpretation of gravity may be hard to come by. Let us nevertheless proceed in this direction for a little while anyway.

Clearly, the gauge group should now not be some 'internal' symmetry group like $U(1)$ or $SU(3)$, but rather a space-time symmetry group itself. Among the gauge groups that have been suggested in this context, one finds

1. the translation group (this is natural because, as we have seen, the generators of coordinate transformations are infinitesimal translations)
2. the Lorentz group (this is natural if one wants to view the Christoffel symbols as the analogues of the gauge fields of gravity)
3. and the Poincaré group (a combination of the two).

However, what - by and large - these investigations have shown is that the more one tries to make a gauge theory look like Einstein gravity the less it looks like a standard gauge theory and vice versa.

The main source of difference between gauge theory and gravity is the fact that in the case of Yang-Mills theory the internal indices bear no relation to the space-time indices whereas in gravity these are the same - contrast $F^a_{\mu\nu}$ with $(F^\lambda_\sigma)_{\mu\nu} = R^\lambda{}_{\sigma\mu\nu}$.

In particular, in gravity one can contract the 'internal' with the space-time indices to obtain a scalar Lagrangian, R, linear in the curvature tensor. This is fortunate because, from the point of view of the metric, this is already a two-derivative object.

For Yang-Mills theory, on the other hand, this is not possible, and in order to construct a Lagrangian which is a singlet under the gauge group one needs to contract the space-time and internal indices separately, i.e. one has a Lagrangian quadratic in the field stregths. This gives the usual two-derivative action for the gauge potentials.

In spite of these and other differences and difficulties, this approach has not been completely abandoned and the gauge theory point of view is still very fruitful and useful provided that one appreciates the crucial features that set gravity apart from standard gauge theories.

The second possibility alluded to above, to realise gauge theories as gravity, is much more radical, but how on earth is one supposed to achieve this? The crucial idea has been known since 1919/20 (T. Kaluza), with important contributions by O. Klein (1926). So what is this idea?

44.2 The Kaluza-Klein Miracle: History and Overview

In the early parts of the last century, the only other fundamental force that was known, in addition to gravity, was electro-magnetism, In 1919, Kaluza submitted a paper (to Einstein) in which he made a number of remarkable observations.

First of all, he stressed the similarity between Christoffel symbols and the Maxwell field strength tensor,

$$\begin{aligned} \Gamma_{\mu\nu\lambda} &= \tfrac{1}{2}(\partial_\nu g_{\mu\lambda} - \partial_\mu g_{\nu\lambda} + \partial_\lambda g_{\mu\nu}) \\ F_{\nu\mu} &= \partial_\nu A_\mu - \partial_\mu A_\nu \ . \end{aligned} \tag{44.1}$$

He then noted that $F_{\mu\nu}$ looks like a truncated Christoffel symbol and proposed, in order to make this more manifest, to introduce a *fifth dimension* with a metric such that $\Gamma_{\mu\nu 5} \sim F_{\mu\nu}$. This is inded possible. If one makes the identification

$$A_\mu = g_{\mu 5} \ , \tag{44.2}$$

and the assumption that $g_{\mu 5}$ is independent of the fifth coordinate x^5, then one finds, using the standard formula for the Christoffel symbols, now extended to five dimensions, that

$$\begin{aligned} \Gamma_{\mu\nu 5} &= \tfrac{1}{2}(\partial_5 g_{\mu\nu} + \partial_\nu g_{\mu 5} - \partial_\mu g_{\nu 5}) \\ &= \tfrac{1}{2}(\partial_\nu A_\mu - \partial_\mu A_\nu) = \tfrac{1}{2} F_{\nu\mu} \ . \end{aligned} \tag{44.3}$$

If this were all, this would not be particularly exciting, but much more than this is true. Kaluza went on to show that when one postulates a five-dimensional metric of the form (hatted quantities will from now on refer to five dimensional quantities)

$$d\hat{s}^2 = g_{\mu\nu} dx^\mu dx^\nu + (dx^5 + A_\mu dx^\mu)^2 \ , \tag{44.4}$$

and calculates the five-dimensional Einstein-Hilbert Lagrangian $\widehat{R}$, one finds *precisely the four-dimensional Einstein-Maxwell Lagrangian*

$$\widehat{R} = R - \tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} \ . \tag{44.5}$$

This fact is affectionately known as the *Kaluza-Klein Miracle*! Moreover, the five-dimensional geodesic equation turns into the four-dimensional Lorentz force equation for a charged particle, and in this sense gravity and Maxwell theory have really been unified in five-dimensional gravity.

However, although this is very nice, rather amazing in fact, and is clearly trying to tell us something deep, there are numerous problems with this and it is not really clear what has been achieved:

1. Should the fifth direction be treated as real or as a mere mathematical device?

2. If it is to be treated as real, why should one make the assumption that the fields are independent of x^5? If, on the other hand, one does not make this assumption, one will not get Einstein-Maxwell theory.

3. Moreover, if the fifth dimension is to be taken seriously, why are we justified in setting $g_{55} = 1$? If we do not do this, we will not get Einstein-Maxwell theory.

4. If the fifth dimension is real, why have we not discovered it yet?

In spite of all this and other questions, related to non-Abelian gauge symmetries or the quantum behaviour of these theories, Kaluza's idea has remained popular ever since or, rather, has periodically created psychological epidemics of frantic activity, interrupted by dormant phases. Today, Kaluza's idea, with its many reincarnations and variations, is an indispensable and fundamental ingredient in the modern theories of theoretical high energy physics (supergravity and string theories) and many of the questions/problems mentioned above have been addressed, understood and overcome.

Let us now look at this more precisely. We consider a five-dimensional space-time with coordinates $\widehat{x}^M = (x^\mu, x^5)$ and a metric of the form (44.4). For later convenience, we will introduce a parameter λ into the metric (even though we will set $\lambda = 1$ for the time being) and write it as

$$d\widehat{s}^2 = g_{\mu\nu}dx^\mu dx^\nu + (dx^5 + \lambda A_\mu dx^\mu)^2 \quad . \tag{44.6}$$

More explictly, we therefore have

$$\begin{aligned} \widehat{g}_{\mu\nu} &= g_{\mu\nu} + A_\mu A_\nu \\ \widehat{g}_{\mu 5} &= A_\mu \\ \widehat{g}_{55} &= 1 \quad . \end{aligned} \tag{44.7}$$

The determinant of the metric is $\widehat{g} = g$, and the inverse metric has components

$$\begin{aligned} \widehat{g}^{\mu\nu} &= g^{\mu\nu} \\ \widehat{g}^{\mu 5} &= -A^\mu \\ \widehat{g}^{55} &= 1 + A_\mu A^\mu \quad . \end{aligned} \tag{44.8}$$

We will (for now) assume that nothing depends on x^5 (in the old Kaluza-Klein literature this assumption is known as the *cylindricity condition*).

Introducing the notation

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \\ B_{\mu\nu} &= \partial_\mu A_\nu + \partial_\nu A_\mu \quad , \end{aligned} \tag{44.9}$$

the Christoffel symbols are readily found to be

$$
\begin{aligned}
\widehat{\Gamma}^{\mu}_{\nu\lambda} &= \Gamma^{\mu}_{\;\nu\lambda} - \tfrac{1}{2}(F^{\mu}_{\;\nu}A_{\lambda} + F^{\mu}_{\;\lambda}A_{\nu}) \\
\widehat{\Gamma}^{5}_{\nu\lambda} &= \tfrac{1}{2}B_{\nu\lambda} - \tfrac{1}{2}A^{\mu}(F_{\nu\mu}A_{\lambda} + F_{\lambda\mu}A_{\nu}) - A^{\mu}\Gamma_{\mu\nu\lambda} \\
\widehat{\Gamma}^{\mu}_{\;5\lambda} &= -\tfrac{1}{2}F^{\mu}_{\;\lambda} \\
\widehat{\Gamma}^{5}_{\;5\mu} &= -\tfrac{1}{2}F_{\mu\nu}A^{\nu} \\
\widehat{\Gamma}^{\mu}_{\;55} &= \widehat{\Gamma}^{5}_{\;55} = 0 \;\; .
\end{aligned} \tag{44.10}
$$

This does not look particularly encouraging, in particular because of the presence of the $B_{\mu\nu}$ term, but Kaluza was not discouraged and proceeded to calculate the Riemann tensor. I will spare you all the components of the Riemann tensor, but the Ricci tensor we need:

$$
\begin{aligned}
\widehat{R}_{\mu\nu} &= R_{\mu\nu} + \tfrac{1}{2}F_{\mu}^{\;\rho}F_{\rho\nu} + \tfrac{1}{4}F^{\lambda\rho}F_{\lambda\rho}A_{\mu}A_{\nu} + \tfrac{1}{2}(A_{\nu}\nabla_{\rho}F_{\mu}^{\;\rho} + A_{\mu}\nabla_{\rho}F_{\nu}^{\;\rho}) \\
\widehat{R}_{5\mu} &= +\tfrac{1}{2}\nabla_{\nu}F_{\mu}^{\;\nu} + \tfrac{1}{4}A_{\mu}F_{\nu\lambda}F^{\nu\lambda} \\
\widehat{R}_{55} &= \tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} \;\; .
\end{aligned} \tag{44.11}
$$

This looks a bit more attractive and covariant but still not very promising. [However, if you work in an orthonormal basis, as introduced for the Kaluza-Klein metric as an example in section 4.8, the result looks much nicer. In such a basis only the first two terms in $\widehat{R}_{\mu\nu}$ and the first term in $\widehat{R}_{5\mu}$ are present and $\widehat{R}_{55}$ is unchanged, so that all the non-covariant looking terms disappear.] Now the miracle happens. Calculating the curvature scalar, all the annoying terms drop out and one finds

$$\widehat{R} = R - \tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} \;\; , \tag{44.12}$$

i.e. the Lagrangian of Einstein-Maxwell theory. For $\lambda \neq 1$, the second term would have been multiplied by λ^2. We now consider the five-dimensional pure gravity Einstein-Hilbert action

$$\widehat{S} = \frac{1}{8\pi\widehat{G}}\int \sqrt{\widehat{g}}d^5x \; \widehat{R} \;\; . \tag{44.13}$$

In order for the integral over x^5 to converge we assume that the x^5-direction is a circle with radius L and we obtain

$$\widehat{S} = \frac{2\pi L}{8\pi\widehat{G}}\int \sqrt{g}d^4x \; (R - \tfrac{1}{4}\lambda^2 F_{\mu\nu}F^{\mu\nu}) \;\; . \tag{44.14}$$

Therefore, if we make the identifications

$$
\begin{aligned}
G_N &= \widehat{G}/2\pi L \\
\lambda^2 &= 8\pi G_N \;\; ,
\end{aligned} \tag{44.15}
$$

we obtain

$$\widehat{S} = \frac{1}{8\pi G_N}\int \sqrt{g}d^4x \; R - \frac{1}{4}\int \sqrt{g}d^4x \; F_{\mu\nu}F^{\mu\nu} \;\; , \tag{44.16}$$

i.e. precisely the four-dimensional Einstein-Maxwell Lagrangian! This amazing fact, that coupled gravity gauge theory systems can arise from higher-dimensional pure gravity, is certainly trying to tell us something.

44.3 ORIGIN OF GAUGE INVARIANCE

In physics, at least, miracles require a rational explanation. So let us try to understand on a priori grounds why the Kaluza-Klein miracle occurs. For this, let us recall Kaluza's ansatz for the line element (44.4),

$$d\widehat{s}^2_{KK} = g_{\mu\nu}(x^\lambda)dx^\mu dx^\nu + (dx^5 + A_\mu(x^\lambda)dx^\mu)^2 \quad . \tag{44.17}$$

and contrast this with the most general form of the line element in five dimensions, namely

$$\begin{aligned} d\widehat{s}^2 &= \widehat{g}_{MN}(x^L)dx^M dx^N \\ &= \widehat{g}_{\mu\nu}(x^\lambda, x^5)dx^\mu dx^\nu + 2\widehat{g}_{\mu 5}(x^\lambda, x^5)dx^\mu dx^5 + \widehat{g}_{55}(x^\mu, x^5)(dx^5)^2 \quad . \end{aligned} \tag{44.18}$$

Clearly, the form of the general five-dimensional line element (44.18) is invariant under arbitrary five-dimensional general coordinate transformations $x^M \to \xi^{M'}(x^N)$. This is not true, however, for the Kaluza-Klein ansatz (44.17), as a general x^5-dependent coordinate transformation would destroy the x^5-independence of $\widehat{g}_{\mu\nu} = g_{\mu\nu}$ and $\widehat{g}_{\mu 5} = A_\mu$ and would also not leave $\widehat{g}_{55} = 1$ invariant.

The form of the Kaluza-Klein line element is, however, invariant under the following two classes of coordinate transformations:

1. There are four-dimensional coordinate transformations

$$\begin{aligned} x^5 &\to x^5 \\ x^\mu &\to \xi^{\nu'}(x^\mu) \end{aligned} \tag{44.19}$$

   Under these transformations, as we know, $g_{\mu\nu}$ transforms in such a way that $g_{\mu\nu}dx^\mu dx^\nu$ is invariant, $A_\mu = \widehat{g}_{\mu 5}$ transforms as a four-dimensional covector, thus $A_\mu dx^\mu$ is invariant, and the whole metric is invariant.

2. There is also another remnant of five-dimensional general covariance, namely

$$\begin{aligned} x^5 &\to \xi^5(x^\mu, x^5) = x^5 + f(x^\mu) \\ x^\mu &\to \xi^\mu(x^\nu) = x^\mu \quad . \end{aligned} \tag{44.20}$$

   Under this transformation, $g_{\mu\nu}$ and g_{55} are invariant, but $A_\mu = g_{\mu 5}$ changes as

$$\begin{aligned} A'_\mu &= \widehat{g}'_{\mu 5} = \frac{\partial x^M}{\partial \xi^\mu}\frac{\partial x^N}{\partial \xi^5}\widehat{g}_{MN} \\ &= \frac{\partial x^M}{\partial x^\mu} g_{\mu 5} \\ &= g_{\mu 5} - \frac{\partial f}{\partial x^\mu} g_{55} \\ &= A_\mu - \partial_\mu f \quad . \end{aligned} \tag{44.21}$$

In other words, the Kaluza-Klein line element is invariant under the shift $x^5 \to x^5 + f(x^\mu)$ accompanied by $A_\mu \to A_\mu - \partial_\mu f$ (and this can of course also be read off directly from the metric).

This is precisely a *gauge transformation* of the vector potential A_μ and we see that in the present context *gauge transformations arise as remnants of five-dimensional general covariance*!

Now it is clear that we are guaranteed to get Einstein-Maxwell theory in four dimensions: First of all, upon integration over x^5, the shift in x^5 is irrelevant and starting with the five-dimensional Einstein-Hilbert action we are bound to end up with an action in four dimensions, depending on $g_{\mu\nu}$ and A_μ, which is

- generally covariant (in the four-dimensional sense),
- second order in derivatives,
- and invariant under gauge transformations of the A_μ.

Then the only possibility is the Einstein-Maxwell action.

A fruitful way of looking at the origin of this gauge invariance is as a consequence of the fact that constant shifts in x^5 are isometries of the metric, i.e. that $\partial/\partial x^5$ is a Killing vector of the metric (44.17). Then the isometry group of the 'internal' circle in the x^5-direction, namely $SO(2)$, becomes the gauge group $U(1) = SO(2)$ of the four-dimensional theory.

From this point of view, the gauge transformation of the vector potential arises from the Lie derivative of $\widehat{g}_{\mu 5}$ along the vector field $f(x^\mu)\partial_5$:

$$\begin{aligned} Y = f(x^\mu)\partial_5 \quad &\Rightarrow \quad Y^\mu = 0 \\ & \qquad\;\; Y^5 = f \\ &\Rightarrow \quad Y_\mu = A_\mu f \\ & \qquad\;\; Y_5 = f \quad . \end{aligned} \tag{44.22}$$

$$\begin{aligned} (L_Y \widehat{g})_{\mu 5} &= \widehat{\nabla}_\mu Y_5 + \widehat{\nabla}_5 Y_\mu \\ &= \partial_\mu Y_5 - 2\widehat{\Gamma}^\mu{}_{5M} Y^M \\ &= \partial_\mu f + F^\nu_{\;\mu} Y_\nu + F_{\mu\nu} A^\nu Y_5 \\ &= \partial_\mu f \\ &\Leftrightarrow \delta A_\mu = -\partial_\mu f \quad . \end{aligned} \tag{44.23}$$

This point of view becomes particularly useful when one wants to obtain non-Abelian gauge symmetries in this way (via a Kaluza-Klein reduction): One starts with a higher-dimensional internal space with isometry group G and makes an analogous ansatz for the metric. Then among the remnants of the higher-dimensional general coordinate transformations there are, in particular, x^μ-dependent 'isometries' of the internal metric. These act like non-Abelian gauge transformations on the off-block-diagonal componenents of the metric and, upon integration over the internal space, one is guaranteed to get, perhaps among other things, the four-dimensional Einstein-Hilbert and Yang-Mills actions.

44.4 Motion of Charged Particles in 4 Dimensions from 5-Dimensional Geodesics

There is something else that works very beautifully in this context, namely the description of the motion of charged particles in four dimensions moving under the combined influence of a gravitational and an electro-magnetic field. As we will see, also these two effects are unified from a five-dimensional Kaluza-Klein point of view.

Let us consider the five-dimensional geodesic equation

$$\ddot{x}^M + \widehat{\Gamma}^M_{NL}\dot{x}^N\dot{x}^L = 0 \quad . \tag{44.24}$$

Either because the metric (and hence the Lagrangian) does not depend on x^5, or because we know that $V = \partial_5$ is a Killing vector of the metric, we know that we have a conserved quantity

$$\frac{\partial\mathcal{L}}{\partial\dot{x}^5} \sim V_M\dot{x}^M = \dot{x}^5 + A_\mu\dot{x}^\mu \quad , \tag{44.25}$$

along the geodesic world lines. We will see in a moment what this quantity corresponds to. The remaining x^μ-component of the geodesic equation is

$$\begin{aligned}
\ddot{x}^\mu + \widehat{\Gamma}^\mu_{NL}\dot{x}^N\dot{x}^L &= \ddot{x}^\mu + \widehat{\Gamma}^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda \\
&+ 2\widehat{\Gamma}^\mu_{\nu 5}\dot{x}^\nu\dot{x}^5 + 2\widehat{\Gamma}^\mu_{55}\dot{x}^5\dot{x}^5 \\
&= \ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda - F^\mu_\nu A_\lambda\dot{x}^\nu\dot{x}^\lambda - F^\mu_\nu\dot{x}^\nu\dot{x}^5 \\
&= \ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda - F^\mu_\nu\dot{x}^\nu(A_\lambda\dot{x}^\lambda + \dot{x}^5) \quad .
\end{aligned} \tag{44.26}$$

Therefore this component of the geodesic equation is equivalent to

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda = (A_\lambda\dot{x}^\lambda + \dot{x}^5)F^\mu_\nu\dot{x}^\nu \quad . \tag{44.27}$$

This is precisely the Lorentz law if one identifies the constant of motion with the ratio of the charge and the mass of the particle,

$$\dot{x}^5 + A_\mu\dot{x}^\mu = \frac{e}{m} \quad . \tag{44.28}$$

Hence electro-magnetic and gravitational forces are indeed unified. The fact that charged particles take a different trajectory from neutral ones is not a violation of the equivalence principle but only reflects the fact that they started out with a different velocity in the x^5-direction!

The equations of motion of the four-dimensional Einstein-Hilbert-Maxwell action will of course give us the coupled Einstein-Maxwell equations

$$
\begin{aligned}
R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R &= 8\pi G_N T_{\mu\nu} \\
\nabla_\mu F^{\mu\nu} &= 0 \ .
\end{aligned}
\tag{44.29}
$$

However, let us now take a look at the equations of motion following from the five-dimensional Einstein-Hilbert action. These are, as we are looking at the vacuum equations, just the Ricci-flatness equations $\widehat{R}_{MN} = 0$. Looking back at (44.11) we see that these are clearly *not* equivalent to the Einstein-Maxwell equations. In particular, $\widehat{R}_{55} = 0$ imposes the constraint

$$
\widehat{R}_{55} = 0 \quad \Rightarrow \quad F_{\mu\nu} F^{\mu\nu} = 0 \ \ , \tag{44.30}
$$

and only then do the remaining equations $\widehat{R}_{\mu\nu} = 0$, $\widehat{R}_{\mu 5} = 0$ become equivalent to the Einstein-Maxwell equations (44.29).

What happened? Well, for one, taking variations and making a particular ansatz for the field configurations in the variational principle are two operations that in general do not commute. In particular, the Kaluza-Klein ansatz is special because it imposes the condition $g_{55} = 1$. Thus in four dimensions there is no equation of motion corresponding to $\widehat{g}_{55}$ whereas $\widehat{R}_{55} = 0$, the additional constraint, is just that, the equation arising from varying $\widehat{g}_{55}$. Thus Einstein-Maxwell theory is not a consistent truncation of five-dimensional General Relativity.

Now we really have to ask ourselves what we have actually achieved. We would like to claim that the five-dimensional Einstein-Hilbert action unifies the four-dimensional Einstein-Hilbert and Maxwell actions, but on the other hand we want to reject the five-dimensional Einstein equations? Then we are not ascribing any dynamics to the fifth dimension and are treating the Kaluza-Klein miracle as a mere kinematical, or mathematical, or bookkeeping device for the four-dimensional fields. This is clearly rather artificial and unsatisfactory.

There are some other unsatisfactory features as well in the theory we have developed so far. For instance we demanded that there be no dependence on x^5, which again makes the five-dimensional point of view look rather artificial. If one wants to take the fifth dimension seriously, one has to allow for an x^5-dependence of all the fields (and then explain later, perhaps, why we have not yet discovered the fifth dimension in every-day or high energy experiments).

With these issues in mind, we will now revisit the Kaluza-Klein ansatz, regarding the fifth dimension as real and exploring the consequences of this. Instead of considering directly the effect of a full (i.e. not restricted by any special ansatz for the metric) five-dimensional metric on four-dimensional physics, we will start with the simpler case of a free massless scalar field in five dimensions.

Let us assume that we have a five-dimensional space-time of the form $M_5 = M_4 \times S^1$ where we will at first assume that M_4 is Minkowski space and the metric is simply

$$d\widehat{s}^2 = -dt^2 + d\vec{x}^2 + (dx^5)^2 \quad , \tag{44.31}$$

with x^5 a coordinate on a circle with radius L. Now consider a massless scalar field $\widehat{\phi}$ on M_5, satisfying the five-dimensional massless Klein-Gordon equation

$$\widehat{\Box}\widehat{\phi}(x^\mu, x^5) = \widehat{\eta}^{MN}\partial_M\partial_N\widehat{\phi}(x^\mu, x^5) = 0 \quad . \tag{44.32}$$

As x^5 is periodic with period $2\pi L$, we can make a Fourier expansion of $\widehat{\phi}$ to make the x^5-dependence more explicit,

$$\widehat{\phi}(x^\mu, x^5) = \sum_n \phi_n(x^\mu)\mathrm{e}^{inx^5/L} \quad . \tag{44.33}$$

Plugging this expansion into the five-dimensional Klein-Gordon equation, we find that this turns into an infinite number of decoupled equations, one for each Fourier mode of ϕ_n of $\widehat{\phi}$, namely

$$(\Box - m_n^2)\phi_n = 0 \quad . \tag{44.34}$$

Here $\Box$ of course now refers to the four-dimensional d'Alembertian, and the mass term

$$m_n^2 = \frac{n^2}{L^2} \tag{44.35}$$

arises from the x^5-derivative ∂_5^2 in $\widehat{\Box}$.

Thus we see that, from a four-dimensional perspective, a massless scalar field in five dimensions give rise to one massless scalar field in four dimensions (the harmonic or constant mode on the internal space) and an infinite number of massive fields. The masses of these fields, known as the Kaluza-Klein modes, have the behaviour $m_n \sim n/L$. In general, this behaviour, an infinite tower of massive fields with mass $\sim 1/$ length scale is characteristic of massive fields arising from dimensional reduction from some higher dimensional space.

Next, instead of looking at a scalar field on Minkowski space-times a circle with the product metric, let us consider the Kaluza-Klein metric,

$$d\widehat{s}^2 = -dt^2 + d\vec{x}^2 + (dx^5 + \lambda A_\mu dx^\mu)^2 \quad , \tag{44.36}$$

and the corresponding Klein-Gordon equation

$$\widehat{\Box}\widehat{\phi}(x^\mu, x^5) = \widehat{g}^{MN}\widehat{\nabla}_M\partial_N\widehat{\phi}(x^\mu, x^5) = 0 \ . \tag{44.37}$$

Rather than spelling this out in terms of Christoffel symbols, it is more convenient to use (5.57) and recall that $\sqrt{\widehat{g}} = \sqrt{g} = 1$ to write this as

$$\begin{aligned} \widehat{\Box} &= \partial_M(\widehat{g}^{MN}\partial_N) \\ &= \partial_\mu\widehat{g}^{\mu\nu}\partial_\nu + \partial_5\widehat{g}^{5\mu}\partial_\mu + \partial_\mu\widehat{g}^{\mu 5}\partial_5 + \partial_5\widehat{g}^{55}\partial_5 \\ &= \eta^{\mu\nu}\partial_\mu\partial_\nu + \partial_5(-\lambda A^\mu\partial_\mu) + \partial_\mu(-\lambda A^\mu\partial_5) + (1+\lambda^2 A_\mu A^\mu)\partial_5\partial_5 \\ &= \eta^{\mu\nu}(\partial_\mu - \lambda A_\mu\partial_5)(\partial_\nu - \lambda A_\nu\partial_5) + (\partial_5)^2 \ . \end{aligned} \tag{44.38}$$

Acting with this operator on the Fourier decomposition of $\widehat{\phi}$, we evidently again get an infinite number of decoupled equations, one for each Fourier mode ϕ_n of $\widehat{\phi}$, namely

$$\left[\eta^{\mu\nu}(\partial_\mu - i\frac{\lambda n}{L}A_\mu)(\partial_\nu - i\frac{\lambda n}{L}A_\nu) - m_n^2\right]\phi_n = 0 \ . \tag{44.39}$$

This shows that the non-constant ($n \neq 0$) modes are not only massive but also charged under the gauge field A_μ. Comparing the operator

$$\partial_\mu - i\frac{\lambda n}{L}A_\mu \tag{44.40}$$

with the standard form of the minimal coupling,

$$\frac{\hbar}{i}\partial_\mu - eA_\mu \ , \tag{44.41}$$

we learn that the electric charge e_n of the n'th mode is given by

$$\frac{e_n}{\hbar} = \frac{n\lambda}{L} \ . \tag{44.42}$$

In particular, these charges are all integer multiples of a basic charge, $e_n = ne$, with

$$e = \frac{\hbar\lambda}{L} = \frac{\sqrt{8\pi G_N}\hbar}{L} \ . \tag{44.43}$$

Thus we get a formula for L, the radius of the fifth dimension,

$$L^2 = \frac{8\pi G_N\hbar^2}{e^2} = \frac{8\pi G_N\hbar}{e^2/\hbar} \ . \tag{44.44}$$

Restoring the velocity of light in this formula, and identifying the present $U(1)$ gauge symmetry with the standard gauge symmetry, we recognise here the fine structure constant

$$\alpha = e^2/4\pi\hbar c \approx 1/137 \ , \tag{44.45}$$

and the *Planck length*

$$\ell_P = \sqrt{\frac{G_N\hbar}{c^3}} \approx 10^{-33}\ \mathsf{cm} \ . \tag{44.46}$$

Thus

$$L^2 = \frac{2\ell_P^2}{\alpha} \approx 274\ell_P^2 \quad . \tag{44.47}$$

This is very small indeed, and it is therefore no surprise that this fifth dimension, if it is the origin of the $U(1)$ gauge invariance of the world we live in, has not yet been seen.

Another way of saying this is that the fact that L is so tiny implies that the masses m_n are huge, not far from the *Planck mass*

$$m_P = \sqrt{\frac{\hbar c}{G_N}} \approx 10^{-5}\ \mathsf{g} \approx 10^{19}\ \mathsf{GeV} \quad . \tag{44.48}$$

These would never have been spotted in present-day accelerators. Thus the massive modes are completely irrelevant for low-energy physics, the non-constant modes can be dropped, and this provides a justification for neglecting the x^5-dependence. However, this also means that the charged particles we know (electrons, protons, ...) cannot possibly be identified with these Kaluza-Klein modes.

The way modern Kaluza-Klein theories address this problem is by identifying the light charged particles we observe with the massless Kaluza-Klein modes. One then requires the standard spontaneous symmetry breaking mechanism to equip them with the small masses required by observation. This still leaves the question of how these particles should pick up a charge (as the zero modes are not only massless but also not charged). This is solved by going to higher dimensions, with non-Abelian gauge groups, for which massless particles are no longer necessarily singlets of the gauge group (they could e.g. live in the adjoint).

44.7 Kinematics of Dimensional Reduction

We have seen above that a massless scalar field in five dimensions gives rise to a massless scalar field plus an infinite tower of massive scalar fields in four dimensions. What happens for other fields (after all, we are ultimately interested in what happens to the five-dimensional metric)?

Consider, for example, a five-dimensional vector potential (covector field) $\widehat{B}_M(x^N)$. From a four-dimensional vantage point this looks like a four-dimensional vector field $B_\mu(x^\nu, x^5)$ and a scalar $\phi(x^\mu, x^5) = B_5(x^\mu, x^5)$. Fourier expanding, one will then obtain in four dimensions:

1. one massless Abelian gauge field $B_\mu(x^\nu)$
2. an infinite tower of massive charged vector fields
3. one massless scalar field $\phi(x^\mu) = B_5(x^\mu)$
4. an infinite tower of massive charged scalar fields

Retaining, for the same reasons as before, only the massless, i.e. x^5-independent, modes we therefore obtain a theory involving one scalar field and one Abelian vector field from pure Maxwell theory in five dimensions. The Lagrangian for these fields would be (dropping all x^5-derivatives)

$$\begin{aligned} F_{MN}F^{MN} &= F_{\mu\nu}F^{\mu\nu} + 2F_{\mu 5}F^{\mu 5} \\ &\to F_{\mu\nu}F^{\mu\nu} + 2(\partial_\mu \phi)(\partial^\mu \phi) \ . \end{aligned} \tag{44.49}$$

This procedure of obtaining Lagrangians in lower dimensions from Lagrangians in higher dimensions by simply dropping the dependence on the 'internal' coordinates is known as dimensional reduction or Kaluza-Klein reduction. However, the terminology is not uniform here - sometimes the latter term is used to indicate the reduction including all the massive modes. Also, in general 'massless' is not the same as 'x^5-independent', and then Kaluza-Klein reduction may refer to keeping the massless modes rather than the x^5-independent modes one retains in dimensional reduction.

Likewise, we can now consider what happens to the five-dimensional metric $\widehat{g}_{MN}(x^L)$. From a four-dimensional perspective, this splits into three different kinds of fields, namely a symmetric tensor $\widehat{g}_{\mu\nu}$, a covector $A_\mu = \widehat{g}_{\mu 5}$ and a scalar $\phi = \widehat{g}_{55}$. As before, these will each give rise to a massless field in four dimensions (which we interpret as the metric, a vector potential and a scalar field) as well as an infinite number of massive fields.

We see that, in addition to the massless fields we considered before, in the old Kaluza-Klein ansatz, we obtain one more massless field, namely the scalar field ϕ. Thus, even if we may be justified in dropping all the massive modes, we should keep this massless field in the ansatz for the metric and the action. With this in mind we now return to the Kaluza-Klein ansatz.

44.8 Kaluza-Klein Ansatz Revisited

Let us once again consider pure gravity in five dimensions, i.e. the Einstein-Hilbert action

$$\widehat{S} = \frac{1}{8\pi\widehat{G}} \int \sqrt{\widehat{g}} d^5x \ \widehat{R} \ . \tag{44.50}$$

Let us now parametrise the full five-dimensional metric as

$$d\widehat{s}^2 = \phi^{-1/3}[g_{\mu\nu}dx^\mu dx^\mu + \phi(dx^5 + \lambda A_\mu dx^\mu)^2] \ , \tag{44.51}$$

where all the fields depend on all the coordinates x^μ, x^5. Any five-dimensional metric can be written in this way and we can simply think of this as a change of variables

$$\widehat{g}_{MN} \to (g_{\mu\nu}, A_\mu, \phi) \ . \tag{44.52}$$

In matrix form, this metric reads

$$(\widehat{g}_{MN}) = \phi^{-1/3} \begin{pmatrix} g_{\mu\nu} + \lambda^2 \phi A_\mu A_\nu & \lambda\phi A_\nu \\ \lambda\phi A_\mu & \phi \end{pmatrix} \tag{44.53}$$

For a variety of reasons, this particular parametrisation is useful. In particular, it reduces to the Kaluza-Klein ansatz when $\phi = 1$ and all the fields are independent of x^5 and the ϕ's in the off-diagonal component ensure that the determinant of the metric is independent of the A_μ.

The only thing that may require some explanation is the strange overall power of ϕ. To see why this is a good choice, assume that the overall power is ϕ^a for some a. Then for $\sqrt{\widehat{g}}$ one finds

$$\sqrt{\widehat{g}} = \phi^{5a/2}\phi^{1/2}\sqrt{g} = \phi^{(5a+1)/2}\sqrt{g} \ . \tag{44.54}$$

On the other hand, for the Ricci tensor one has, schematically,

$$\hat{R}_{\mu\nu} = R_{\mu\nu} + \dots \ , \tag{44.55}$$

and therefore

$$\begin{aligned} \hat{R} &= \widehat{g}^{\mu\nu} R_{\mu\nu} + \dots \\ &= \phi^{-a} g^{\mu\nu} R_{\mu\nu} + \dots \\ &= \phi^{-a} R + \dots \ . \end{aligned} \tag{44.56}$$

Hence the five-dimensional Einstein-Hilbert action reduces to

$$\begin{aligned} \sqrt{\widehat{g}}\widehat{R} &\sim \phi^{(5a+1)/2}\phi^{-a}\sqrt{g}R + \dots \\ &= \phi^{(3a+1)/2}\sqrt{g}R + \dots \ . \end{aligned} \tag{44.57}$$

Thus, if one wants the five-dimensional Einstein-Hilbert action to reduce to the standard four-dimensional Einstein-Hilbert action (plus other things), without any non-minimal coupling of the scalar field ϕ to the metric, one needs to choose $a = -1/3$ which is the choice made in (44.51,44.53).

Making a Fourier-mode expansion of all the fields, plugging this into the Einstein-Hilbert action

$$\frac{1}{8\pi\widehat{G}} \int \sqrt{\widehat{g}} d^5x \ \hat{R} \ , \tag{44.58}$$

integrating over x^5 and retaining only the constant modes $g_{(0)\mu\nu}$, $A_{(0)\mu}$ and $\phi_{(0)}$, one obtains the action

$$S = \int \sqrt{g} d^4x \ \left[\frac{1}{8\pi G_N} R(g_{(0)\mu\nu}) - \frac{1}{4}\phi_{(0)} F_{(0)\mu\nu} F^{\mu\nu}_{(0)} - \frac{1}{48\pi G_N} \phi^{-2}_{(0)} g^{\mu\nu}_{(0)} \partial_\mu \phi_{(0)} \partial_\nu \phi_{(0)} \right] \ . \tag{44.59}$$

Here we have once again made the identifications (44.15). This action may not look as nice as before, but it is what it is. It is at least generally covariant and gauge invariant, as expected. We also see very clearly that it is inconsistenst with the equations of motion for $\phi_{(0)}$,

$$\Box \log \phi_{(0)} = \frac{3}{4} 8\pi G_N \phi_{(0)} F_{(0)\mu\nu} F^{\mu\nu}_{(0)} \ , \tag{44.60}$$

to set $\phi_{(0)} = 1$ as this would imply $F_{(0)\mu\nu}F^{\mu\nu}_{(0)} = 0$, in agreement with our earlier observations regarding $\hat{R}_{55} = 0$.

However, the configuration $g_{(0)\mu\nu} = \eta_{\mu\nu}$, $A_{(0)\mu} = 0$, $\phi_{(0)} = 1$ is a solution to the equations of motion and defines the 'vacuum' or ground state of the theory. From this point of view the zero mode metric, (44.53) with the fields replaced by their zero modes, i.e. the Kaluza-Klein ansatz with the inclusion of ϕ, has the following interpretation: as usual in quantum theory, once one has chosen a vacuum, one can consider fluctuations around that vacuum. The fields $g_{(0)\mu\nu}, A_{(0)\mu}, \phi_{(0)}$ are then the massless fluctuations around the vacuum and are the fields of the low-energy action. The full classical or quantum theory will also contain all the massive and charged Kaluza-Klein modes.

44.9 Non-Abelian Generalisation and Outlook

Even though in certain respects the Abelian theory we have discussed above is atypical, it is rather straightforward to generalise the previous considerations from Maxwell theory to Yang-Mills theory for an arbitrary non-Abelian gauge group. Of course, to achieve that, one needs to consider higher-dimensional internal spaces, i.e. gravity in $4+d$ dimensions, with a space-time of the form $M_4 \times M_d$. The crucial observation is that gauge symmetries in four dimensions arise from isometries (Killing vectors) of the metric on M_d.

Let the coordinates on M_d be x^a, denote by g_{ab} the metric on M_d, and let K_i^a, $i = 1,\ldots,n$ denote the n linearly independent Killing vectors of the metric g_{ab}. These generate the Lie algebra of the isometry group G via the Lie bracket

$$[K_i, K_j]^a \equiv K_i^b \partial_b K_j^a - K_j^b \partial_b K_i^a = f_{ij}^k K_k^a \ . \tag{44.61}$$

M_d could for example be the group manifold of the Lie group G itself, or a homogeneous space G/H for some subgroup $H \subset G$.

Now consider the following Kaluza-Klein ansatz for the metric,

$$d\hat{s}^2 = g_{\mu\nu}dx^\mu dx^\nu + g_{ab}(dx^a + K_i^a A_\mu^i dx^\mu)(dx^b + K_j^b A_\nu^j dx^\nu) \ . \tag{44.62}$$

Note the appearance of fields with the correct index structure to act as non-Abelian gauge fields for the gauge group G, namely the A_μ^i. Again these should be thought of as fluctuations of the metric around its 'ground state', $M_4 \times M_d$ with its product metric $(g_{\mu\nu}, g_{ab})$.

Now consider an infinitesimal coordinate transformation generated by the vector field

$$V^a(x^\mu, x^b) = f^i(x^\mu)K_i^a(x^b) \quad , \tag{44.63}$$

i.e.

$$\delta x^a = f^i(x^\mu)K_i^a(x^b) \quad . \tag{44.64}$$

This leaves the form of the metric invariant, and

$$\delta\widehat{g}_{\mu a} = L_V\widehat{g}_{\mu a} \tag{44.65}$$

can be seen to imply

$$\delta A_\mu^i = D_\mu f^i \equiv \partial_\mu f^i - f^i_{jk}A^j_\mu f^k \quad , \tag{44.66}$$

i.e. precisely an infinitesimal non-Abelian gauge transformation. The easiest way to see this is to use the form of the Lie derivative not in its covariant form,

$$L_V\widehat{g}_{\mu a} = \widehat{\nabla}_\mu V_a + \widehat{\nabla}_a V_\mu \tag{44.67}$$

(which requires knowledge of the Christoffel symbols) but in the form

$$L_V\widehat{g}_{\mu a} = V^c\partial_c\widehat{g}_{\mu a} + \partial_\mu V^c\widehat{g}_{ca} + \partial_a V^c\widehat{g}_{\mu c} \quad . \tag{44.68}$$

Inserting the definitions of $\widehat{g}_{\mu a}$ and V^a, using the fact that the K_i^a are Killing vectors of the metric g_{ab} and the relation (44.61), one finds

$$L_V\widehat{g}_{\mu a} = g_{ab}K_i^b D_\mu f^i \quad , \tag{44.69}$$

and hence (44.66).

One is then assured to find a Yang-Mills like term

$$L_{YM} \sim F^i_{\mu\nu}F^{j\,\mu\nu}K_i^a K_j^b g_{ab} \tag{44.70}$$

in the reduction of the Lagrangian from $4+d$ to 4 dimensions.

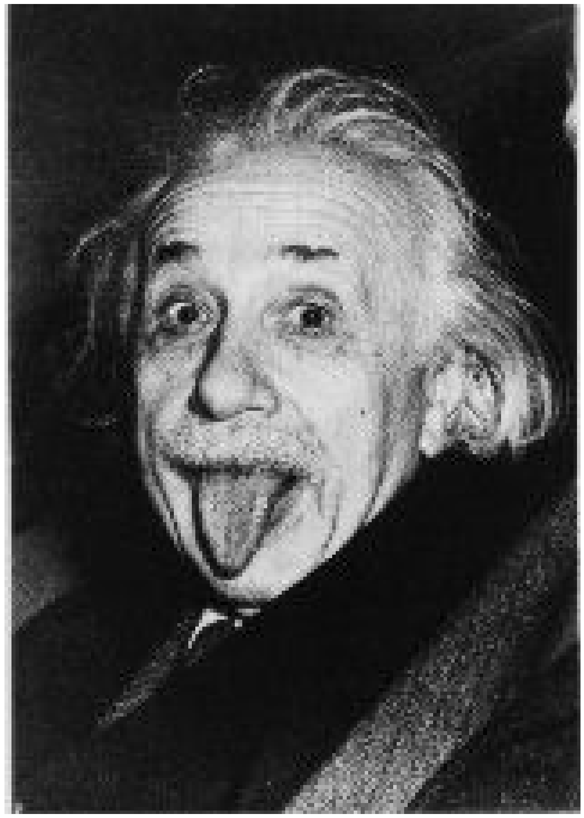
The problem with this scenario (already prior to worrying about the inclusion of scalar fields, of which there will be plenty in this case, one for each component of g_{ab}) is that the four-dimensional space-time cannot be chosen to be flat. Rather, it must have a huge cosmological constant. This arises because the dimensional reduction of the $(4+d)$-dimensional Einstein-Hilbert Lagrangian $\hat{R}$ will also include a contribution from the scalar curvature R_d of the metric on M_d. For a compact internal space with non-Abelian isometries this scalar curvature is non-zero and will therefore lead to an effective cosmological constant in the four-dimensional action. This cosmological constant could be cancelled 'by hand' by introducing an appropriate cosmological constant of the opposite sign into the $(d+4)$-dimensional Einstein-Hilbert action, but this looks rather contrived and artificial.

Nevertheless, this and other problems have not stopped people from looking for 'realistic' Kaluza-Klein theories giving rise to the standard model gauge group in four dimension. Of course, in order to get the standard model action or something resembling it, fermions need to be added to the $(d+4)$-dimensional action.

An interesting observation in this regard is that the lowest possible dimension for a homogenous space with isometry group $G = SU(3) \times SU(2) \times U(1)$ is seven, so that the dimension of space-time is eleven. This arises because the maximal compact subgroup H of G, giving rise to the smallest dimensional homogeneous space G/H of G, is $SU(2) \times U(1) \times U(1)$. As the dimension of G is $8+3+1 = 12$ and that of H is $3+1+1 = 5$, the dimension of G/H is $12 - 5 = 7$. This is intriguing because eleven is also the highest dimension in which supergravity exists (in higher dimensions, supersymmetry would require the existence of spin > 2 particles). That, plus the hope that supergravity would have a better quantum behaviour than ordinary gravity, led to an enourmous amount of activity on Kaluza-Klein supergravity in the early 80's.

Unfortunately, it turned out that not only was supergravity sick at the quantum level as well but also that it is impossible to get a chiral fermion spectrum in four dimensions from pure gravity plus spinors in $(4+d)$ dimensions. One way around the latter problem is to include explicit Yang-Mills fields already in $(d+4)$-dimensions, but that appeared to defy the purpose of the whole Kaluza-Klein idea.

Today, the picture has changed and supergravity is regarded as a low-energy approximation to string theory which is believed to give a consistent description of quantum gravity. These string theories typically live in ten dimensions, and thus one needs to 'compactify' the theory on a small internal six-dimensional space, much as in the Kaluza-Klein idea. Even though non-Abelian gauge fields now typically do not arise from Kaluza-Klein reduction but rather from explicit gauge fields in ten dimensions (or objects called *D-branes*), in all other respects Kaluza's old idea is alive, doing very well, and an indispensable part of the toolkit of modern theoretical high energy physics.



THE END