Fourier Analysis and Related Topics

J. Korevaar

Preface

For many years, the author taught a one-year course called "Mathematical Methods". It was intended for beginning graduate students in the physical sciences and engineering, as well as for mathematics students with an interest in applications. The aim was to provide mathematical tools used in applications, and a certain theoretical background that would make other parts of mathematical analysis accessible to the student of physical science. The course was taken by a large number of students at the University of Wisconsin (Madison), the University of California San Diego (La Jolla), and finally, the University of Amsterdam. At one time the author planned to turn his elaborate lecture notes into a multi-volume book, but only one volume appeared [**68**]. The material in the present book represents a selection from the lecture notes, with emphasis on Fourier theory. Starting with the classical theory for well-behaved functions, and passing through L^1 and L^2 theory, it culminates in distributional theory, with applications to bounday value problems.

At the International Congress of Mathematicians (Cambridge, Mass) in 1950, many people became interested in the Generalized Functions or "Distributions" of field medallist Laurent Schwartz; cf. [**110**]. Right after the congress, Michael Golomb, Merrill Shanks and the author organized a yearlong seminar at Purdue University to study Schwartz's work. The seminar led the author to a more concrete approach to distributions [**66**], which he included in applied mathematics courses at the University of Wisconsin. (The innovation was recognized by a Reynolds award in 1956.)

It took the mathematical community a while to agree that distributions were useful. This happened only when the theory led to major new developments; see the five books on generalized functions by Gelfand and coauthors [**37**], and especially the four volumes by Hörmander [**52**] on partial differential equations.

A detailed description of the now classical material in the present textbook may be found in the introductions to the various chapters. The survey in Chapter 1 mentions work of Euler and Daniel Bernoulli, which preceded the elaborate work of Fourier related to the heat equation. Dirichlet's rigorous convergence theory for Fourier series of "good" functions is covered in Chapter 2. The possible divergence in the case of continuous functions is treated, as well as the remarkable Gibbs phenomenon. Chapter 3 shows how such problems were overcome around 1900 by the use of summability methods, notably by Fejér. Soon thereafter, the notion of square integrable functions in the sense of Lebesgue would lead to an elegant treatment of Fourier series as orthogonal series. However, even summability methods and L^2 theory were not general enough to satisfy the needs of applications. Many of these needs were finally met by Schwartz's distributional theory (Chapter 4). The classical restrictions on many operations, such as differentiation and termwise integration or differentiation of infinite series, could be removed.

After some general results on metric and normed spaces, including a construction of completion, Chapter 5 discusses inner product spaces and Hilbert spaces. It thus provides the theoretical setting for a good treatment of general orthogonal series and orthogonal bases (Chapter 6). Chapter 7 is devoted to important classical orthogonal systems such as the Legendre polynomials and the Hermite functions. Most of these orthogonal systems arise also as systems of eigenfunctions of Sturm–Liouville eigenvalue problems for differential operators, as shown in Chapter 8. That chapter ends with results on Laplace's equation (Dirichlet problem) and spherical harmonics. Chapter 9 treats Fourier transformation for well-behaved integrable functions on $\mathbb{R}$. Among the well-behaved functions the Hermite functions stand out; here they appear as eigenfunctions of the linear harmonic oscillator in quantum mechanics.

At this stage the student should be well-prepared for a general theory of Fourier integrals. The basic questions are to represent larger or unruly functions by trigonometric integrals, and to make Fourier inversion widely possible. A convenient class to work with are the so-called tempered distributions, which include all functions of at most polynomial growth, as well as their (generalized) derivatives of arbitrary order. A good starting point to prove unlimited inversion is the observation that the Fourier transform operator $\mathcal{F}$ commutes with the Hermite operator $\mathcal{H} = x^2 - D^2$, where D stands for differentiation, d/dx. It follows that the two operators have the same eigenfunctions. Now the normalized eigenfunctions of $\mathcal{H}$ are the Hermite functions h_n, which form an orthonormal basis of L^2. Tempered distributions also have a unique representation $\sum c_n h_n$; see Chapter 10. The (normalized) Fourier operator $\mathcal{F}$ transforms the series $\sum c_n h_n$ into $\sum (-i)^n c_n h_n$, while the reflected Fourier operator $\mathcal{F}_R$ multiplies the expansion coefficients by i^n. Thus $\mathcal{F}$ is inverted by $\mathcal{F}_R$; cf. Chapter 11. For L^2 this approach goes back to Wiener [**124**]. [The author has used Hermite series to extend Fourier theory to a much larger class of generalized functions than tempered distributions; see [**67**], and cf. Zhang [**126**].]

Chapter 12 first deals with one-sided integral transforms such as the Laplace transform, which are important for initial value problems. Next come multiple Fourier transforms. The most important application is to so-called fundamental solutions of certain partial differential equations. In the case of the wave equation one thus obtains the response to a sharply time-limited signal at time zero at the origin. As a striking result one finds that communication governed by that equation works poorly in even dimensions, and works really well only in $\mathbb{R}^3$!

The short final Chapter 13 sketches the theory of general Schwartz distributions and two-sided Laplace transforms.

Acknowledgements. Thanks are due to University of Amsterdam colleague Jan van de Craats, who converted my sketches into the nice figures in the text. I also thank former and present colleagues who have encouraged me to write the present "Mathematical Methods" book. Last but not least, I thank the many students who have contributed to the exposition by their questions and comments; it was a pleasure to work with them! Both categories come together in Jan Wiegerinck, who also became a good friend, and director of the Korteweg–de Vries Institute for Mathematics, a nice place to work.

Amsterdam, Spring, 2011 Jaap Korevaar

Contents

CHAPTER 1

Introduction and survey

Trigonometric series began to play a role in mathematics through the work of the Swiss mathematicians Leonhard Euler (1707–1783, St. Petersburg, Berlin; [**29**]) and Daniel Bernoulli (1700–1782, Basel; [**6**]). Systematic applications of trigonometric series and integrals to problems of *mathematical physics* were made by Joseph Fourier (1768–1830, Paris, "Théorie analytique de la chaleur", 1822; [**33**]). A first rigorous *convergence theory* for Fourier series was developed by Johann P.G.L. Dirichlet (1805–1859, Germany; [**25**]). It applied to "good" periodic functions, for example, piecewise monotonic functions. Later, it was discovered that there are rapidly oscillating *continuous* functions whose Fourier series do not converge in the ordinary sense. However, Lipót Fejér (1880–1959, Budapest; [**30**]) could show that there is a *summability method* that reproduces *every* continuous function from its Fourier series (1904). A little later, with the introduction of the Lebesgue integral, there arose a beautiful theory of Fourier series as *orthogonal series*. Even this theory was not general enough to satisfy the needs of applications. Around 1945, Laurent Schwartz (1915–2002, France; [**109**]) introduced a powerful theory of Fourier series and integrals based on his so-called *distributions* or generalized functions.

There are many books on Fourier analysis, see the Internet; a few are mentioned at the end of the chapter.

1.1. Power series and trigonometric series

Trigonometric series arise when we consider a power series or Laurent series $\sum c_n z^n$ on a circle $x = re^{it}$, $-\pi < t \leq \pi$.

Example 1.1.1. In Complex Analysis one encounters the principal value (p.v.) of the logarithm of a complex number $w \neq 0$:

$$\text{p.v.} \log w \stackrel{\text{def}}{=} \log |w| + i \, \text{p.v.} \arg w,$$

where the principal value of the argument is $> -\pi$ and $\leq +\pi$. This formula defines an analytic function outside the (closed) negative real axis with derivative $1/w$. For $w = 1 + z$ with $|z| < 1$ one may represent the principal value by an integral along the segment from 0 to z, and hence by a power series:

$$\begin{aligned} \text{p.v. } \log(1+z) &= \int_0^z \frac{ds}{1+s} = \int_0^z (1 - s + s^2 - s^3 + \cdots)\, ds \\ &= z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \frac{1}{4}z^4 + \cdots . \end{aligned}$$

Setting $z = re^{it}$ and letting $r \nearrow 1$, one *formally* [that is, without regard to convergence] obtains

$$\begin{aligned} \text{p.v. } \log(1+e^{it}) &= \log\left|1+e^{it}\right| + i\,\text{p.v. } \arg\left(1+e^{it}\right) \\ &= \log\left|2\cos\frac{1}{2}\,t\right| + i\,\frac{1}{2}\,t \\ &= e^{it} - \frac{1}{2}\,e^{2it} + \frac{1}{3}\,e^{3it} - \frac{1}{4}\,e^{4it} + \cdots , \quad |t| < \pi. \end{aligned} \tag{1.1.1}$$

Assuming that the series in (1.1.1) *is* convergent, and then separating real and imaginary parts, one finds that

$$\log\left|2\cos\frac{1}{2}t\right| = \cos t - \frac{1}{2}\cos 2t + \frac{1}{3}\cos 3t - \frac{1}{4}\cos 4t + \cdots , \tag{1.1.2}$$

$$\frac{1}{2}\,t = \sin t - \frac{1}{2}\sin 2t + \frac{1}{3}\sin 3t - \frac{1}{4}\sin 4t + \cdots , \quad |t| < \pi. \tag{1.1.3}$$

Are these manipulations permitted? A *continuity theorem* of Niels H. Abel (Norway, 1802–1829; [**1**]) will be helpful.

THEOREM 1.1.2. *Let $f(z) = \sum_{n=0}^\infty c_n z^n$ for $|z| < R$ and suppose that the power series converges at the point z_0 on the circle $C(0,R) = \{|z| = R\}$. Then the sum of the series at the point z_0 can be obtained as a radial limit:*

$$\sum_{n=0}^\infty c_n z_0^n = \lim_{r \nearrow 1} f(rz_0).$$

With this theorem the question of the validity of (1.1.1) is reduced to the question whether the series

$$\sum_{n=1}^\infty \frac{(-1)^{n-1}}{n}\, e^{int} \tag{1.1.4}$$

is convergent. Since we do not have absolute convergence, this is a delicate matter. Here one can use *partial summation*:

LEMMA 1.1.3. (i) *For complex numbers a_n, b_n, $n \in \mathbb{N}$ and the partial sums $A_n = a_1 + a_2 + \cdots + a_n$* [with $A_0 = 0$] *one has*

$$\sum_{n=j+1}^{k} a_n b_n = \sum_{n=j}^{k-1} A_n(b_n - b_{n+1}) + A_k b_k - A_j b_j \qquad (k > j).$$

(ii) *If $|A_n| \le M < \infty$ for all n and $b_n \searrow 0$* (monotonicity!), *then the infinite series $\sum_{n=1}^{\infty} a_n b_n$ is convergent, and*

$$\sum_{n=j+1}^{\infty} a_n b_n = \sum_{n=j}^{\infty} A_n(b_n - b_{n+1}) - A_j b_j.$$

Application to the series in (1.1.1). Take $a_n = (-1)^{n-1}e^{int}$, $b_n = \frac{1}{n}$. Then

$$A_n = e^{it} - e^{2it} + \cdots + (-1)^{n-1}e^{int} = e^{it}\,\frac{1-(-e^{it})^n}{1-(-e^{it})},$$

so that

$$|A_n| \le \frac{2}{|1+e^{it}|} = \frac{1}{|\cos\frac{1}{2}t|}. \tag{1.1.5}$$

Thus by Lemma 1.1.3, the series (1.1.4) converges for $|t| < \pi$. The sum of the series in (1.1.1) can now be obtained from Abel's theorem:

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\,e^{int} &= \lim_{r\nearrow 1}\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\,r^n e^{int}\\ &= \lim_{r\nearrow 1}\text{p.v. }\log\left(1+re^{it}\right) = \text{p.v. }\log\left(1+e^{it}\right), \quad |t| < \pi.\end{aligned}$$

Exercises 1.1.1. Verify Lemma 1.1.3.

1.1.2. Use Lemma 1.1.3 to prove Theorerm 1.1.2.

Hint. One may take $R = 1$ and $z_0 = 1$; by changing c_0 one may also suppose that $\sum_0^\infty c_n = 0$.

1.1.3. Use formula (1.1.1) to calculate the sum $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$.

1.1.4. Compute the sums of the series

$$\sum_1^\infty \frac{\cos nx}{n} \quad \text{and} \quad \sum_1^\infty \frac{\sin nx}{n},$$

first for $0 < x < 2\pi$, and next for general $x \in \mathbb{R}$. Sketch the graphs of the sum functions.

1.1.5. What do you think of Euler's formulas

$$\sum_{n=-\infty}^{\infty} e^{inx} = 0 \quad \text{for } 0 < x < 2\pi; \quad 1 - 2 + 2^2 - 2^3 + \cdots = \frac{1}{3}\,?$$

1.2. New series by integration or differentiation

EXAMPLE 1.2.1. Formal termwise integration of the series for $\frac{1}{2}t$ in formula (1.1.3) gives

$$-\cos t + \frac{1}{2^2}\cos 2t - \frac{1}{3^2}\cos 3t + \frac{1}{4^2}\cos 4t - \cdots = \frac{1}{4}t^2 + C. \tag{1.2.1}$$

Would this be correct for $|t| < \pi$? Perhaps even for $|t| \le \pi$? If so, we can evaluate C and also

$$S = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots, \tag{1.2.2}$$

simply by setting $t = 0$ and $t = \pi$:

$$C = -1 + \frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{4^2} - \cdots,$$

$$S = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{1}{4}\pi^2 + C. \tag{1.2.3}$$

Indeed, addition would give

$$C + S = \frac{2}{2^2} + \frac{2}{4^2} + \frac{2}{6^2} + \cdots = \frac{2}{2^2}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots\right) = \frac{1}{2}S,$$

so that $S = -2C$, and hence by (1.2.3),

$$C = -\frac{1}{12}\pi^2, \quad S = \frac{1}{6}\pi^2 \quad \text{(a famous result of Euler!)}. \tag{1.2.4}$$

But is this allowed? The simplest theorem that justifies *termwise integration* involves *uniform convergence*.

THEOREM 1.2.2. *Suppose that the series* $\sum_1^\infty g_n(t)$, *with continuous functions* $g_n(t)$, *is uniformly convergent on the finite closed interval* $a \le t \le b$. *Then the sum* $f(t)$ *of the series is continuous on* $[a, b]$, *and for* c, $t \in [a, b]$,

$$\sum_1^\infty \int_c^t g_n(s)ds = \int_c^t \sum_1^\infty g_n(s)\,ds = \int_c^t f(s)ds.$$

Application to Example 1.2.1. We will show that the complex series in (1.1.1) is uniformly convergent for $|t| \le b < \pi$; the same will then be true for the series in (1.1.2), (1.1.3) which are obtained by taking real and imaginary parts.

Accordingly, set

$$g_n(t) = (-1)^{n-1} e^{int} \cdot \frac{1}{n} = a_n \cdot b_n, \quad a_1 + \cdots + a_n = A_n.$$

Denoting the k-th partial sum $\sum_1^k g_n(t)$ by $S_k(t)$, partial summation as in Lemma 1.1.3 with $j < k$ gives

$$S_k(t) - S_j(t) = \sum_{n=j+1}^{k} a_n b_n = \sum_{n=j}^{k-1} A_n(b_n - b_{n+1}) + A_k b_k - A_j b_j.$$

Using inequality (1.1.5) we thus obtain the estimate

$$\begin{aligned} |S_k(t) - S_j(t)| &\le \sum_{n=j}^{k-1} |A_n|\,|b_n - b_{n+1}| + |A_k|\,|b_k| + |A_j|\,|b_j| \\ &\le \frac{1}{|\cos\frac{1}{2}t|}\left\{\sum_{n=j}^{k-1}\left(\frac{1}{n} - \frac{1}{n+1}\right) + \frac{1}{k} + \frac{1}{j}\right\} \le \frac{1}{|\cos\frac{1}{2}t|}\,\frac{2}{j}. \end{aligned}$$

It follows that $S_k(t) - S_j(t) \to 0$ as $j, k \to \infty$, uniformly for $|t| \le b < \pi$. Hence by a criterion of Augustin-Louis Cauchy (France, 1789–1857; [**12**]), the series $\sum_1^\infty g_n(t)$ in (1.1.1) is uniformly convergent for $|t| \le b$.

The same is true for the series in (1.1.3) which is $\sum_1^\infty \operatorname{Im} g_n(t)$. Integrating from 0 to b we now obtain from Theorem 1.2.2 that

$$\frac{1}{4}\,b^2 = (-\cos b + 1) + \left(\frac{1}{2^2}\cos 2b - \frac{1}{2^2}\right) + \left(\frac{1}{3^2}\cos 3b - \frac{1}{3^2}\right) + \cdots .$$

Replacing b by t we obtain (1.2.1) for $0 \le t < \pi$; by symmetry it will be true for $|t| < \pi$. Formula (1.2.1) will also hold for $|t| = \pi$, since both sides of (1.2.1) will represent continuous functions on $[-\pi, \pi]$ (by uniform convergence of the series!).

Example 1.2.3. Formal termwise differentiation of the series in (1.1.3) would give

$$\frac{1}{2} = \cos t - \cos 2t + \cos 3t - \cos 4t + \cdots . \tag{1.2.5}$$

Is this a correct result? Is the new series uniformly convergent? No, it is not even convergent, since the terms do not tend to zero (take $t = 0$ for example)! Can one attach a meaning to (1.2.5)? Formulas of this type occur in the work of Euler, but Abel [a hundred years later] had no use for divergent series. The contemporary view is that (1.2.5) makes sense with appropriate interpretation. One could apply a suitable summability method, or one may consider convergence in the generalized sense of distribution theory; see Chapters 3 and 4.

Exercises 1.2.1. Prove that the series

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n}$$

is uniformly convergent for $\delta \leq x \leq 2\pi - \delta$ (where $0 < \delta < \pi$). Is the series uniformly convergent for $-\delta \leq x \leq \delta$?

1.2.2. Use partial summation to show that the partial sums

$$S_k(x) = \sum_{n=1}^{k} \frac{\sin nx}{n}$$

remain bounded on $-\delta \leq x \leq \delta \, (< \pi)$, hence on $\mathbb{R}$. Is this also true for the corresponding cosine series?

1.2.3. Compute the sums of the series

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}, \quad \sum_{n=1}^{\infty} \frac{\sin nx}{n^3}, \quad \sum_{n=1}^{\infty} \frac{\cos nx}{n^4}.$$

1.2.4. What formulas do you obtain by termwise differentiation of the results obtained in Exercise 1.1.4?

1.2.5. *Other manipulations.* Use the result

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2} \quad \text{for } \ 0 < x < 2\pi$$

to sum the series

$$\sum_{n=1}^{\infty} \frac{\sin 2nx}{2n} \quad \text{and} \quad \sum_{k=1}^{\infty} \frac{\sin(2k-1)x}{2k-1} \quad \text{on } \ (0, \pi).$$

Next verify the following representation for the *signum function*

$$\operatorname{sgn} x \stackrel{\text{def}}{=} \left\{ \begin{array}{ll} 1 & \text{for } x > 0 \\ -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \end{array} \right\} = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\sin(2k-1)x}{2k-1}$$

on $(-\pi, \pi)$. Derive that on the same interval

$$|x| = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos(2k-1)x}{(2k-1)^2}.$$

1.2.6. Compute the sum of the series

$$\cos x - \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x - \cdots \quad \text{on} \quad (-\pi, \pi).$$

1.3. Vibrating string and sine series. A controversy

The one-dimensional wave equation. We consider a tightly stretched homogeneous string, whose equilibrium position is the interval $[0, L]$ of the X-axis, and whose ends are kept fixed. Idealizing, one supposes that the string only carries out transverse vibrations in the "vertical" (X, U)-plane (a reasonable approximation when the displacements are small). The point of the string with coordinates $(x, 0)$ in the equilibrium position has transverse displacement $u = u(x, t)$ at time t. At time t, the generic point P of the string has coordinates $(x, u) = (x, u(x, t))$.

It is also supposed that the tension $T = T(x, u)$ in the string is large and that the string is perfectly flexible. Then the force exerted by the part of the string to the left of the point P upon the part to the right of P will be tangential to the string. The horizontal component of that force will thus be $T \cos\alpha$, the vertical component $T \sin\alpha$, where α is the angle of the string with the horizontal at P (see Figure 1.1). We suppose furthermore that there are no external forces: no gravity, no damping, etc.

Let us now focus our attention on the part of the string "above" the interval $(x, x+\Delta x)$ of the X-axis. Since there are no horizontal displacements, the net horizontal force on our part must be zero:

$$(T + \Delta T)\cos(\alpha + \Delta\alpha) - T\cos\alpha = 0, \quad \text{hence } \; T\cos\alpha = \text{const} = T_0,$$

say. The net vertical force will be

$$(T + \Delta T)\sin(\alpha + \Delta\alpha) - T\sin\alpha = T_0 \tan(\alpha + \Delta\alpha) - T_0 \tan\alpha.$$

This force will give rise to "vertical" motion by Newton's second law: force = mass $\times$ acceleration, applied at the center of mass (x', u'). Since the

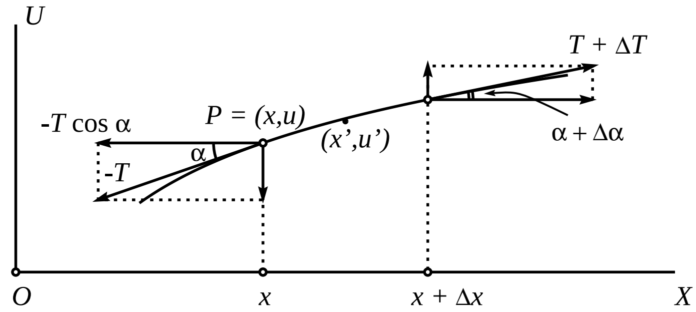

FIGURE 1.1

mass of our part is the same as in the equilibrium position, where it equals density $\times$ length $= \rho_0 \Delta x$, say, we obtain

$$T_0 \tan(\alpha + \Delta\alpha) - T_0 \tan\alpha = \rho_0 \Delta x \cdot \frac{\partial^2 u}{\partial t^2}(x', t).$$

Now $\tan\alpha = \partial\alpha/\partial x$; dividing both sides by Δx and letting $\Delta x \to 0$, we obtain the *one-dimensional wave equation:*

$$(1.3.1) \qquad T_0 \frac{\partial^2 u}{\partial x^2} = \rho_0 \frac{\partial^2 u}{\partial t^2} \quad \text{or} \quad u_{xx} = \frac{1}{c^2} u_{tt}, \quad 0 < x < L, \ \ t \in \mathbb{R},$$

where $c = \sqrt{T_0/\rho_0}$. Observe that c has the dimension of a velocity. This is confirmed by dimensional analysis: $\{(ml/t^2)/(m/l)\}^{\frac{1}{2}} = l/t$.

In the physical situation, the requirement that the ends of the string be kept fixed imposes the *boundary conditions*

$$(1.3.2) \qquad u(0, t) = 0, \quad u(L, t) = 0, \quad \forall\, t.$$

PROBLEM 1.3.1. *Initial value problem for the string with fixed ends.* Let us consider the initial value problem for our string in the situation where the string is released at time $t = 0$ from an arbitrary starting position:

$$(1.3.3) \qquad u(x, 0) = f(x), \quad 0 \le x \le L;$$

cf. Figure 1.2. Here we must of course ask that f be continuous and that $f(0) = f(L) = 0$. For $t = 0$, each point of the string has velocity zero:

$$(1.3.4) \qquad \frac{\partial u}{\partial t}(x, 0) = 0, \quad 0 \le x \le L.$$

The question is if Problem 1.3.1, given by (1.3.1)–(1.3.4), always has a solution, and if it is unique.

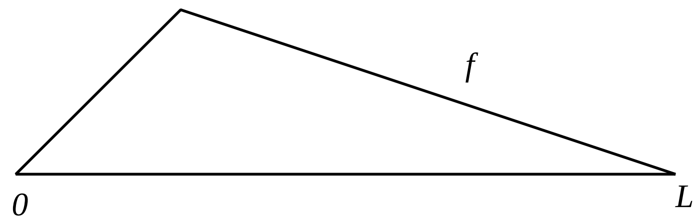

FIGURE 1.2

Having seen vibrating strings, one would probably say that the simplest initial position is given by a sinusoid:

$$u(x,0) = \sin\frac{\pi}{L}x.$$

For this initial position there is a *standing wave* solution of our problem, that is, a product solution

$$u(x,t) = v(x)\cdot w(t).$$

Taking $v(x) = u(x,0) = \sin\frac{\pi}{L}x$, our conditions lead to the following requirements for $w(t)$:

$$w'' = -\frac{\pi^2c^2}{L^2}\,w, \quad w(0) = 1, \quad w'(0) = 0.$$

Thus $w(t) = \cos\frac{\pi}{L}ct$ and

$$u(x,t) = \sin\frac{\pi}{L}\,x\,\cos\frac{\pi}{L}\,ct.$$

This formula describes the so-called *fundamental mode* of vibration of the string, which produces the "fundamental tone". The period of this vibration (the time it takes for $\frac{\pi}{L}ct$ to increase by 2π) is $\frac{2L}{c}$. Thus the "fundamental frequency" (the number of vibrations per second) equals

$$\frac{c}{2L} = \frac{1}{2L}\sqrt{\frac{T_0}{\rho_0}}.$$

By change of scale we may assume that the length L of the string is equal to π. Making this simplifying assumption from here on, we have $u(x,0) = \sin x$ and the fundamental mode becomes

$$u(x,t) = \sin x\cos ct;$$

cf. Figure 1.3. Analogously, the initial position $u(x,t) = \sin 2x$ of the string

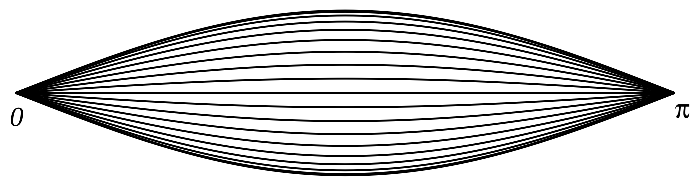

FIGURE 1.3

leads to the standing wave solution $u(x,t) = \sin 2x \cos 2ct$. More generally, the initial position $u(x,0) = \sin nx$ leads to the *standing wave solution*

$$u(x,t) = \sin nx \cos nct, \quad n \in \mathbb{N}. \tag{1.3.5}$$

The frequency in this mode of vibration is precisely n times the fundamental frequency – what we hear is the n-th harmonic overtone.

Exercises 1.3.1. Show that the vibrating string problem (1.3.1), (1.3.2), (1.3.4) with $L = \pi$ has no standing wave solutions $u(x,t) = v(x)w(t)$ other than (1.3.5), apart from constant multiples.

Hint. "Separating variables", the differential equation (1.3.1) requires that

$$\frac{v''(x)}{v(x)} = \frac{1}{c^2}\,\frac{w''(t)}{w(t)} = \lambda, \quad \text{a constant.}$$

Thus $v(x)$ has to be an "eigenfunction" for the problem

$$v'' = \lambda v, \ \ 0 < x < \pi, \quad v(0) = v(\pi) = 0; \ \ \text{cf. (1.3.2).}$$

Returning to the initial value problem 1.3.1 with general $f(x)$ (but $L = \pi$), we observe that the conditions (1.3.1), (1.3.2), (1.3.4) are *linear*. Thus *superpositions* of solutions to that part of the problem are also solutions. More precisely, any finite linear combination

$$u_k(x,t) = \sum_{n=1}^{k} b_n \sin nx \cos nct$$

of solutions (1.3.5) is also a solution of (1.3.1), (1.3.2), (1.3.4). This combination will solve the whole problem – including (1.3.3) – if the initial position of the string has the special form $f(x) = \sum_{n=1}^{k} b_n \sin nx$. Boldly going to infinite sums, it seems plausible that the expression

$$u(x,t) = \sum_{n=1}^{\infty} b_n \sin nx \cos nct \tag{1.3.6}$$

will solve the Initial value Problem 1.3.1, provided the initial position of the string can be represented in the form

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx, \quad 0 \le x \le \pi. \tag{1.3.7}$$

A controversy. Around 1750, the problem of the vibrating string with fixed end points, Problem 1.3.1, was considered by Jean le Rond d'Alembert (Paris, 1717–1783; [**3**]), Euler and Daniel Bernoulli. The latter claimed that every mode of vibration can be represented in the form (1.3.6), that is, every mode can be obtained by superposing (multiples of) the fundamental mode and higher harmonics. The implication would be that every *geometrically* given initial shape $f(x)$ of the string can be represented by a sine series (1.3.7). Euler found it difficult to accept this. He did not believe that every geometrically given initial shape $f(x)$ on $(0, \pi)$ could be equal to (what to him looked like) an *analytic expression* $\sum_{n=1}^{\infty} b_n \sin nx$. Euler's authority was such that Bernoulli's proposition was rejected. Several years later, Fourier made Bernoulli's ideas more plausible. He gave many examples of functions with representations (1.3.7) and related "Fourier series", but a satisfactory proof of the representations under fairly general conditions on f had to wait for Dirichlet (around 1830).

1.4. Heat conduction and cosine series

Heat or thermal energy is transferred from warmer to cooler parts of a solid by conduction. One speaks of heat flow, in analogy to fluid flow or diffusion. Denoting the temperature at the point P and the time t by $u = u(P, t)$, the *basic postulate of heat condution* is that the heat flow vector $\vec{q}$ at P is proportional to $-\mathrm{grad}\,\mathrm{u}$:

$$\vec{q} = -\lambda\, \mathrm{grad}\, \mathrm{u} = -\lambda \left(\frac{\partial \mathrm{u}}{\partial \mathrm{x}}, \frac{\partial \mathrm{u}}{\partial \mathrm{y}}, \frac{\partial \mathrm{u}}{\partial \mathrm{z}} \right).$$

Here λ is called the thermal conductivity (at P and t). Thus the heat flow across a small surface element ΔS at P over a small time interval $[t, t + \Delta t]$, and to the side indicated by the normal $\vec{N}$, is approximately equal to $-\lambda(\partial u/\partial N)\Delta S \Delta t$; cf. Figure 1.4

Here we will consider the heat flow in a *thin homogeneous rod*, occupying the segment $[0, L]$ of the X-axis. We suppose that there are no heat sources in the rod and that heat flows only in the X-direction (there is no heat

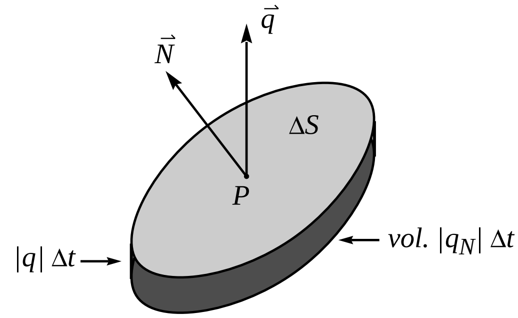

FIGURE 1.4

flow across the lateral surface of the rod). [One would have similar one-dimensional heat flow in an infinite slab, bounded by the parallel planes $\{x = 0\}$ and $\{x = L\}$ in space.] We now concentrate on the element $[x, x + \Delta x]$ of the rod; cf. Figure 1.5. The quantity of heat entering this element across the left-hand face, over the small time interval $[t, t + \Delta t]$, will be approximately $-\lambda(\partial u/\partial x)(x, t)\Delta S\Delta t$, where ΔS denotes the area of the cross section of the rod. Similarly, the heat leaving the element across the right-hand face will be $-\lambda(\partial u/\partial x)(x + \Delta x, t)\Delta S\Delta t$. Thus the net amount of heat flowing into the element over the time interval $[t, t+\Delta t]$ is approximately

$$\Delta Q = \lambda \left[\frac{\partial u}{\partial x}(x + \Delta x, t) - \frac{\partial u}{\partial x}(x, t) \right] \Delta S \Delta t.$$

The heat flowing into our element will increase the temperature, say by Δu. This temperature increase Δu will require a number of calories $\Delta Q'$ proportional to Δu and to the volume $\Delta S \Delta x$ of the element, hence

$$\Delta Q' \approx c \Delta u \Delta S \Delta x,$$

where c is the specific heat of the material.

Equating $\Delta Q'$ to ΔQ and dividing by $\Delta S \Delta x \Delta t$, one finds the approximate equation

$$c\,\frac{\Delta u}{\Delta t} = \lambda\,\frac{\dfrac{\partial u}{\partial x}(x + \Delta x, t) - \dfrac{\partial u}{\partial x}(x, t)}{\Delta x}.$$

Passing to the limit as $\Delta x \to 0$ and $\Delta t \to 0$, we obtain the *one-dimensional heat* or *diffusion equation*:

$$\frac{\partial u}{\partial t} = \beta\,\frac{\partial^2 u}{\partial x^2}, \quad 0 < x < L, \quad t \in \mathbb{R}, \tag{1.4.1}$$

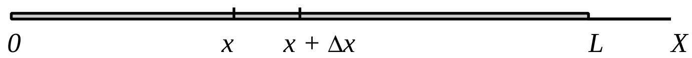

FIGURE 1.5

where $\beta = \lambda/c > 0$. For the homogeneous rod it is reasonable to treat β as a constant.

We could now prescribe the temperature at the ends of the rod and study corresponding heat flow(s). The simplest case would involve constant temperatures $u(0,t)$ and $u(L,t)$ at the ends. Subtracting a suitable linear function of x from $u(x,t)$, we might as well require that $u(0,t) = 0$ and $u(L,t) = 0$ for all t. Then we would have the same boundary conditions as in (1.3.2), and this would again lead to sine functions and sine series. A different situation arises when one keeps the *ends* of the rod *insulated*. There will then be no heat flow across the ends. The resulting boundary conditions are

$$\frac{\partial u}{\partial x}(0,t) = 0, \quad \frac{\partial u}{\partial x}(L,t) = 0, \quad \forall\, t. \tag{1.4.2}$$

PROBLEM 1.4.1. *Rod with insulated ends.* Let us consider the problem where the temperature along the rod is prescribed at time $t = 0$:

$$u(x,0) = f(x), \quad 0 \le x \le L. \tag{1.4.3}$$

In view of (1.4.2) we will now require that $f'(0) = f'(L) = 0$. The question is if Problem 1.4.1, given by (1.4.1)–(1.4.3), always has a solution, and if it is unique.

Just as in Section 1.3, we may and will take $L = \pi$. Time-independent solutions $u(x,t) = v(x)$ of (1.4.2) must then satisfy the conditions $v'(0) = v'(\pi) = 0$. This suggests cosine functions for $v(x)$ instead of sines:

$$v(x) = 1,\ \cos x,\ \cos 2x,\ \cdots,\ \cos nx,\ \cdots.$$

Corresponding *stationary mode* solutions, or product solutions, $u(x,t) = v(x)w(t) = (\cos nx)w(t)$ of (1.4.1) must satisfy the condition

$$(\cos nx)w'(t) = \beta\,(-n^2 \cos nx)w(t).$$

This leads to the following solutions of problem (1.4.1), (1.4.2) with $L = \pi$:

$$u(x,t) = (\cos nx)e^{-n^2\beta t}, \quad n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}. \tag{1.4.4}$$

Indeed, w has to satisfy the conditions $w' = -n^2\beta w$, $w(0) = 1$.

Superpositions of solutions (1.4.4) also satisfy (1.4.1), (1.4.2) (with $L = \pi$). We immediately take an infinite sum

$$u(x,t) = \sum_{n=0}^{\infty} a_n (\cos nx) e^{-n^2 \beta t}, \tag{1.4.5}$$

and ask if with such a sum, we can satisfy the general initial condition (1.4.3). In other words, can every (reasonable) function $f(x)$ on $[0, \pi]$ be represented by a cosine series,

$$f(x) = u(x,0) = \sum_{n=0}^{\infty} a_n \cos nx, \quad 0 \le x \le \pi\,? \tag{1.4.6}$$

Exercises 1.4.1. Show that the heat flow problem (1.4.1), (1.4.2) with $L = \pi$ has no stationary mode solutions $u(x,t) = v(x)w(t)$ other than (1.4.4), apart from constant multiples. [Which eigenvalue problem for v is involved?]

1.5. Fourier series

If a function f on $\mathbb{R}$ is for every x equal to the sum of a sine series (1.3.7), then f is *odd*: $f(-x) = -f(x)$, and *periodic* with period 2π: $f(x + 2\pi) = f(x)$. Similarly, if a function f on $\mathbb{R}$ is for every x equal to the sum of a cosine series (1.4.6), then f is *even*: $f(-x) = f(x)$, and *periodic* with period 2π. Suppose now that every (reasonable) function f on $(0, \pi)$ can be represented both by a sine series and by a cosine series. Then every odd 2π-periodic function on $\mathbb{R}$ can be represented (on all of $\mathbb{R}$) by a sine series, every even 2π-periodic function by a cosine series. It will then follow that every (reasonable) 2π-periodic function on $\mathbb{R}$ can be represented by a *trigonometric series*

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx). \tag{1.5.1}$$

Indeed, every function f on $\mathbb{R}$ is equal to the sum of its even part and its odd part, and if f has period 2π, so do those parts:

$$f(x) = \frac{1}{2}\{f(x) + f(-x)\} + \frac{1}{2}\{f(x) - f(-x)\}.$$

Conversely, if every 2π-periodic function f on $\mathbb{R}$ has a representation (1.5.1), then every function f on $(0, \pi)$ can be represented by a sine series [as well as by a cosine series]. Indeed, any given f on $(0, \pi)$ can be extended to an odd function of period 2π, and for the extended function f, (1.5.1) would imply

$$\begin{aligned}
f(x) = -f(-x) &= \frac{1}{2}\{f(x) - f(-x)\} \\
&= \frac{1}{2}\left\{a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)\right. \\
&\qquad \left. -a_0 - \sum_{n=1}^{\infty}(a_n \cos nx - b_n \sin nx)\right\} \\
&= \sum_{n=1}^{\infty} b_n \sin nx.
\end{aligned}$$

[To obtain a cosine series, one would extend f to an even function of period 2π.]

It is often useful to consider a function f of period 2π as a function on the unit circle $C(0,1)$ in the complex plane:

$$C(0,1) = \{z \in \mathbb{C} : z = e^{it}, \quad -\pi < t \leq \pi\}.$$

Using independent variable t instead of x, the 2π-periodic function f may be represented in the form

$$f(t) = g(e^{it}), \quad t \in \mathbb{R}, \tag{1.5.2}$$

where $g(z)$ is defined on the unit circumference. For readers with a basic knowledge of Complex Analysis we can now discuss a (rather strong) condition on $f(t) = g(e^{it})$ which ensures that there is a representation (1.5.1) [with t instead of x]. Note that is customary to replace the constant term a_0 in (1.5.1) by $\frac{1}{2}a_0$ in order to obtain uniform formulas for the coefficients a_n.

THEOREM 1.5.1. *Let $f(t) = g(e^{it})$ be a function on $\mathbb{R}$ with period 2π such that $g(z)$ has an analytic extension from the unit circle $C(0,1) = \{|z| = 1\}$ to some annulus $A(0; r, R) = \{r < |z| < R\}$ with $r < 1 < R$. Then*

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int} = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt), \tag{1.5.3}$$

where

$$
\begin{aligned}
c_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt, \quad \forall\, n \in \mathbb{Z},\\
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt, \quad n = 0, 1, 2, \cdots,\\
b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt, \quad n = 1, 2, \cdots.
\end{aligned} \tag{1.5.4}
$$

PROOF. An analytic function $g(z)$ on the annulus $A(0; r, R)$ can be represented by the *Laurent series*

$$
g(z) = \sum_{n=-\infty}^{\infty} c_n z^n, \quad r < |z| < R,
$$

where

$$
c_n = \frac{1}{2\pi i}\int_{C(0,1)^+} g(z)z^{-n-1}dz = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(e^{it})e^{-int}dt, \quad \forall\, n \in \mathbb{Z}.
$$

This result from Complex Analysis implies the first representation for $f(t) = g(e^{it})$ in (1.5.3) with c_n as in (1.5.4). Here the series for $f(t)$ will be absolutely convergent. In fact, the coefficients c_n will satisfy an inequality of the form $|c_n| \le Me^{-\delta|n|}$ with $\delta > 0$; cf. Exercise 1.5.6.

In order to obtain the second representation in (1.5.3) one combines the terms in the first series corresponding to n (> 0) and its negative. Thus

$$
\begin{aligned}
&c_n e^{int} + c_{-n}e^{-int}\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)e^{-ins}ds \cdot e^{int} + \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)e^{ins}ds \cdot e^{-int}\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\cdot 2\cos n(s-t)\, ds\\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} f(s)\cos ns\, ds \cdot \cos nt + \frac{1}{\pi}\int_{-\pi}^{\pi} f(s)\sin ns\, ds \cdot \sin nt\\
&= a_n \cos nt + b_n \sin nt,
\end{aligned} \tag{1.5.5}
$$

with a_n, b_n as in (1.5.4). Finally taking $n = 0$, one finds that

$$
c_0\, e^{i0t} = c_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)ds = \frac{1}{2}a_0. \tag{1.5.6}
$$

$\square$

Definition 1.5.2. Let f on $\mathbb{R}$ be 2π-periodic and integrable over a period. Then the numbers a_n, b_n computed with the aid of (1.5.4) are called the *Fourier coefficients* of f, and the second series in (1.5.3), formed with these coefficients, is called the *Fourier series* for f. We write

$$f(t) \sim \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt), \tag{1.5.7}$$

with the *symbol* $\sim$, to emphasize that the series on the right is the Fourier series of $f(t)$, but that nothing is implied about convergence. The numbers c_n determined by (1.5.4) are called the complex Fourier coefficients of f and the first series in (1.5.3), formed with these coefficients, is called the *complex Fourier series* for f.

Question 1.5.3. The basic problem is: under what conditions, and in what sense, will the Fourier series of f converge to f? We would of course want conditions weaker than the analyticity condition in Theorem 1.5.1.

For clarity, the Fourier coefficients of f will often be written as $a_n[f]$, $b_n[f]$, $c_n[f]$. The partial sums of the Fourier series for f will be denoted by $s_k[f]$; the sum $s_k[f]$ will also be equal to the *symmetric* partial sum of the complex Fourier series:

$$\begin{aligned} s_k[f](t) &\stackrel{\text{def}}{=} \frac{1}{2}a_0[f] + \sum_{n=1}^{k} (a_n[f] \cos nt + b_n[f] \sin nt) \\ &= \sum_{n=-k}^{k} c_n[f] e^{int}; \end{aligned} \tag{1.5.8}$$

cf. (1.5.5), (1.5.6). Instead of variable t one may of course use x or any other letter. In ch 2 we will derive an integral formula for $s_k[f]$. From that formula we will among others obtain a convergence theorem for the case of piecewise smooth functions.

Definition 1.5.4. For any integrable function f on $(-\pi, \pi)$ or on $(0, 2\pi)$, the Fourier series is defined as the Fourier series for the 2π-periodic extension. For integrable f on $(0, \pi)$, the *Fourier cosine series* and the *Fourier sine series*,

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx \quad \text{and} \quad \sum_{n=1}^{\infty} b_n \sin nx,$$

are defined as the Fourier series for the *even* extension of f with period 2π, and the *odd* extension, respectively.

Exercises 1.5.1. Prove that the Fourier series for an even 2π-periodic function is a cosine series, and that the Fourier series for an odd 2π-periodic function is a sine series.

1.5.2. Let f be integrable on $(0, \pi)$. Prove that for the Fourier cosine and sine series of f,

$$a_n = \frac{2}{\pi} \int_0^\pi f(t) \cos nt \, dt, \quad b_n = \frac{2}{\pi} \int_0^\pi f(t) \sin nt \, dt.$$

1.5.3. Determine the Fourier cosine and sine series for $f(x) = 1$ on $(0, \pi)$.

1.5.4. Same question for $f(x) = x$ on $(0, \pi)$.

1.5.5. Do you see a connection between the series in Exercises 1.5.3, 1.5.4 and certain trigonometric series which we encountered earlier?

1.5.6. Let $f(t) = g(e^{it})$, where $g(z)$ is analytic on the annulus given by $e^{-\delta} \le |z| \le e^\delta$ and in absolute value bounded by M. Use Cauchy's theorem [**14**] and suitable circles of integration to show that $|c_n[f]| \le Me^{-\delta|n|}$ for all n.

1.5.7. Let $U(x, y)$ denote a stationary temperature distribution in a planar domain D. In polar coordinates, the temperature becomes a function of r and θ, $U(r \cos\theta, r \sin\theta) = u(r, \theta)$, say. It will satisfy Laplace's equation, named after the French mathematician-astronomer Pierre-Simon Laplace (1749–1827; [**73**]):

$$\Delta U \overset{\text{def}}{=} \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0.$$

In the case of $D = B(0, 1)$, the unit disc, the geometry implies a periodicity condition, $u(r, \theta + 2\pi) = u(r, \theta)$. Also, $u(r, \theta)$ must remain finite as $r \searrow 0$. Show that in polar coordinates, Laplace's equation on $B(0, 1)$ has product solutions $u(r, \theta)$ of the form $v_n(r) \cos n\theta$, $n \in \mathbb{N}_0$, and $v_n(r) \sin n\theta$, $n \in \mathbb{N}$. Determine $v_n(r)$ if $v_n(1) = 1$. What are the most general product solutions $u(r, \theta) = v(r)w(\theta)$ of Laplace's equation on the disc $B(0, 1)$?

1.5.8. (Continuation) We wish to solve the so-called Dirichlet problem for Laplace's equation on the unit disc:

$$\Delta U = 0 \text{ on } B(0, 1), \quad U = F \text{ on } C(0, 1).$$

[Stationary temperature distribution in the disc corresponding to prescribed boundary temperatures.] Assuming that the boundary function F, written as $f(\theta)$, can be represented by a Fourier series, one asks for a solution $u(r,\theta)$ in the form of an infinite series.

1.6. Fourier series as orthogonal series

A function f will be called square-integrable on (a,b) if f is integrable over every finite subinterval, and $|f|^2$ is integrable over the whole interval (a,b); cf. Section 5.5. If f and g are square-integrable on (a,b) the product $f\overline{g}$ will have a finite integral over (a,b). Square-integrable functions f and g are called *orthogonal* on (a,b), and we write $f \perp g$, if

$$\int_a^b f\overline{g} = \int_a^b f(x)\overline{g(x)}dx = 0. \tag{1.6.1}$$

One may introduce a related abstract *inner product* by the formula

$$(u,v) = \int_a^b u\overline{v} = \int_a^b u(x)\overline{v(x)}dx. \tag{1.6.2}$$

Definition 1.6.1. A family $\phi_1, \phi_2, \phi_3, \cdots$ of square-integrable functions on (a,b) is called an *orthogonal system* on (a,b) if the functions are pairwise orthogonal and none of them is (equivalent to) the zero function:

$$\int_a^b \phi_n\overline{\phi}_k = 0, \quad k \neq n; \qquad \int_a^b |\phi_n|^2 > 0, \quad \forall\, n.$$

Other index sets than $\mathbb{N}$ will occur, and if (a,b) is finite, we may also speak of an orthogonal system on $[a,b]$.

Examples 1.6.2. Each of the systems

$$\begin{aligned}&\frac{1}{2},\ \cos x,\ \cos 2x,\ \cdots,\ \cos nx,\ \cdots,\\ &\sin x,\ \sin 2x,\ \cdots,\ \sin nx,\ \cdots,\end{aligned}$$

is orthogonal on $(0,\pi)$ [and also on $(-\pi,\pi)$]. Each of the systems

$$\begin{aligned}&\frac{1}{2},\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots,\ \cos nx,\ \sin nx,\ \cdots,\\ &1,\ e^{ix},\ e^{-ix},\ e^{2ix},\ e^{-2ix},\ \cdots,\ e^{inx},\ e^{-inx},\ \cdots\end{aligned}$$

is orthogonal on $(-\pi,\pi)$ [and also on every other interval of length 2π]. We will verify the orthogonality of the first and the last system:

$$\int_0^\pi \cos nx \cos kx\, dx = \frac{1}{2}\int_0^\pi \{\cos(n+k)x + \cos(n-k)x\}\, dx$$
$$= \frac{1}{2}\left[\frac{\sin(n+k)x}{n+k} + \frac{\sin(n-k)x}{n-k}\right]_0^\pi = 0 \;\; for \;\; k \neq n \;\; (n,\, k \geq 0);$$
$$\int_{-\pi}^\pi e^{inx}e^{-ikx}dx = \left[\frac{e^{i(n-k)x}}{i(n-k)}\right]_{-\pi}^\pi = 0 \;\; for \;\; k \neq n.$$

If $\{\phi_n\}$, $n \in \mathbb{N}$ is an orthogonal system, a series $\sum_{n=1}^\infty c_n\phi_n$ with constant coefficients c_n will be called an *orthogonal series* [the terms in the series are pairwise orthogonal]. Fourier cosine series and Fourier sine series are orthogonal series on $(0,\pi)$. Complex Fourier series are orthogonal series on $(-\pi,\pi)$, and so are real Fourier series.

If an orthogonal series converges in an appropriate sense, the coefficients can be expressed in terms of the sum function in a simple way:

LEMMA 1.6.3. *Let $\{\phi_n\}$, $n \in \mathbb{N}$ be an orthogonal system of piecewise continuous functions on the bounded closed interval $[a,b]$. Suppose that a certain series $\sum_{n=1}^\infty c_n\phi_n$ converges uniformly on $[a,b]$ to a piecewise continuous function f:*

$$\sum_{n=1}^\infty c_n\phi_n(x) = f(x), \quad uniformly\ on\ [a,b]. \tag{1.6.3}$$

Then

$$c_n = \frac{\int_a^b f\overline{\phi}_n}{\int_a^b |\phi_n|^2}, \quad \forall\, n. \tag{1.6.4}$$

PROOF. Since the function $\overline{\phi}_k$ will be bounded on $[a,b]$, it follows from the hypothesis that the series

$$\sum_{n=1}^\infty c_n\phi_n\overline{\phi}_k \text{ converges uniformly to } f\overline{\phi}_k \text{ on } [a,b].$$

Thus we may integrate term by term to obtain

$$\int_a^b f\overline{\phi}_k = \sum_{n=1}^\infty c_n \int_a^b \phi_n\overline{\phi}_k = c_k\int_a^b |\phi_k|^2.$$

In the final step we have used the orthogonality of the system $\{\phi_n\}$. The result gives (1.6.4) [with k instead of n]. □

The lemma shows that for given $\{\phi_n\}$ and f, there is at most one orthogonal representation (1.6.3) [with uniform convergence].

DEFINITION 1.6.4. For a given orthogonal system $\{\phi_n\}$ and given square-integrable f on (a,b), the numbers c_n computed with the aid of (1.6.4) are called the *expansion coefficients* of f with respect to the system $\{\phi_n\}$. The corresponding series $\sum_{n=1}^{\infty} c_n\phi_n$ is called the (*orthogonal*) *expansion* of f with respect to the system $\{\phi_n\}$. To emphasize that there is no implication of convergence we write

$$f \sim \sum_{n=1}^{\infty} c_n\phi_n \quad \text{or also} \quad f(x) \sim \sum_{n=1}^{\infty} c_n\phi_n(x). \tag{1.6.5}$$

QUESTIONS 1.6.5. The basic problems are: under what conditions, and in what sense, do orthogonal expansions converge, and if they converge, will they converge to the given function f? We would aim for conditions weaker than the one in Lemma 1.6.3.

These questions are best treated in the context of inner product spaces, preferably complete inner product spaces or so-called *Hilbert spaces*; cf. Chapters 5, 7. (Such spaces are named after the German mathematician David Hilbert, 1862–1943; [**48**].) The square-integrable functions on (a,b) with the inner product given by (1.6.2) form an inner product space. It is best to use integrability in the sense of Lebesgue here (see Section 2.1), because then the square-integrable functions on (a,b) form a Hilbert space, the space $L^2(a,b)$.

Fourier series can be considered as orthogonal expansions. Thus the complex Fourier series of a square-integrable function f on $(-\pi,\pi)$ is the same as its expansion with respect to the orthogonal system $\{e^{inx}\}$, $n = 0, \pm1, \pm2, \cdots$:

$$c_n[f] \stackrel{\text{def}}{=} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx = \frac{\int_{-\pi}^{\pi} f(x)e^{-inx}dx}{\int_{-\pi}^{\pi} |e^{inx}|^2dx}.$$

Besides sines, cosines and complex exponentials, there are many orthogonal systems of practical importance. We mention orthogonal systems of polynomials and more general orthogonal systems of eigenfunctions; cf. Chapter 7.

Exercises 1.6.1. Show that the Fourier cosine series $\frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx$ of a square-integrable function f on $(0, \pi)$ is also its orthogonal expansion with respect to the system $\frac{1}{2}$, $\cos x$, $\cos 2x$, $\cdots$ on $(0, \pi)$; cf. Exercise 1.5.2.

1.6.2. State and prove the corresponding result for the Fourier sine series.

1.6.3. Write down the expansion of the function $f(x) = 1$ on $(-\pi, \pi)$ with respect to the orthogonal system $\sin x$, $\sin 2x$, $\cdots$ on $(-\pi, \pi)$. Does the expansion converge? Does it converge to $f(x)$?

1.6.4. Same questions for the expansion of the function $f(x) = 1 + x$ on $(-\pi, \pi)$ with respect to the orthogonal system $\frac{1}{2}$, $\cos x$, $\cos 2x$, $\cdots$ on $(-\pi, \pi)$.

1.6.5. Determine the expansion of the function $f(x) = e^{\alpha x}$ on $(0, 2\pi)$ with respect to the orthogonal system $\{e^{inx}\}$, $n \in \mathbb{Z}$ on $(0, 2\pi)$.

1.7. Fourier integrals

Many boundary value problems for (partial) differential equations involve infinite media and for such problems one needs an analog to Fourier series for infinite intervals. We will indicate how Fourier series go over into Fourier integrals as the basic interval expands to the whole line $\mathbb{R}$.

For a locally integrable function f on $\mathbb{R}$ with period $2L$ instead of 2π one obtains the Fourier series by a simple change of scale. Indeed, $f\left(\frac{L}{\pi}t\right)$ will now have period 2π as a function of t. Hence it has the following Fourier series:

$$f\left(\frac{L}{\pi}t\right) \sim \sum_{n=-\infty}^{\infty} c_n(L) e^{int} \quad \text{on} \quad (-\pi, \pi), \quad \text{where}$$

$$c_n(L) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(\frac{L}{\pi}t\right) e^{-int} dt.$$

Changing scale, one obtains the Fourier series for $f(x)$ on $(-L, L)$:

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n(L) e^{in(\pi/L)x} \quad \text{on} \quad (-L, L), \quad \text{where}$$

$$c_n(L) = \frac{1}{2L} \int_{-L}^{L} f(x) e^{-in(\pi/L)x} dx. \tag{1.7.1}$$

Suppose now that $f(x)$ is defined on $\mathbb{R}$, not periodic but relatively small as $x \to \pm\infty$, and so well-behaved that for every $L > 0$, the restriction of

f to $(-L, L)$ is equal to the sum of its Fourier series for that interval. For large x we will now use the approximation

$$c_n(L) \approx \frac{1}{2L} \int_{-\infty}^{\infty} f(x) e^{-in(\pi/L)x} dx.$$

[If $f(x)$ vanishes outside some finite interval $(-b, b)$, the approximation will be exact if we take $L \geq b$.] At this point it is convenient to introduce the so-called *Fourier transform* of f on $\mathbb{R}$:

$$g(\xi) = \hat{f}(\xi) = (\mathcal{F}f)(\xi) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx, \quad \xi \in \mathbb{R}. \tag{1.7.2}$$

In terms of g,

$$c_n(L) \approx \frac{1}{2L}\, g\left(n\frac{\pi}{L}\right).$$

Hence for large L and $-L < x < L$, the postulated equality for our $f(x)$ in (1.7.1) will give the approximate formula

$$f(x) \approx \frac{1}{2L} \sum_{n=-\infty}^{\infty} g\left(n\frac{\pi}{L}\right) e^{in(\pi/L)x} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} g\left(n\frac{\pi}{L}\right) e^{in(\pi/L)x}\, \frac{\pi}{L}. \tag{1.7.3}$$

For fixed x, the final sum may be considered as an *infinite Riemann sum*

$$\sum_{n=-\infty}^{\infty} G(\xi_n)\Delta\xi_n, \quad \text{with } \xi_n = n\frac{\pi}{L}, \;\; \Delta\xi_n = \frac{\pi}{L}, \tag{1.7.4}$$

and

$$G(\xi) = G(\xi, x) = g(\xi)e^{i\xi x}, \quad -\infty < \xi < \infty.$$

For suitably well-behaved functions $G(\xi)$, sums (1.7.4) will approach the integral $\int_{-\infty}^{\infty} G(\xi)d\xi$ as $L \to \infty$. It is therefore plausible that for fixed $x \in \mathbb{R}$, the limit may be taken in (1.7.3) as $L \to \infty$ to obtain the following integral representation for $f(x)$ in terms of g:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\xi, x) d\xi = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\xi) e^{i\xi x} d\xi. \tag{1.7.5}$$

Observe that the final integral resembles the Fourier transform $\hat{g}(x)$ of $g(\xi)$. The latter would have x instead of $-x$, or $-x$ instead of x. Thus the integral in (1.7.5) equals $\hat{g}(-x)$; it is the reflection of the Fourier transform of g, or the so-called *reflected Fourier transform*, $(\mathcal{F}_R g)(x)$. Hence we arrive at the important formula for *Fourier inversion*:

$$\text{If } \; g = \hat{f} = \mathcal{F}f, \quad \text{then} \quad f = \frac{1}{2\pi}\, \widehat{g}_R = \frac{1}{2\pi}\, \mathcal{F}_R g. \tag{1.7.6}$$

Precise conditions for the validity of the inversion theorem will be obtained in Chapters 9 and 10.

It may be of interest to *observe* that the factor $1/(2\pi)$ in formula (1.7.6) is related to the famous "Cauchy factor" $1/(2\pi i)$ of Complex Analysis:

EXAMPLE 1.7.1. Let $f(x) = e^{-a|x|}$ where $a > 0$. We compute the Fourier transform:

$$g(\xi) = \int_{-\infty}^{\infty} e^{-a|x|}e^{-i\xi x}dx = \int_{0}^{\infty} e^{-(a+i\xi)x}dx + \int_{-\infty}^{0} e^{(a-i\xi)x}dx$$

$$= \frac{1}{a+i\xi} + \frac{1}{a-i\xi} = \frac{2a}{\xi^2+a^2}. \tag{1.7.7}$$

In this case one can verify the inversion formula (1.7.6) with the aid of Complex Analysis. Indeed, introducing a complex variable $\zeta = \xi + i\eta$, one may write

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} g(\xi)e^{ix\xi}d\xi = \lim_{R\to\infty}\frac{1}{2\pi i}\int_{[-R,R]}\frac{2ia}{\zeta^2+a^2}\,e^{ix\zeta}d\zeta. \tag{1.7.8}$$

Now choose $R > a$. For $x \geq 0$ we attach to the real segment $[-R, R]$ the semicircle C_R, given by $\zeta = Re^{it}$, $0 \leq t \leq \pi$. This semicircle lies in the upper half-plane $\{\operatorname{Im}\zeta \geq 0\}$, where $|e^{ix\zeta}| = e^{-x\eta} \leq 1$. For the closed path $W_R = [-R,R] + C_R$, the Cauchy integral formula [**13**] (or the residue theorem) gives

$$\frac{1}{2\pi i}\int_{W_R}\frac{1}{\zeta-ia}\,\frac{2ia}{\zeta+ia}\,e^{ix\zeta}d\zeta$$

$$= \left\{\text{value of } \frac{2ia}{\zeta+ia}\,e^{ix\zeta} \text{ at the point } \zeta = ia\right\} = e^{-ax}. \tag{1.7.9}$$

Since

$$\left|\frac{1}{2\pi i}\int_{C_R}\frac{2ia}{\zeta^2+a^2}\,e^{ix\zeta}d\zeta\right| \leq \frac{aR}{R^2-a^2} \to 0 \text{ as } R\to\infty,$$

(1.7.9) implies that the limit on the right-hand side of (1.7.8) has the value e^{-ax}:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} g(\xi)e^{ix\xi}d\xi = e^{-ax} = e^{-a|x|} \quad (x \geq 0).$$

For $x < 0$ one may augment the segment $[-R, R]$ by a semicircle in the lower half-plane $\{\operatorname{Im}\zeta < 0\}$ to obtain the answer $e^{ax} = e^{-a|x|}$.

For the applications, the most useful property of Fourier transformation is the fact that *differentiation* goes over into ("maps to") *multiplication* by a simple function:

$$f'(x) = Df(x) \xmapsto{\mathcal{F}} i\xi \hat{f}(\xi). \tag{1.7.10}$$

Repeated application of the rule gives

$$p(D)f(x) \xmapsto{\mathcal{F}} p(i\xi)\hat{f}(\xi)$$

for any polynomial $p(t)$ with constant coefficients. Thus Fourier transformation changes an ordinary linear differential equation $p(D)u = f$ with constant coefficients into the algebraic equation $p(i\xi)\hat{u} = \hat{f}$. Applying Fourier transformation to one or more of the variables, linear partial differential equations with constant coefficients go over into differential equations with fewer independent variables. Applications to boundary value problems will be discussed in Chapters 9–12.

Exercises 1.7.1. Show that the Fourier transform of

$$f(x) = \frac{1}{x^2 + a^2} \quad \text{is equal to} \quad \frac{\pi}{a} e^{-a|\xi|} \quad (a > 0).$$

1.7.2. Prove that the Fourier transform of an even function is even, that of an odd function, odd.

1.7.3. Compute the Fourier transform $g(\xi)$ of the function

$$f(x) = \begin{cases} 1 & \text{for } |x| \le a, \\ 0 & \text{for } |x| > a. \end{cases}$$

Next use the improper integral

$$\int_{\mathbb{R}} \frac{\sin t\xi}{\xi}\, d\xi = \pi \operatorname{sgn} t$$

[for sgn see Exercise 1.2.5] to show that, also in the present case,

$$\frac{1}{2\pi} \int_{\mathbb{R}} g(\xi) e^{ix\xi} d\xi = f(x).$$

1.7.4. Let f be a "good" function: smooth, and small at $\pm\infty$. Use integration by parts to prove that $(\mathcal{F}f')(\xi) = i\xi(\mathcal{F}f)(\xi)$.

1.7.5. Prove the "dual" rule: If f is small at $\pm\infty$ and $\mathcal{F}f = g$, then $\mathcal{F}[xf(x)](\xi) = ig'(\xi)$.

1.7.6. Use the rules above to compute the Fourier transform of $f(x) = e^{-ax^2}$ where $a > 0$. Hint: One has $f'(x) = -2axf(x)$.

Books. There are many books on Fourier analysis; see the Internet for standard texts. The author mentions only some of the authors here; their books are listed in chronological order. See the bibliography for full titles.

1931 Wolff [**125**] Fourier series (in German), very short introduction

1933 Wiener [**124**] and 1934 Paley and Wiener [**88**], original work on Fourier integrals

1937 Titchmarsh [**120**], basic book on Fourier integrals

1944 Hardy and Rogosinski [**45**], Fourier series, short, scholarly

1950/1966 Schwartz [**110**], his basic work on distributions

1960 Lighthill [**81**], short, Fourier asymptotics

1971 Stein and Weiss [**114**], Fourier analysis on Euclidean spaces

1972 Dym and McKean [**28**], Fourier integrals and applications

1983 Hörmander [**52**], vol 1, his treatise on distributions for partial differential equations

1989 Körner [**70**], refreshingly different

1992 Folland [**32**], balance of theory and applications

2002 Zygmund [**129**] (predecessor 1935), two-volume standard work on Fourier series

2010 Duistermaat and Kolk [**27**], advanced text

CHAPTER 2

Pointwise convergence of Fourier series

For *smooth* periodic functions [functions of class $\mathcal{C}^1$, that is, continuously differentiable functions], the Fourier series is pointwise convergent to the function, even uniformly convergent. The smoother the function, the faster the Fourier series will converge. Pointwise convergence holds also for piecewise smooth functions, provided such functions are suitably normalized at points of discontinuity. However, for arbitrary continuous functions the Fourier series need not converge in the ordinary sense.

2.1. Integrable functions. Riemann–Lebesgue lemma

Let $[a,b]$ be a bounded closed interval. A function f on $[a,b]$ or (a,b) will be called *piecewise continuous* if there is a finite partioning $a = a_0 \leq a_1 \leq \cdots \leq a_p = b$ of $[a,b]$ as follows. On each (nonempty) open subinterval (a_{k-1}, a_k) the function f is continuous, *and* has finite right-hand and left-hand limits $f(a_{k-1}+)$ and $f(a_k-)$. The value $f(a_k)$ may be different from limits $f(a_k-)$ and/or $f(a_k+)$; this is in particular the case if $a_k = a_{k-1}$.

One similarly defines a *piecewise constant* or *step* function s; it will be *constant* on intervals (a_{k-1}, a_k). It is clear what the integral of such a function on $[a,b]$ will be; values $s(a_k)$ different from limits $s(a_k-)$ or $s(a_k+)$ will have no effect.

A real function f on $[a,b]$ will be *Riemann integrable* (after Bernhard Riemann, Germany, 1826–1866; [**98**]) if there are sequences of step functions $\{s_n\}$ and $\{s'_n\}$ such that

$$s_n(x) \leq f(x) \leq s'_n(x), \quad \forall\, n \text{ and } \forall\, x \in [a,b], \text{ and}$$
$$\int_a^b (s'_n - s_n) = \int_a^b \{s'_n(x) - s_n(x)\} dx \to 0 \quad \text{as } n \to \infty.$$

The *Riemann integral* $\int_a^b f = \int_a^b f(x)dx$ will then be the common limit of the integrals of s_n and s'_n; cf. [**99**]. For complex-valued functions one would separately consider the real and imaginary part.

In this book the statement: "f is integrable over (a, b)" (or $[a, b]$) shall mean that f is *integrable in the sense of Lebesgue* (after Henri Lebesgue, France, 1875–1941; [**76**]); notation: $f \in \mathcal{L}(a, b)$. For a Riemann integrable function on a finite interval the Lebesgue integral has the same value as the Riemann integral. However, the class of Lebesgue integrable functions is larger than the class of (properly) Riemann integrable functions, and that has certain advantages, for example, when it comes to termwise integration of infinite series; cf. Section 5.4. If a function has an improper Riemann integral on (a, b) that is absolutely convergent, it also has Lebesgue integral equal to the Riemann integral.

For an integrable function f on (a, b) and any $\varepsilon > 0$, there is a step function $s = s_\varepsilon$ on $[a, b]$ such that

$$\int_a^b |f(x) - s(x)| dx < \varepsilon. \tag{2.1.1}$$

[This holds also for unbounded intervals (a, b), but then one will require that s be equal to zero outside a finite subinterval.]

*For a *definition* of Lebesgue integrability one may start with the notion of a negligible set or *set of measure zero*. A set $E \subset \mathbb{R}$ has (Lebesgue) measure zero if for every $\varepsilon > 0$, it can be covered by a countable family of intervals of total length $< \varepsilon$. If a property holds everywhere on (a, b) outside a set of measure zero, one says that it holds *almost everywhere*, notation a.e., on (a, b). A real or complex function f on (a, b) will be Lebesgue integrable if there is a sequence of step functions $\{s_n\}$ that converges to f a.e. on (a, b), and is such that

$$\int_a^b |s_m - s_n| \to 0 \quad \text{as } m, n \to \infty.$$

The Lebesgue integral of f over (a, b) is then defined as

$$\int_a^b f = \lim_{n \to \infty} \int_a^b s_n;$$

cf. [**68**]. Using this approach, a subset E of $\mathbb{R}$ may be called (Lebesgue) measurable if its characteristic function h_E is integrable; the (Lebesgue) measure $m(E)$ is then given by the integral of h_E. [By definition, h_E has the value 1 on E and 0 outside E.] A function f may be called (Lebesgue) measurable if it is a pointwise limit a.e. of step functions.

*For a treatment of integration based on measure theory, which is essential in Mathematical Statistics, see for example [**77**].

We can now show that the Fourier coefficients $a_n = a_n[f]$, $b_n = b_n[f]$ and $c_n = c_n[f]$ of a periodic integrable function f [Section 1.5] tend to zero as $|n| \to \infty$. Likewise, the Fourier transform $g(\xi)$ of an integrable function f on $\mathbb{R}$ [Section 1.7] will tend to zero as $|\xi| \to \infty$.

LEMMA 2.1.1. (Riemann–Lebesgue) *Let f be integrable over (a, b) and let λ be a positive real parameter. Then as $\lambda \to \infty$,*

$$\int_a^b f(x)\cos\lambda x\,dx \to 0, \qquad \int_a^b f(x)\sin\lambda x\,dx \to 0,$$
$$\int_a^b f(x)e^{\pm i\lambda x}\,dx \to 0.$$

PROOF FOR $\int_a^b f(x)e^{i\lambda x}dx$. The integral will exist because the product of an integrable function and a bounded continuous function is integrable [see Integration Theory].

(i) We first consider the case where f is the characteristic function h_J of a finite subinterval J of (a, b) with end-points α and β. Clearly

$$\left|\int_a^b h_J(x)e^{i\lambda x}dx\right| = \left|\int_\alpha^\beta e^{i\lambda x}dx\right| = \left|\frac{e^{i\lambda\beta} - e^{i\lambda\alpha}}{i\lambda}\right| \le \frac{2}{\lambda} \to 0$$

as $\lambda \to \infty$.

(ii) Suppose now that f is a piecewise constant function s on (a, b) [which is different from zero only on a finite subinterval]. The function s can be represented as a finite linear combination of characteristic functions:

$$s(x) = \sum_{k=1}^{p} \gamma_k h_{J_k}(x), \quad J_k \in (a, b) \text{ finite.} \tag{2.1.2}$$

[Here some of the intervals J_k might reduce to a point.] Thus

$$\left|\int_a^b s(x)e^{i\lambda x}dx\right| = \left|\sum_{k=1}^{p} \gamma_k \int_{J_k} e^{i\lambda x}dx\right| \le \frac{2}{\lambda}\sum_{k=1}^{p} |\gamma_k| \to 0 \tag{2.1.3}$$

as $\lambda \to \infty$.

(iii) For arbitrary integrable f and given $\varepsilon > 0$, one first approximates f by a pieceweise constant function s such that inequality (2.1.1) is satisfied.

Representing s in the form (2.1.2), it now follows from (2.1.3) that

$$\left|\int_a^b f(x)e^{i\lambda x}dx\right| = \left|\int_a^b \{f(x)-s(x)\}e^{i\lambda x}dx + \int_a^b s(x)e^{i\lambda x}dx\right|$$
$$\leq \int_a^b |f(x)-s(x)|dx + \frac{2}{\lambda}\sum_{k=1}^p |\gamma_k| < 2\varepsilon \quad \text{provided} \;\; \lambda > \lambda_0.$$

□

How rapidly do Fourier coefficients or Fourier transforms tend to zero? That will depend on the degree of smoothness of f. Let $\mathcal{C}^p$ denote the class of p times continuously differentiable functions.

Lemma 2.1.2. *For some $p \geq 1$ let f be of class $\mathcal{C}^p_{2\pi}$, that is, f has period 2π and is of class $\mathcal{C}^p$ on $\mathbb{R}$* (not just on $[-\pi,\pi]$!). *Then*

$$c_n[f'] = inc_n[f], \; \cdots, \; c_n[f^{(p)}] = (in)^p c_n[f],$$

so that

$$|c_n[f]| \leq \frac{\sup|f^{(p)}|}{|n|^p} \quad \text{and} \;\; n^p c_n[f] \to 0 \;\; \text{as} \;\; |n| \to \infty.$$

Proof. Integration by parts shows that

$$c_n[f'] = \frac{1}{2\pi}\int_{-\pi}^{\pi} f'(x)e^{-inx}dx = \frac{1}{2\pi}\left[f(x)e^{-inx}\right]_{-\pi}^{\pi}$$
$$+\, in\,\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx = inc_n[f]. \tag{2.1.4}$$

Indeed, the integrated term drops out by the periodicity of f. For $p \geq 2$ one will use repeated integration by parts. The inequality for $c_n[f]$ now follows from a straightforward estimate of the integral for $c_n[f^{(p)}]$. The final result follows from Lemma 2.1.1 applied to $f^{(p)}$. □

Remarks 2.1.3. The first result in Lemma 2.1.2 may be expressed by saying that for $f \in \mathcal{C}^1_{2\pi}$, the complex Fourier series for f' may be obtained from that of f by *termwise differentiation*:

$$f(x) \sim \sum c_n e^{inx} \quad\Longrightarrow\quad f'(x) \sim \sum inc_n[f]e^{inx}. \tag{2.1.5}$$

[Recall that the symbol $\sim$ is to be read as "has the Fourier series".] The implication (2.1.5) is independent of questions of convergence. Similarly, for the "real" Fourier series of $f \in \mathcal{C}^1_{2\pi}$,

$$f(x) \sim \frac{1}{2}a_0[f] + \sum_{n=1}^{\infty}(a_n[f]\cos nx + b_n[f]\sin nx)$$

$$\Longrightarrow f'(x) \sim \sum_{n=1}^{\infty}(nb_n[f]\cos nx - na_n[f]\sin nx). \tag{2.1.6}$$

The computation (2.1.4) is valid whenever f is in $\mathcal{C}_{2\pi}$ *and* can be written as *an indefinite integral* of its derivative, which we suppose to be integrable:

$$f(x) = f(0) + \int_0^x f'(t)dt. \tag{2.1.7}$$

Representation (2.1.7) will in particular hold if f is continuous and piecewise smooth.

For Fourier analysis, an important class of functions is given by the so-called functions of *bounded variation*, or finite total variation:

DEFINITION 2.1.4. The total variation $V = V_f[a,b]$ of a function f on a (finite) closed interval $[a,b]$ is defined as

$$V = \sup_{a=x_0<x_1<\cdots<x_p=b}\{|f(x_1 - f(x_0)| + |f(x_2 - f(x_1)| + \cdots$$
$$+ |f(x_p - f(x_{p-1})|\}, \tag{2.1.8}$$

where the supremum is taken over all finite partitionings of $[a,b]$.

Simple examples of functions of bounded variation are given by bounded monotonic functions, and by "indefinite integrals" (2.1.7) of integrable functions on a finite interval. Total variation is additive: if $a < c < b$ then $V_f[a,b] = V_f[a,c] + V_f[c,b]$. One may deduce that $\phi(x) = V_f[a,x]$ is continuous wherever $f(x)$ is. A real function f of bounded variation can be represented as the difference of two increasing (nondecreasing) functions on $[a,b]$, whose total variations add up to that of f:

$$f = \frac{1}{2}(\phi + f) - \frac{1}{2}(\phi - f).$$

In particular a function f of bounded variation has a finite right-hand and left-hand limit at every point $c \in [a,b)$, and $c \in (a,b]$, respectively. For finite $[a,b]$ such a function will be Riemann integrable.

Exercises. 2.1.1. Prove the implication (2.1.6) for piecewise smooth $f \in \mathcal{C}_{2\pi}$.

2.1.2. Let f be in $\mathcal{C}^2_{2\pi}$ or at least, $f \in \mathcal{C}^1_{2\pi}$ with f' piecewise smooth. Prove that the Fourier coefficients of f are $\mathcal{O}(1/n^2)$, and deduce that the Fourier series for f is uniformly convergent.

2.1.3. Verify the following Fourier series for functions on $(-\pi, \pi]$:

$$x \sim 2\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sin nx, \quad x^2 \sim \frac{1}{3}\pi^2 + 4\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nx.$$

Explain why the Fourier coefficients for $f(x) = x^2$ tend to zero faster than those for $g(x) = x$. [Both functions are infinitely differentiable on $(-\pi, \pi]$!

2.1.4. Compute (or estimate) the total variation $V_f[-\pi, \pi]$ for (i) a monotonic function; (ii) an indefinite integral; (iii) the functions $\cos nx$, $\sin nx$, e^{inx}; (iii) an arbitrary $\mathcal{C}^1$ function f.

2.1.5. For f of bounded variation on $[a, b]$ set $\phi(x) = V_f[a, x]$. Assuming f real, show that $\phi + f$ and $\phi - f$ are nondecreasing.

2.1.6. Show that the set of discontinuities of a function of bounded variation is countable.

2.1.7. Let f be of finite total variation V on $[a, b]$ and let g be of class $\mathcal{C}^1[a, b]$. Prove that

$$\left| \int_a^b f g' \right| \le V \sup |g| + |f(b)g(b) - f(a)g(a)|.$$

Hint. It will be sufficient to prove the result for piecewise constant functions f.

2.1.8. Let f be a 2π-periodic function of finite total variation V on $[-\pi, \pi]$. Prove that

$$|c_n[f]| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \right| \le \frac{1}{2\pi} \frac{V}{|n|}, \quad \forall\, n \ne 0. \tag{2.1.9}$$

Obtain corresponding inequalities for the Fourier coefficients $a_n[f]$ and $b_n[f]$. Are the inequalities sharp?

2.2. Partial sum formula. Dirichlet kernel

Let f be an integrable function on $(-\pi, \pi)$; we extend f to a 2π-periodic function on $\mathbb{R}$. As in Section 1.5 the Fourier coefficients for f are denoted by $a_n = a_n[f]$ etc., and the k-th partial sum of the Fourier series is called

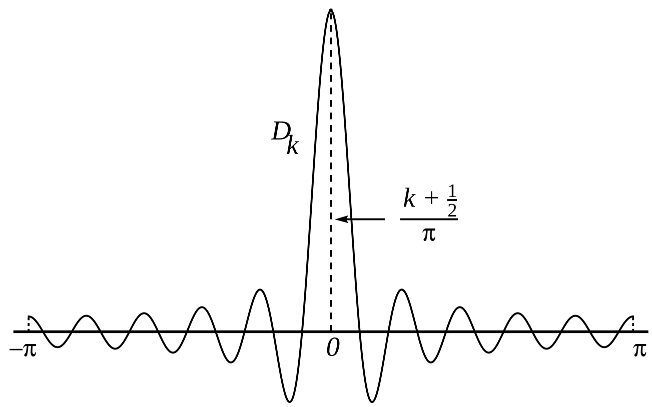

FIGURE 2.1

s_k or $s_k[f]$:

$$
\begin{aligned}
s_k[f](x) &= \frac{1}{2}a_0[f] + \sum_{n=1}^{k} (a_n[f]\cos nx + b_n[f]\sin nx) \\
&= \sum_{n=-k}^{k} c_n[f]\, e^{inx}.
\end{aligned}
$$

We will express $s_k(x)$ as an integral by substituting the defining integrals for the complex Fourier coefficients c_n, this time using variable of integration u in order to keep t in reserve for $x-u$:

$$
\begin{aligned}
s_k(x) &= \sum_{n=-k}^{k} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u)e^{-inu}du \cdot e^{inx} \\
&= \int_{-\pi}^{\pi} f(u)\, \frac{1}{2\pi} \sum_{n=-k}^{k} e^{in(x-u)}\, du \\
&= \int_{x-\pi}^{x+\pi} f(x-t)\, \frac{1}{2\pi} \sum_{n=-k}^{k} e^{int}\, dt. \qquad (2.2.1)
\end{aligned}
$$

The "kernel" by which $f(x-t)$ is multiplied is called the *Dirichlet kernel*. It will be expressed in closed form below, and is illustrated in Figure 2.1.

LEMMA 2.2.1. *For $k = 0, 1, 2, \cdots$ and all $t \in \mathbb{R}$,*

$$D_k(t) \stackrel{\text{def}}{=} \frac{1}{2\pi} \sum_{n=-k}^{k} e^{int} = \frac{\sin(k+\frac{1}{2})t}{2\pi \sin \frac{1}{2}t}. \tag{2.2.2}$$

[Here the right-hand side is defined by its limit value at the points $t = 2\nu\pi$.]

The proof follows from the sum formula for geometric series:

$$\begin{aligned}\sum_{n=-k}^{k} e^{int} &= e^{-ikt}(1 + e^{it} + \cdots + e^{2kit}) = e^{-ikt}\,\frac{e^{(2k+1)it} - 1}{e^{it} - 1} \\ &= \frac{e^{(k+\frac{1}{2})it} - e^{-(k+\frac{1}{2})it}}{e^{\frac{1}{2}it} - e^{-\frac{1}{2}it}} = \frac{2i \sin(k+\frac{1}{2})t}{2i \sin \frac{1}{2}t}.\end{aligned}$$

LEMMA 2.2.2. *The kernel D_k is even and has period 2π, and*

$$\int_{-\pi}^{\pi} D_k(t)dt = 1, \quad \forall\, k \in \mathbb{N}_0.$$

The proof follows from the definition of D_k; note that $\int_{-\pi}^{\pi} e^{int}dt = 0$ for all $n \neq 0$.

THEOREM 2.2.3. (Dirichlet) *Let f be 2π-periodic and integrable over a period. Then the partial sums of the Fourier series for f are equal to the following integrals involving the kernel D_k:*

$$\begin{aligned}s_k[f](x) &= \int_{-\pi}^{\pi} f(x-t)D_k(t)dt = \int_{-\pi}^{\pi} f(x+t)D_k(t)dt \\ &= \int_{-\pi}^{\pi} \frac{f(x+t) + f(x-t)}{2}\, D_k(t)dt, \quad \forall\, k \in \mathbb{N}_0.\end{aligned} \tag{2.2.3}$$

PROOF. For fixed x the function $g(t) = g(t,x) = f(x-t)D_k(t)$ is integrable and has period 2π. Thus the integral of g over every interval of length 2π has the same value. Formulas (2.2.1) and (2.2.2) now show that

$$\begin{aligned}s_k(x) &= \int_{x-\pi}^{x+\pi} f(x-t)D_k(t)dt = \int_{-\pi}^{\pi} f(x-t)D_k(t)dt \\ &= -\int_{\pi}^{-\pi} f(x+v)D_k(-v)dv = \int_{-\pi}^{\pi} f(x+v)D_k(v)dv\end{aligned} \tag{2.2.4}$$

because D_k is even. The final formula in (2.2.3) follows by averaging. $\square$

Exercises. 2.2.1. Let $f(x) = e^{imx}$ with $m \in \mathbb{N}$. Determine $s_k[f](x)$ for $k = 0, 1, 2, \cdots$: (i) directly from the Fourier series; (ii) by formula (2.2.3).

2.2.2. Same questions for $g(x) = \cos mx$.

2.2.3. Compute

$$\int_0^\pi x \, \frac{\sin(k+\frac{1}{2})x}{\sin\frac{1}{2}x} \, dx,$$

and determine the limit as $k \to \infty$.

2.2.4. (A function f bounded by 1 with some large partial sums.) Define

$$f(x) = \sin\left(p + \frac{1}{2}\right)|x|, \quad -\pi \le x \le \pi \qquad (p \in \mathbb{N}).$$

Prove that there is a constant C (independent of p) such that

$$s_p[f](0) > \frac{1}{\pi}\log p - C, \quad |s_k[f](0)| \le C \text{ whenever } |k-p| \ge \frac{1}{2}p.$$

Hint. Using the inequality

$$\left|\frac{1}{\sin\frac{1}{2}t} - \frac{1}{\frac{1}{2}t}\right| \le C_1 \quad \text{on } [-\pi, \pi],$$

one can show that

$$\pi s_p(0) = \int_0^\pi \frac{\sin^2(p+\frac{1}{2})t}{\sin\frac{1}{2}t}\, dt > \int_0^\pi \frac{1-\cos(2p+1)t}{t}\, dt - C_2.$$

2.3. Theorems on pointwise convergence

Let f be an integrable 2π-periodic function. For given x one obtains a useful integral for the difference $s_k[f](x) - f(x)$ by using the second integral (2.2.3) for $s_k(x)$ and writing $f(x)$ as $\int_{-\pi}^{\pi} f(x) D_k(t) dt$:

$$\begin{aligned} s_k(x) - f(x) &= \int_{-\pi}^{\pi} \{f(x+t) - f(x)\} D_k(t) dt \\ &= \int_{-\pi}^{\pi} \frac{f(x+t)-f(x)}{2\pi \sin\frac{1}{2}t} \sin\left(k+\frac{1}{2}\right) t \, dt. \end{aligned} \tag{2.3.1}$$

Keeping x fixed, it is natural to introduce the auxiliary function

$$\phi(t) = \phi(t, x) \stackrel{\text{def}}{=} \frac{f(x+t)-f(x)}{2\pi\sin\frac{1}{2}t}, \quad t \ne 2\nu\pi. \tag{2.3.2}$$

If $\phi(t)$ would be integrable over $(-\pi, \pi)$, formula (2.3.1) and the Riemann–Lebesgue Lemma 2.1.1 would immediately show that $s_k(x) - f(x) \to 0$ or

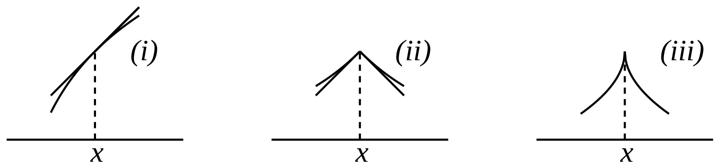

FIGURE 2.2

$s_k(x) \to f(x)$ as $k \to \infty$. In other words, the Fourier series for f at the point x would then converge to the value $f(x)$.

The difficulty is that $\phi(t)$ has a singularity at the point $t = 0$. The question is whether the difference $f(x+t) - f(x)$ is small enough for t close to 0 to make $\phi(t)$ integrable. A simple sufficient condition would be the following:

$$f \in \mathcal{C}_{2\pi} \ \ \textit{and} \ \ f \ \text{ is differentiable at the point } \ x.$$

Indeed, in that case one can make $\phi(t)$ continuous on $[-\pi, \pi]$ by defining

$$\phi(0) = \lim_{t \to 0} \phi(t) = f'(x)/\pi.$$

We will see that for the integrability of $\phi(t)$, a "Hölder–Lipschitz condition" (after Otto Hölder, Germany, 1859–1937; [**51**] and Rudolf Lipschitz, Germany, 1832–1903; [**83**]) on f will suffice.

DEFINITION 2.3.1. One says that f satisfies a Hölder–Lipschitz condition at the point x if there exist positive constants M, α and δ such that

$$|f(x+t) - f(x)| \leq M|t|^{\alpha} \quad \text{for} \quad -\delta < t < \delta. \tag{2.3.3}$$

THEOREM 2.3.2. *Let f be 2π-periodic and integrable over a period. Then each of the following conditions is sufficient for the convergence of the Fourier series for f at the point x to the value $f(x)$:*

(i) *f is differentiable at the point x;*

(ii) *f is continuous at x and has a finite right-hand and left-hand derivative at x:*

$$\lim_{t \searrow 0} \frac{f(x+t) - f(x)}{t} = f'_{+}(x), \quad \lim_{t \nearrow 0} \frac{f(x+t) - f(x)}{t} = f'_{-}(x);$$

(iii) *f satisfies a Hölder–Lipschitz condition at the point x.*

Cf. Figure 2.2.

PROOF. Since (i) and (ii) imply (iii) with $\alpha = 1$, it is sufficient to deal with the latter case. Let $\varepsilon > 0$ be given. Observe that $|\sin \frac{1}{2}t| \geq |t|/\pi$ for $|t| \leq \pi$, and take $\delta \leq \pi$. Then inequality (2.3.3) shows that for all k,

$$\begin{aligned} &\left| \int_{-\delta}^{\delta} \{f(x+t) - f(x)\} \frac{\sin(k+\frac{1}{2})t}{2\pi \sin \frac{1}{2}t}\, dt \right| \\ &\leq \int_{-\delta}^{\delta} \frac{M|t|^\alpha}{2\pi \sin \frac{1}{2}t}\, dt \leq \frac{1}{2} M \int_{-\delta}^{\delta} |t|^{\alpha-1} dt = \frac{M}{\alpha}\, \delta^\alpha. \end{aligned} \tag{2.3.4}$$

We may decrease δ until the final bound is $\leq \varepsilon$. Keeping δ fixed from here on, we note that $\phi(t)$ in (2.3.2) is integrable over $[-\pi, -\delta]$ and $[\delta, \pi]$: it is the quotient of an integrable function and a continuous function that stays away from zero. Thus by the Riemann–Lebesgue Lemma 2.1.1,

$$\left| \left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} \right) \phi(t) \sin\left(k + \frac{1}{2}\right) t\, dt \right| < \varepsilon \tag{2.3.5}$$

for all k larger than some k_0. Combination of (2.3.1), (2.3.4) and (2.3.5) shows that $|s_k(x) - f(x)| < 2\varepsilon$ for all $k > k_0$. □

EXAMPLES 2.3.3. The Fourier series for the 2π-periodic function $f(x)$ equal to x on $(-\pi, \pi]$ (Exercise 2.1.3) will converge to $f(x)$ at each point $x \in (-\pi, \pi)$ [but not at $x = \pi$]. See condition (i) of the Theorem.

The Fourier series for the 2π-periodic function $f(x)$ equal to x^2 on $(-\pi, \pi]$ (Exercise 2.1.3) will converge to $f(x)$ for all x. At $x = \pm\pi$, condition (ii) of the Theorem is satisfied.

The Fourier series for the 2π-periodic function $f(x)$ equal to 0 on $(-\pi, 0)$ and to $\sqrt{x}$ on $[0, \pi]$ will converge to $f(x)$ on $(-\pi, \pi)$: at $x = 0$, condition (iii) of the Theorem is satisfied.

We can also deal with functions that have simple discontinuities.

THEOREM 2.3.4. *Let f be an integrable 2π-periodic function. Suppose that f has a right-hand limit $f(x+)$ and a left-hand limit $f(x-)$ at the point x, and in addition suppose that there are positive constants M, α and δ such that*

$$|f(x+t) - f(x+)| \leq M|t|^\alpha \ \ \textit{and} \ \ |f(x-t) - f(x-)| \leq M|t|^\alpha \tag{2.3.6}$$

for $0 < t < \delta$. Then the Fourier series for f converges at the point x to the average of the right-hand and the left-hand limit, $\frac{1}{2}\{f(x+) + f(x-)\}$.

Cf. Figure 2.3.

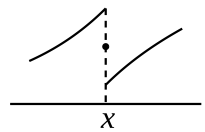

FIGURE 2.3

PROOF. By the final integral for $s_k[f](x)$ in (2.2.3),

$$s_k(x) - \frac{f(x+) + f(x-)}{2} = \int_{-\pi}^{\pi} \left\{ \frac{f(x+t) + f(x-t)}{2} - \frac{f(x+) + f(x-)}{2} \right\} D_k(t)dt.$$

For our fixed x we now define a 2π-periodic function g such that

$$g(t) = \frac{f(x+t) + f(x-t)}{2}, \quad 0 < |t| \leq \pi, \quad g(0) = \frac{f(x+) + f(x-)}{2}.$$

Then

$$s_k[f](x) - \frac{f(x+) + f(x-)}{2} = s_k[g](0) - g(0),$$

and the function g will satisfy condition (iii) of Theorem 2.3.2 at the point $t = 0$. Hence $s_k[g](0) \to g(0)$ as $k \to \infty$, which implies the desired result. $\square$

EXAMPLE 2.3.5. By inspection, the Fourier series for the 2π-periodic function $f(x)$ equal to x on $(-\pi, \pi]$ (Exercise 2.1.3) converges to 0 at the point π. This is precisely the average of $f(\pi+) = f(-\pi+) = -\pi$ and $f(\pi-) = f(\pi) = \pi$.

Exercises. 2.3.1. Let $f(x) = 0$ for $-\pi < x < 0$, $= 1$ for $0 \leq x \leq \pi$. Determine the Fourier series for f. Where on $\mathbb{R}$ does it converge? Give a precise description of the sum function on $\mathbb{R}$. Which theorems have you used?

2.3.2. Same questions about the complex Fourier series for $f(x) = e^{\alpha x}$ on $[0, 2\pi)$.

2.4. Uniform convergence

In some cases one can establish uniform convergence of a Fourier series by studying the coefficients. Thus by Cauchy's criterion [Section 1.2], the convergence of one of the series

$$\sum_{n=1}^{\infty}(|a_n[f]|+|b_n[f]|) \quad \text{or} \quad \sum_{n=-\infty}^{\infty}|\,c_n[f]|$$

implies uniform convergence of the Fourier series for f. Indeed, for $k>j$,

$$\bigl|s_k[f](x)-s_j[f](x)\bigr| \le \sum_{n=j+1}^{k}(|a_n[f]|+|b_n[f]|).$$

Partial summation is another useful tool. For example, the Fourier series $\sum_{n=1}^{\infty}\frac{1}{n}\sin nx$ for the 2π-periodic function $f(x)$, equal to $\frac{1}{2}(\pi-x)$ on $[0,2\pi)$, is uniformly convergent on $[\delta,2\pi-\delta]$ for every $\delta\in(0,\pi)$; cf. Section 1.2. We will see that such results on locally uniform convergence can be obtained directly from properties of the function f with the aid of the partial sum formula. We begin with a refinement of the Riemann–Lebesgue Lemma that involves uniform convergence.

Lemma 2.4.1. *Let f be 2π-periodic and integrable over a period. Then for any continuous function g on $(0,2\pi)$ and $\delta\in(0,\pi)$,*

$$I_f(x,\lambda)=\int_{\delta}^{2\pi-\delta}f(x+t)g(t)\sin\lambda t\,dt\to 0 \quad \text{as} \ \ \lambda\to\infty, \tag{2.4.1}$$

uniformly in x.

In the application below we will take $g(t)=1/(2\pi\sin\frac{1}{2}t)$.

Proof of the lemma. As in the case of Lemma 2.1.1, the proof may be reduced to the case where f is (the periodic extension of) the characteristic function h_J of an interval.

(i) To given $\varepsilon>0$ we choose a piecewise constant function s on $(-\pi,\pi)$ such that $\int_{-\pi}^{\pi}|f(t)-s(t)|dt<\varepsilon$. Extending s to a 2π-periodic function, we may conclude that

$$\int_0^{2\pi}|f(x+t)-s(x+t)|dt<\varepsilon, \quad \forall\, x\in\mathbb{R}.$$

It now follows that for all x and λ,

$$\left|\int_{\delta}^{2\pi-\delta}\{f(x+t)-s(x+t)\}g(t)\sin\lambda t\,dt\right|\le\varepsilon\sup_{[\delta,2\pi-\delta]}|g(t)|.$$

(ii) On $(-\pi,\pi)$, s is a finite linear combination $\sum_{k=1}^{p}\gamma_k h_{J_k}$, where the intervals J_k are nonoverlapping subintervals of $(-\pi,\pi)$. In order to prove that $I_s(x,\lambda) \to 0$ uniformly in x as $\lambda \to \infty$, it is sufficient to consider the case where $s = h_J$ on $(-\pi,\pi)$, so that on $\mathbb{R}$, our function s is equal to the periodic extension $\tilde{h}_J$ of h_J.

(iii) For $J = (\alpha,\beta) \in (-\pi.\pi)$, $h_J(x+t)$ is the characteristic function of the t-interval $\alpha - x < t < \beta - x$. Thus $\tilde{h}_J(x+t)$ is the characteristic function of the union of the t-intervals $\alpha - x + \nu \cdot 2\pi < t < \beta - x + \nu \cdot 2\pi$, $\nu \in \mathbb{Z}$. This union will intersect $(\delta, 2\pi - \delta)$ in one or perhaps two intervals $(\alpha(x), \beta(x))$. Thus

$$I_{\tilde{h}_J}(x,\lambda) = \int_\delta^{2\pi-\delta} \tilde{h}_J(x+t) g(t) \sin \lambda t \, dt = \int_{\alpha(x)}^{\beta(x)} g(t) \sin \lambda t \, dt,$$

plus perhaps another integral like it. Since g is continuous on $[\delta, 2\pi - \delta]$ we may approximate it by a piecewise constant function there, etc. The argument of the proof for Lemma 2.1.1 now readily shows that $I_{\tilde{h}_J}(x,\lambda) \to 0$ uniformly in x as $\lambda \to \infty$. $\square$

We still need an interval analog to the Hölder–Lipschitz condition.

Definition 2.4.2. A function f on $[a,b]$ is said to be Hölder continuous if there are positive constants M and α such that

$$|f(x') - f(x'')| \le M|x' - x''|^\alpha, \quad \forall\, x',\, x'' \in [a,b]. \tag{2.4.2}$$

Theorem 2.4.3. *Let f be 2π-periodic, integrable over a period, and Hölder continuous on $[a,b]$. Then the Fourier series for f converges to f uniformly on every interval $[a+r, b-r]$ with $0 < r < \frac{1}{2}(b-a)$.* [Thus if $b - a > 2\pi$, the Fourier series will by periodicity converge uniformly on $\mathbb{R}$.]

Proof. Restricting x to $[a+r, b-r]$ and taking $0 < \delta \le r$, we split the integral for $s_k(x) - f(x)$ as follows:

$$\begin{aligned} s_k(x) - f(x) &= \int_{-\delta}^{2\pi-\delta} \frac{f(x+t) - f(x)}{2\pi \sin \frac{1}{2} t} \sin\left(k + \frac{1}{2}\right) t \, dt \\ &= \int_{-\delta}^{\delta} + \int_\delta^{2\pi-\delta} = I_1(x,k) + I_2(x,k), \end{aligned}$$

say. Here by the Hölder continuity of f on $[a,b]$,

$$|I_1(x,k)| \le \int_{-\delta}^{\delta} \frac{M|t|^\alpha}{2|t|} \, dt = M \frac{\delta^\alpha}{\alpha}.$$

To given $\varepsilon > 0$, we now decrease δ until $M\delta^\alpha/\alpha < \varepsilon$. Keeping δ fixed from here on and setting $B = \sup|f(x)|$ on $[a, b]$, we estimate I_2:

$$\begin{aligned}|I_2(x,k)| &\leq \left|\int_\delta^{2\pi-\delta} \frac{f(x+t)}{2\pi \sin \frac{1}{2}t} \sin\left(k+\frac{1}{2}\right) t\, dt\right| \\ &\quad + \left|f(x) \int_\delta^{2\pi-\delta} \frac{1}{2\pi \sin \frac{1}{2}t} \sin\left(k+\frac{1}{2}\right) t\, dt\right| \\ &\leq \left|I_f\left(x, k+\frac{1}{2}\right)\right| + B\left|I_1\left(x, k+\frac{1}{2}\right)\right|,\end{aligned}$$

where I_f is as in (2.4.1) with $g(t) = 1/(2\pi \sin \frac{1}{2}t)$ and I_1 refers to $f \equiv 1$. By Lemma 2.4.1 we can choose k_0 such that $|I_f(x, k+\frac{1}{2})| < \varepsilon$ and $|I_1(x, k+\frac{1}{2})| < \varepsilon$ for all $k > k_0$ and all x. Thus in conclusion,

$$|s_k(x) - f(x)| < (2+B)\varepsilon, \quad \forall\, k > k_0 \text{ and } \forall\, x \in [a+r, b-r].$$

$\square$

EXAMPLES 2.4.4. The functions in Examples 2.3.3 are Hölder continuous on $(-\pi, \pi]$, hence their Fourier series will converge uniformly on $[-\pi + r, \pi - r]$ for $0 < r < \pi$. The series for $f(x) = x^2$ on $(-\pi, \pi]$ will converge uniformly on $\mathbb{R}$.

2.5. Divergence of certain Fourier series

Dirichlet proved around 1830 that the Fourier series of every bounded *monotonic* [or piecewise monotonic] function on $(-\pi, \pi]$ is *everywhere convergent*. The same will be true for linear combinations of bounded monotonic functions. It then came as a surprise to the mathematical world when Paul du Bois-Reymond (Germany, 1831–1889; [**8**]) showed that there exist oscillating *continuous* functions whose Fourier series *diverge* at certain points. The basic reason is that

$$\int_{-\pi}^{\pi} |D_k(t)| dt \to \infty \quad \text{as } k \to \infty.$$

Now for special continuous functions f with $\sup|f| \leq 1$, there is a resonance between f and certain D_p's, and this results in large corresponding values of $s_p[f]$.

EXAMPLE 2.5.1. A simple building block for a "bad function" f is given by

$$f_p(x) = \sin\left(p+\frac{1}{2}\right)|x|, \quad -\pi \le x \le \pi. \tag{2.5.1}$$

As was indicated in Exercise 2.2.4 there is an absolute constant C such that

$$s_p[f_p](0) = \frac{1}{\pi}\int_0^\pi \frac{\sin^2(p+\frac{1}{2})t}{\sin\frac{1}{2}t}\,dt > \frac{1}{\pi}\log p - C, \tag{2.5.2}$$

while for $|q-p| \ge \frac{1}{2}p$,

$$\big|s_q[f_p](0)\big| = \left|\frac{1}{\pi}\int_0^\pi \frac{\sin(p+\frac{1}{2})t\,\sin(q+\frac{1}{2})t}{\sin\frac{1}{2}t}\,dt\right| \le C. \tag{2.5.3}$$

Let us now consider functions f of the form

$$f(x) \stackrel{\text{def}}{=} \sum_{j=1}^\infty \frac{1}{j(j+1)}\, f_{p_j}(x), \quad |x| \le \pi, \tag{2.5.4}$$

where $\{p_j\}$ is a rapidly increasing sequence of positive integers such that $|p_k - p_j| \ge \frac{1}{2}p_j$ whenever $k \ne j$. For this it suffices to take $p_{j+1} \ge 2p_j$ for all j.

By uniform convergence, functions f as in (2.5.4) are continuous and clearly $\sup|f| \le 1$. Also by uniform convergence, and using (2.5.2), (2.5.3),

$$\begin{aligned} s_{p_k}[f](0) &= \sum_{j=1}^\infty \frac{1}{j(j+1)}\, s_{p_k}[f_{p_j}](0) \\ &\ge \frac{1}{k(k+1)}\, s_{p_k}[f_{p_k}](0) - \sum_{j\ne k} \frac{1}{j(j+1)}\,\big|s_{p_k}[f_{p_j}](0)\big| \\ &\ge \frac{1}{k(k+1)}\left(\frac{1}{\pi}\log p_k - C\right) - \sum_{j\ne k}\frac{1}{j(j+1)}\,C = \frac{\log p_k}{\pi k(k+1)} - C. \end{aligned} \tag{2.5.5}$$

Conclusion. Suppose we take $p_j = 2^{j^3}$ for all $j \in \mathbb{N}$. Then

$$s_{p_k}[f](0) \ge \frac{\log 2}{2\pi}\,k - C \to \infty \quad \text{as } k \to \infty. \tag{2.5.6}$$

In this case, the Fourier series for the continuous function f *fails to converge* at the point 0.

Combining functions f with different “critical points” and different sequences $\{p_j\}$, one can construct continuous functions whose Fourier series diverge on arbitrary finite sets of points. It is much more difficult to prove, as Andrey Kolmogorov (Russia, 1903–1987; [**65**]) did close to 1930, that there exist integrable functions whose Fourier series diverge *everywhere*. For a long time, it was an open question whether such divergence can occur in the case of continuous functions. Finally, in 1966, Lennart Carleson (Sweden, born 1928; [**11**]), proved that for a *continuous* function, the Fourier series converges to the function *almost everywhere*, that is, the exceptional set must have measure zero.

After du Bois-Reymond, one gradually realized that ordinary pointwise convergence is not the most suitable concept of convergence in Fourier analysis. In Chapter 3 we will see that certain summability methods are effective for all continuous functions. Even more important is the concept of convergence in the mean of order two. In Chapter 6 it will be shown that

$$\int_{-\pi}^{\pi} \left|f(x) - s_k[f](x)\right|^2 dx \to 0 \quad \text{as } \ k \to \infty$$

for all continous, and in fact, for all square-integrable functions f. Finally, for integrable functions and for so-called distributions or generalized functions, one may use weak or distributional convergence to sum the Fourier series (Chapter 4).

Exercises. 2.5.1. Prove successively that

$$\int_0^{\pi} \left(\frac{1}{\pi t} - \frac{1}{2\pi \sin \frac{1}{2} t}\right) \sin\left(k + \frac{1}{2}\right) t \, dt \to 0 \quad \text{as } \ k \to \infty;$$

$$\int_0^{(k+\frac{1}{2})\pi} \frac{\sin x}{\pi x} \, dx = \int_0^{\pi} \frac{\sin(k+\frac{1}{2})t}{\pi t} \, dt \to \frac{1}{2} \quad \text{as } \ k \to \infty;$$

$$\int_0^{\infty-} \frac{\sin x}{x} \, dx \stackrel{\text{def}}{=} \lim_{A\to\infty} \int_0^{A} \frac{\sin x}{x} \, dx = \frac{1}{2}\pi; \tag{2.5.7}$$

and that there is an absolute constant C such that for all k and x,

$$\left|\int_0^{x} D_k(t) dt\right| \leq C.$$

2.5.2. Compute

$$\lim_{A\to\infty} \int_0^1 e^x \, \frac{\sin Ax}{x} \, dx.$$

2.5.3. Let $f(x) = |x|$ for $-\pi \le x \le \pi$. Describe the sum function of the Fourier series on $\mathbb{R}$ and prove that the Fourier series is uniformly convergent.

2.5.4. Same questions for $f(x) = (\pi^2 - x^2)^{\frac{1}{2}}$, $-\pi \le x \le \pi$.

2.5.5. Develop each of the following functions on $(0, \pi)$ into a Fourier sine series as well as a Fourier cosine series. Which of the two series converges faster? Explain by using properties of the functions and their appropriate extensions. Describe the sum functions on $\mathbb{R}$. Indicate intervals of uniform and non-uniform convergence:

$$\begin{aligned} &(i)\ \ f(x) = 1; &&(ii)\ \ f(x) = 1 \ \text{ on } (0, c), \ = 0 \ \text{ on } (c, \pi); \\ &(iii)\ \ f(x) = x; &&(iv)\ \ f(x) = \sin \alpha x. \end{aligned}$$

2.5.6. (*Principle of localization*) Let f and g be 2π-periodic and integrable over a period. Suppose that $f(x) = g(x)$ on (a, b). Prove that the Fourier series for f and g are either both convergent at $c \in (a, b)$ (to the same sum), or both divergent. Finally prove that the Fourier series for f is uniformly convergent on $[a + r, b - r] \subset (a, b)$ whenever the series for g is.

2.5.7. Let f be in $\mathcal{C}^1_{2\pi}$ and set $\sup |f'(x)| = M$. Prove that $|s_k(x) - f(x)| \le 8M/\sqrt{k}$ for all k.

Hint. Choose a good δ to treat the integral $\int_{-\delta}^{\delta}$.

2.5.8. Let f be 2π-periodic and of finite total variation over a period. Prove that the Fourier series for f converges to $\frac{1}{2}\{f(x+) + f(x-)\}$ at every point x.

Hint. If f is continuous at the point x one has $V_f[x, x+t] \to 0$ as $t \searrow 0$.

2.5.9. Show that the Fourier series for a continuous 2π-periodic function, which is of bounded variation on $[-\pi, \pi]$, is uniformly convergent.

2.5.10. Prove that for $k \to \infty$,

$$\int_{-\pi}^{\pi} |D_k(t)| dt \sim \frac{2}{\pi} \int_0^{(k+\frac{1}{2})\pi} \frac{|\sin u|}{u}\, du \sim \frac{4}{\pi^2} \log k.$$

Here the symbol $\sim$ stands for "asymptotically equal". Two functions are called asymptotically equal if their quotient has limit 1.

2.6. The Gibbs phenomenon

Here we will discuss a remarkable example of non-uniform convergence that was first analyzed by Henry Wilbraham (England) around 1850, and later studied by the physicist J. Willard Gibbs (USA, 1839–1903; [**38**]). We begin by discussing the *integral sine function* (Figure 2.4):

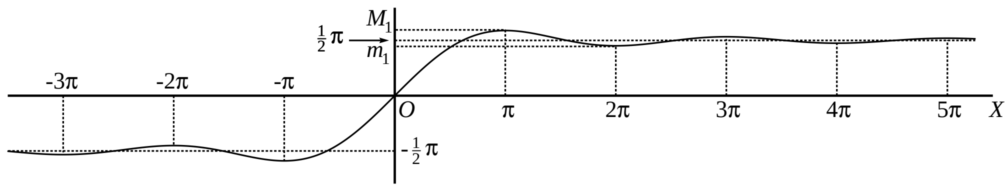

FIGURE 2.4

$$\mathrm{Si}(x) \stackrel{\text{def}}{=} \int_0^x \frac{\sin t}{t}\,dt, \quad x \in \mathbb{R}. \tag{2.6.1}$$

It is clear that the integral sine is odd. Calculus shows that it has relative maxima

$$M_1 \approx 1.85 > M_2 > \cdots \quad \text{at the points } \pi,\, 3\pi,\, \cdots$$

and relative minima

$$m_1 \approx 1.41 < m_2 < \cdots \quad \text{at the points } 2\pi,\, 4\pi,\, \cdots.$$

Moreover as $k \to \infty$,

$$\begin{aligned} \lim \mathrm{Si}\left\{\left(k+\frac{1}{2}\right)\pi\right\} &= \lim \int_0^\pi \frac{\sin(k+\frac{1}{2})t}{t}\,dt \\ &= \lim \int_0^\pi \frac{\sin(k+\frac{1}{2})t}{2\sin\frac{1}{2}t}\,dt = \frac{1}{2}\,\pi; \end{aligned}$$

cf. Exercise 2.5.1. From this it readily follows that

$$\lim \mathrm{Si}(k\pi) = \frac{1}{2}\,\pi, \quad \lim_{x\to\infty} \mathrm{Si}(x) = \frac{1}{2}\,\pi. \tag{2.6.2}$$

Gibbs phenomenon, cf. [**39**]. Let f be a bounded piecewise monotonic function on $(-\pi,\pi)$ whose periodic extension has a jump discontinuity at the point x_0: $f(x_0+) \neq f(x_0-)$. Then the Fourier series for f cannot converge uniformly in a neighborhood of x_0 [why not?]. The non-uniformity is of an interesting type. For x close to x_0, the graph of the partial sums $s_k[f]$ exhibits oscillations around the graph of f which approach the vertical through x_0 as $k \to \infty$, but whose amplitudes approach nonzero limits. Normalizing f so that x_0 becomes 0 and the jump at 0 becomes equal to π, the differences $s_k[f](x) - f(x)$ will behave near 0 in about the same way as the difference $\mathrm{Si}(x) - \frac{1}{2}\pi\,\mathrm{sgn}\,x$, except that the horizontal scale is shortened by a factor of about $1/k$.

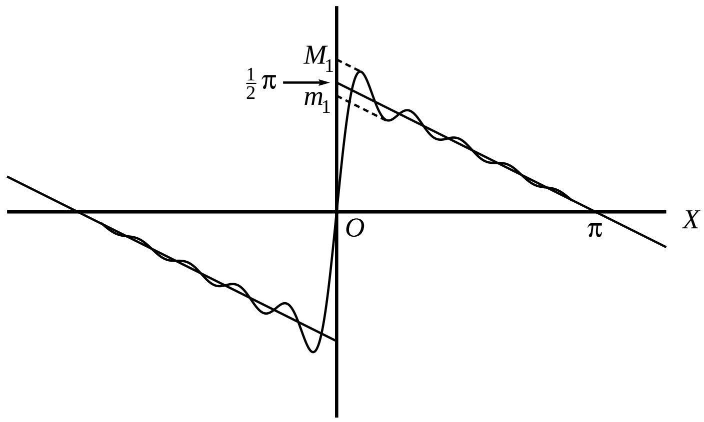

FIGURE 2.5

We will analyze the phenomenon for the function

$$f(x) = \frac{\pi}{2} \operatorname{sgn} x - \frac{1}{2} x = \sum_{n=1}^{\infty} \frac{\sin nx}{n}, \quad |x| \le \pi; \tag{2.6.3}$$

cf. Figure 2.5. By the preceding analysis

$$\begin{aligned}\frac{1}{2} x + s_k(x) &= \int_0^x \left(\frac{1}{2} + \sum_{n=1}^{k} \cos nt \right) dt = \int_0^x \frac{\sin(k+\frac{1}{2})t}{2 \sin \frac{1}{2} t} \, dt \\ &= \int_0^x \frac{\sin(k+\frac{1}{2})t}{t} \, dt + r_k(x) = \operatorname{Si} \left\{ \left(k + \frac{1}{2} \right) x \right\} + r_k(x),\end{aligned}$$

where $r_k(x) \to 0$ uniformly for $|x| \le \pi$ as $k \to \infty$. It follows that

$$s_k(x) - f(x) = \operatorname{Si} \left\{ \left(k + \frac{1}{2} \right) x \right\} - \frac{1}{2} \pi \operatorname{sgn} x + r_k(x).$$

In particular for $k \to \infty$,

$$s_k \left(\frac{\pi}{k+\frac{1}{2}} \right) - f \left(\frac{\pi}{k+\frac{1}{2}} \right) \to \operatorname{Si}(\pi) - \frac{1}{2}\pi = M_1 - \frac{1}{2}\pi \approx 0.28, \tag{2.6.4}$$

$$s_k \left(\frac{2\pi}{k+\frac{1}{2}} \right) - f \left(\frac{2\pi}{k+\frac{1}{2}} \right) \to \operatorname{Si}(2\pi) - \frac{1}{2}\pi = m_1 - \frac{1}{2}\pi \approx -0.16. \tag{2.6.5}$$

Cf. Figure 2.6. At the first maximum point the "overshoot" is about 18%. This happens at the jump discontinuities of any reasonable function.

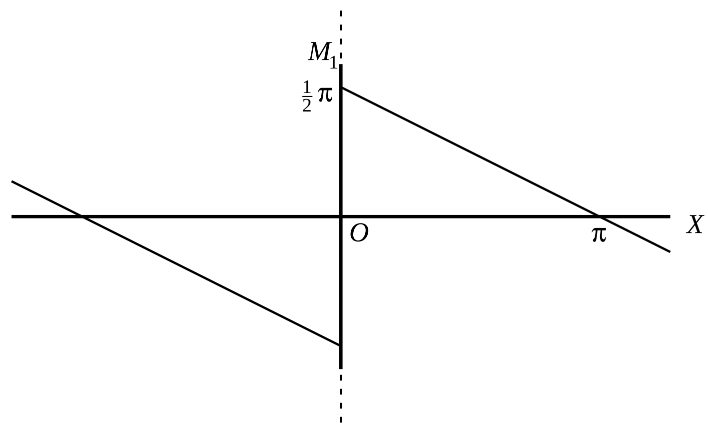

FIGURE 2.6

Exercises. 2.6.1. Let f be a piecewise smooth function on $(-\pi, \pi]$ with a jump at the point 0. Show that for large k, the apparent jump of the partial sum $s_k[f]$ around 0 is about 18% larger than that of f.

CHAPTER 3

Summability of Fourier series

Important summability methods for (possibly divergent) infinite series are Cesàro's method of arithmetic means (named after Ernesto Cesàro, Italy, 1859–1906; [**16**]), and the power series method that is related to Abel's Continuity Theorem 1.1.2. The power series method is usually called the "Abel method", although Abel himself rejected the use of divergent series. We will see that the Fourier series for a continuous 2π-periodic function is uniformly summable to the function by Cesàro's method. As an application we derive Weierstrass's theorems (after Karl Weierstrass, Germany, 1815–1897; [**123**]) on the uniform approximation of continuous functions by trigonometric and ordinary polynomials. The Fourier series for a continuous function is also summable by the Abel method. As an application we give a careful treatment of the Dirichlet problem for Laplace's equation on a disc.

3.1. Cesàro and Abel summability

Good summability methods assign a reasonable generalized sum to many divergent (= nonconvergent) series, while summing every convergent series to its usual sum. The methods in this section will both assign the generalized sum $\frac{1}{2}$ to the divergent series

$$1 - 1 + 1 - 1 + 1 - 1 + \cdots . \tag{3.1.1}$$

*This famous (or should we say: infamous?) series has long fascinated nonmathematicians. At one time it was facetiously used to illustrate the story of Creation. One would write

$$1 - 1 + 1 - 1 + 1 - 1 + \cdots = 1 - 1 + 1 - 1 + 1 - 1 + \cdots ,$$

and then insert parentheses as follows:

$$\begin{aligned}&(1-1) + (1-1) + (1-1) + \cdots \\ &= 1 + (-1+1) + (-1+1) + (-1+\cdot) + \cdots .\end{aligned}$$

The conclusion would be:

$$0 = 1,$$

hence it would be possible to "create" something out of nothing!

Notation. In this chapter we write infinite series of real or complex numbers in the form

$$u_0 + u_1 + u_2 + \cdots, \quad \text{or equivalently,} \quad \sum_{n=0}^{\infty} u_n. \tag{3.1.2}$$

The partial sums will be denoted by s_k:

$$s_k = u_0 + u_1 + \cdots + u_k = \sum_{n=0}^{k} u_n, \quad k = 0, 1, 2, \cdots. \tag{3.1.3}$$

We also introduce the arithmetic means or $(C, 1)$ means [Cesàro means of order one] σ_k of the first k partial sums:

$$\begin{aligned} \sigma_k &= \frac{s_0 + s_1 + \cdots + s_{k-1}}{k} \\ &= \frac{u_0 + (u_0 + u_1) + \cdots + (u_0 + \cdots + u_{k-1})}{k} \\ &= u_0 + \left(1 - \frac{1}{k}\right)u_1 + \cdots + \left(1 - \frac{n}{k}\right)u_n + \cdots + \frac{1}{k}u_{k-1}, \end{aligned} \tag{3.1.4}$$

$k = 1, 2, \cdots$.

By definition a series (3.1.2) is convergent, and has sum s, if $\lim s_k$ exists and is equal to s. Observe that if $\lim s_k = s$, then also $\lim \sigma_k = s$. Indeed, if $|s - s_n| < \varepsilon$ for all $n \geq p \geq 1$ and $\sup_n |s - s_n| = M$, then for $k > p$,

$$\begin{aligned} &|\sigma_k - s| \\ &= \left| \frac{(s_0 - s) + \cdots + (s_{p-1} - s)}{k} + \frac{(s_p - s) + \cdots + (s_{k-1} - s)}{k} \right| \\ &\leq \frac{p}{k} M + \frac{k-p}{k} \varepsilon < 2\varepsilon \quad \text{for all} \ \ k \geq \frac{pM}{\varepsilon}. \end{aligned}$$

On the other hand, it may happen that σ_k has a limit while s_k does not.

EXAMPLE 3.1.1. For the series (3.1.1) we have

$$\begin{aligned} &s_0 = 1,\ s_1 = 0,\ s_2 = 1,\ s_3 = 0,\ s_4 = 1,\ \cdots; \quad \lim s_k \ \text{ does not exist,} \\ &\sigma_1 = 1,\ \sigma_2 = \frac{1}{2},\ \sigma_3 = \frac{2}{3},\ \sigma_4 = \frac{1}{2},\ \sigma_5 = \frac{3}{5},\ \cdots; \quad \lim \sigma_k = \frac{1}{2}. \end{aligned}$$

Indeed, $\sigma_k = \frac{1}{2}$ if k is even, $\sigma_k = \frac{1}{2} + \frac{1}{2k}$ if k is odd.

DEFINITION 3.1.2. The series $\sum_{n=0}^{\infty} u_n$ is *C-summable* or $(C,1)$-summable [Cesàro summable of order one, or summable by the method of arithmetic means] if $\lim \sigma_k$ exists. If $\lim \sigma_k = \sigma$ one calls σ the C-sum [Cesàro sum] of the series.

Cesàro actually introduced a family of summability methods (C,k); cf. [**17**].

By the preceding every convergent series, with sum s, is C-summable, with C-sum σ equal to s. The divergent series (3.1.1) has C-sum $\frac{1}{2}$.

Suppose now that for the series (3.1.2), the corresponding power series $\sum_{n=0}^{\infty} u_n r^n$ converges (at least) for $|r| < 1$. Then by absolute convergence,

$$(1 + r + r^2 + \cdots)(u_0 + u_1 r + u_2 r^2 + \cdots) = s_0 + s_1 r + s_2 r^2 + \cdots$$

as long as $|r| < 1$. The *Abel means* of the partial sums s_k are given by

$$\begin{aligned} A_r \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} u_n r^n &= \frac{s_0 + s_1 r + s_2 r^2 + \cdots}{1 + r + r^2 + \cdots} \\ &= (1-r)\sum_{n=0}^{\infty} s_n r^n, \quad 0 \le r < 1. \qquad (3.1.5) \end{aligned}$$

If $\lim s_k = s$ as $k \to \infty$, then also $\lim A_r = s$ as $r \nearrow 1$. Indeed, assuming $s_k \to s$, the numbers $u_n = s_n - s_{n-1}$ form a bounded sequence, hence the series $\sum_{n=0}^{\infty} u_n r^n$ will converge at least for $|r| < 1$. Furthermore

$$A_r - s = (1-r)\sum_{n=0}^{\infty} (s_n - s) r^n = (1-r)\left(\sum_{n<p} + \sum_{n \ge p}\right)(s_n - s) r^n$$

will be small when p is large and r is close to 1.

EXAMPLES 3.1.3. For the divergent series (3.1.1) we have

$$A_r = 1 - r + r^2 - r^3 + \cdots = \frac{1}{1+r} \to \frac{1}{2} \quad \text{as } r \nearrow 1.$$

That $\lim A_r = \lim \sigma_k$ for the series (3.1.1) is no coincidence. Indeed, if for any series $\sum u_n$ one has $\sigma_k \to \sigma$, then also $A_r \to \sigma$; see below. The converse of this is not true, as may be derived from the example of the series

$$1 - 2 + 3 - 4 + 5 - \cdots . \qquad (3.1.6)$$

This series is not C-summable, but it *is* "Abel summable":

$$A_r = 1 - 2r + 3r^2 - 4r^3 + \cdots = \frac{d}{dr}\left(-1 + r - r^2 + r^3 - r^4 + \cdots\right)$$
$$= \frac{d}{dr}\frac{-1}{1+r} = \frac{1}{(1+r)^2} \to \frac{1}{4} \quad \text{as } r \nearrow 1.$$

DEFINITION 3.1.4. The series $\sum_{n=0}^{\infty} u_n$ is *A-summable* [Abel summable, or summable by the method of power series] if the Abel means $A_r = \sum_{n=0}^{\infty} u_n r^n$ exist for $0 \le r < 1$ and approach a limit as $r \nearrow 1$. If $\lim A_r = \alpha$ as $r \nearrow 1$ one calls α the A-sum [*Abel sum*] of the series.

By the preceding every convergent series, with sum s, is A-summable, with A-sum α equal to s. Every C-summable series, with C-sum σ, will be A-summable, with A-sum α equal to σ (cf. Exercise 3.1.4). However, not all A-summable series are C-summable. The A-method of summability is "stronger" than the C-method.

For some applications it is useful to consider inverse theorems for summability methods, so-called *Tauberian theorems*; cf. [**2**]. The prototype was the following theorem of Alfred Tauber (Austria, 1866–1942; [**118**]):

THEOREM 3.1.5. *Let the series $\sum_{n=0}^{\infty} u_n$ be Abel summable and suppose that $nu_n \to 0$ as $n \to \infty$. Then $\sum_{n=0}^{\infty} u_n$ is convergent.*

Cf. Exercise 3.1.7.

Exercises. 3.1.1. Supposing that $s_k \to \infty$, prove that also $\sigma_k \to \infty$. Deduce that a C-summable series of nonnegative terms must be convergent.

3.1.2. Show that the series (3.1.6) is not C-summable.

3.1.3. Express s_n in terms of numbers σ_k and deduce that the terms u_n of a C-summable series must satisfy the condition $u_n/n \to 0$.

3.1.4. Setting

$$s_0 + s_1 + \cdots + s_n = s_n^{(-1)} \quad [= (n+1)\sigma_{n+1}],$$

show that

$$A_r = (1-r)\sum_{n=0}^{\infty} s_n r^n = (1-r)^2 \sum_{n=0}^{\infty} s_n^{(-1)} r^n,$$
$$A_r - \sigma = (1-r)^2 \sum_{n=0}^{\infty} (n+1)(\sigma_{n+1} - \sigma) r^n.$$

Deduce that if $\sigma_k \to \sigma$ as $k \to \infty$, then also $A_r \to \sigma$ as $r \nearrow 1$.

3.1.5. Determine the A-sum of the series $1^2 - 2^2 + 3^2 - 4^2 + \cdots$.

3.1.6. Determine the A-means and the A-sums for the series

$$\frac{1}{2} + \cos x + \cos 2x + \cdots, \quad \sin x + \sin 2x + \cdots \quad (0 < x < 2\pi).$$

3.1.7. Prove Tauber's Theorem 3.1.5.

Hint. Setting $\sum_{n=0}^{\infty} u_n r^n = f(r)$ one has

$$s_k - f(r) = \sum_{n=1}^{k} u_n(1 - r^n) - \sum_{n=k+1}^{\infty} u_n r^n.$$

3.1.8. (*Discrete Taylor formula*) Taking u_n real, show that for $h \in \mathbb{N}$,

$$s_{k+h}^{(-1)} = s_k^{(-1)} + hs_k + \frac{1}{2}h(h+1)u_\xi^*,$$

where u_ξ^* is a number such that

$$\min_{k+1 \le n \le k+h} u_n \le u_\xi^* \le \max_{k+1 \le n \le k+h} u_n.$$

3.1.9. (*Tauberian theorem* of Godfrey H. Hardy (England, 1877–1947; [**44**]). Suppose that $\sum_{n=0}^{\infty} u_n$ is C-summable and that $|u_n| \le B/n$ for all $n \ge 1$. Prove that $\sum_{n=0}^{\infty} u_n$ is convergent.

Hint. Decreasing the original u_0 by σ, one may assume that $\sigma = 0$, so that $|\sigma_k| < \varepsilon$ for all $k \ge k_0$. Then $\left|s_n^{(-1)}\right| < \varepsilon(n+1)$ for all $n \ge n_0$. Now estimate $|s_k|$ from Exercise 3.1.8 by choosing h appropriately.

[This simple approach to Hardy's theorem is due to Hendrik D. Kloosterman (Netherlands, 1900–1968; [**64**]). Soon after Hardy obtained his result, John E. Littlewood (England, 1885–1977; [**84**]) proved a corresponding (more difficult) Tauberian theorem for Abel summability. Subsequently, Hardy and Littlewood jointly obtained a very large number of Tauberian theorems; cf. Korevaar's book [**69**].]

3.2. Cesàro means. Fejér kernel

Let f be an integrable function on $(-\pi, \pi)$; as usual we extend f to a 2π-periodic function. The partial sum $s_k[f]$ of the Fourier series for f is given by Dirichlet's integrals (Theorem 2.2.3). For the arithmetic mean $\sigma_k[f]$ of the first k partial sums we thus obtain the formula

$$\begin{aligned} \sigma_k(x) = \sigma_k[f](x) &= \frac{s_0(x) + \cdots + s_{k-1}(x)}{k} \\ &= \int_{-\pi}^{\pi} f(x \pm t)\, \frac{D_0(t) + \cdots + D_{k-1}(t)}{k}\, dt. \end{aligned} \tag{3.2.1}$$

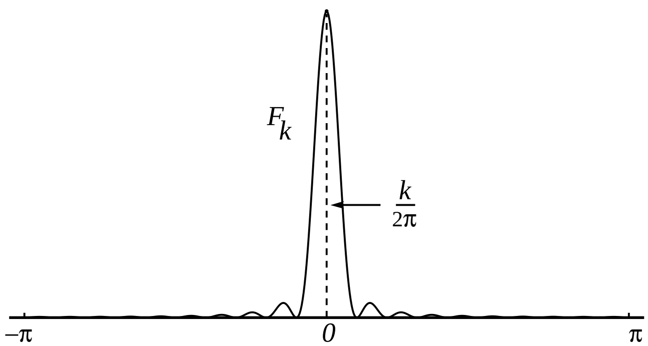

FIGURE 3.1

The "kernel" by which $f(x \pm t)$ is multiplied is called the *Fejér kernel*. It may be expressed in closed form and is illustrated in Figure 3.1.

LEMMA 3.2.1. *For $k = 1, 2, \cdots$ and all $t \in \mathbb{R}$,*

$$F_k(t) \stackrel{\text{def}}{=} \frac{D_0(t) + \cdots + D_{k-1}(t)}{k} = \frac{1}{\pi}\left\{\frac{1}{2} + \sum_{n=1}^{k}\left(1 - \frac{n}{k}\right)\cos nt\right\}$$

$$= \frac{1}{k}\sum_{n=0}^{k-1}\frac{\sin(n+\frac{1}{2})t}{2\pi\sin\frac{1}{2}t} = \frac{\sin^2\frac{1}{2}kt}{2\pi k\sin^2\frac{1}{2}t}. \tag{3.2.2}$$

[At the points $t = 2\nu\pi$ the last two fractions are defined by their limit values.]

PROOF. By its definition (cf. Lemma 2.2.1), the Dirichlet kernel $D_n(t)$ is equal to the nth order partial sum of the trigonometric series

$$\frac{1}{\pi}\left(\frac{1}{2} + \cos t + \cos 2t + \cdots\right).$$

Hence $F_k(t)$ is the kth Cesàro mean for this series; cf. formula (3.1.4). In view of Lemma 2.2.1, it only remains to derive the final identity in (3.2.2). For this we have to evaluate the sum $\sum_{n=0}^{k-1}\sin(n+\frac{1}{2})t$. Writing $\sin(n+\frac{1}{2})t$ as the imaginary part of $e^{(n+\frac{1}{2})it}$, we first work out the corresponding sum of exponentials:

$$\begin{aligned} e^{\frac{1}{2}it} + e^{\frac{3}{2}it} + \cdots + e^{(k-\frac{1}{2})it} &= e^{\frac{1}{2}it}\,\frac{e^{kit}-1}{e^{it}-1} \\ &= \frac{e^{kit}-1}{e^{\frac{1}{2}it}-e^{-\frac{1}{2}it}} = i\,\frac{1-e^{kit}}{2\sin\frac{1}{2}t}. \end{aligned}$$

Thus, taking imaginary parts,

$$\sin\frac{1}{2}t + \sin\frac{3}{2}t + \cdots + \sin\left(k-\frac{1}{2}\right)t = \frac{1-\cos kt}{2\sin\frac{1}{2}t} = \frac{\sin^2\frac{1}{2}kt}{\sin\frac{1}{2}t}.$$

The final identity in (3.2.2) follows upon division by $2\pi k\sin\frac{1}{2}t$. □

The Fejér kernel behaves much better than the Dirichlet kernel. By the preceding it has the following nice properties:

LEMMA 3.2.2. *F_k is* nonnegative, *even and periodic with period 2π. As $k\to\infty$, $F_k(t)$ tends to* 0 uniformly *on the intervals $[\delta,\pi]$ and $[-\pi,-\delta]$ for any $\delta\in(0,\pi)$, while*

$$\int_{-\pi}^{\pi} F_k(t)dt = 1, \quad \forall\, k\in\mathbb{N}. \tag{3.2.3}$$

Formulas (3.2.1) and (3.2.2) readily give

THEOREM 3.2.3. *Let f be 2π-periodic and integrable over a period. Then*

$$\begin{aligned} \sigma_k(x) = \sigma_k[f](x) &= \frac{s_0(x)+\cdots+s_{k-1}(x)}{k} \\ &= \frac{1}{2}a_0 + \sum_{n=1}^{k-1}\left(1-\frac{n}{k}\right)(a_n\cos nx + b_n\sin nx) \qquad (3.2.4) \\ &= \int_{-\pi}^{\pi} f(x\pm t)F_k(t)dt = \int_{-\pi}^{\pi}\frac{f(x+t)+f(x-t)}{2}\,F_k(t)dt, \quad \forall\, k. \end{aligned}$$

Exercises. 3.2.1. Let f be real, 2π-periodic, integrable over a period and bounded: $m\le f(x)\le M$, $\forall\, x$. Prove that $m\le\sigma_k[f](x)\le M$, $\forall\, x$, k. Deduce that the averages $\sigma_k[f](x)$ cannot exhibit a "divergence phenomenon" as in Exercise 2.2.4, nor a "Gibbs phenomenon" as in Section 2.6.

3.2.2. Determine the C-means and the C-sum for the series

$$\frac{1}{2} + \cos x + \cos 2x + \cdots : \ (i) \text{ for } 0<x<2\pi,\ (ii) \text{ for } x=0.$$

3.2.3. Same questions for the series $\sin x + \sin 2x + \cdots$.

Hint. Show that the nth order partial sum is equal to

$$\frac{\cos\frac{1}{2}x - \cos(n+\frac{1}{2})x}{2\sin\frac{1}{2}x}.$$

3.3. Cesàro summability: Fejér's theorems

We assume throughout that f is an integrable function on $(-\pi,\pi]$ which has been made periodic with period 2π.

THEOREM 3.3.1. (Pointwise summability) *Suppose that f satisfies one of the following conditions:*

(i) *f is continuous at the point x;*

(ii) *f has a finite right-hand limit $f(x+)$ and a finite left-hand limit $f(x-)$ at x, but these two are different.*

Then the Fourier series for f is C-summable at the point x to the value $f(x)$, and to the value $\frac{1}{2}\{f(x+)+f(x-)\}$, respectively.

PROOF. Case (i). By Theorem 3.2.3 and Lemma 3.2.2,

$$\sigma_k[f](x) - f(x) = \int_{-\pi}^{\pi}\{f(x+t)-f(x)\}F_k(t)dt. \tag{3.3.1}$$

For given x and $\varepsilon > 0$ we choose $\delta \in (0,\pi)$ such that

$$|f(x+t)-f(x)| < \varepsilon \quad \text{for} \quad -\delta < t < \delta. \tag{3.3.2}$$

Then for all k,

$$\left|\int_{-\delta}^{\delta}\{f(x+t)-f(x)\}F_k(t)dt\right| < \varepsilon\int_{-\delta}^{\delta}F_k(t)dt < \varepsilon \tag{3.3.3}$$

by (3.2.3). On the other hand

$$\begin{aligned} &\left|\left(\int_{-\pi}^{-\delta}+\int_{\delta}^{\pi}\right)\{f(x+t)-f(x)\}F_k(t)dt\right| \\ &\le \max_{\delta\le|t|\le\pi} F_k(t)\left(\int_{-\pi}^{-\delta}+\int_{\delta}^{\pi}\right)|f(x+t)-f(x)|dt \\ &\le \frac{1}{2\pi k\sin^2\frac{1}{2}\delta}\left(\int_{-\pi}^{\pi}|f(u)|du + 2\pi|f(x)|\right). \end{aligned} \tag{3.3.4}$$

Combining the above relations one finds that $|\sigma_k(x) - f(x)| < 2\varepsilon$ for all $k \ge k_0(x,\delta)$.

The proof in case (ii) is similar, but this time one would use the final integral in (3.2.4). □

We now come to the most important theorem of Fejér:

THEOREM 3.3.2. (Uniform summability) (i) *For* $f \in \mathcal{C}_{2\pi}$, *the Cesàro means* $\sigma_k[f]$ *converge to* f *uniformly on* $[-\pi, \pi]$ (hence on $\mathbb{R}$).

(ii) *If* f *is continous on* (a, b), *then* $\sigma_k[f] \to f$ *uniformly on every closed subinterval* $[\alpha, \beta]$ *of* (a, b).

PROOF. We only consider case (i). Let $\varepsilon > 0$. Since our continuous periodic function f will be uniformly continuous, we can choose $\delta \in (0, \pi)$ such that (3.3.2) and hence (3.3.3) hold for all x and k. Setting $\sup_{\mathbb{R}} |f| = M$, the final member of (3.3.4) will be bounded by $2M/(k \sin^2 \frac{1}{2}\delta)$ for all x. Conclusion:

$$|\sigma_k(x) - f(x)| < 2\varepsilon \quad \text{for all } \; x \in \mathbb{R} \; \text{ when } \; k \geq 2M \Big/ \left(\varepsilon \sin^2 \frac{1}{2}\delta\right).$$

□

THEOREM 3.3.3. (Summability in the mean of order one) *For any integrable function* f *on* $(-\pi, \pi)$,

$$\int_{-\pi}^{\pi} \big|\sigma_k[f](x) - f(x)\big| dx \to 0 \quad \textit{as } \; k \to \infty.$$

*PROOF. Here we need the theorem of Guido Fubini (Italy, 1879–1943; [**34**]) which allows inversion of the order of integration in an absolutely convergent repeated integral [see Integration Theory]. By (3.3.1), making f periodic,

$$\begin{aligned}
\Delta_k &\stackrel{\text{def}}{=} \int_{-\pi}^{\pi} \big|\sigma_k[f](x) - f(x)\big| dx \\
&= \int_{-\pi}^{\pi} \left| \int_{-\pi}^{\pi} \{f(x+t)) - f(x)\} F_k(t) dt \right| dx \\
&\leq \int_{-\pi}^{\pi} \left\{ \int_{-\pi}^{\pi} |f(x+t)) - f(x)| F_k(t) dt \right\} dx \\
&= \int_{-\pi}^{\pi} \left\{ \int_{-\pi}^{\pi} |f(x+t)) - f(x)| dx \right\} F_k(t) dt.
\end{aligned} \tag{3.3.5}$$

Setting

$$g(t) = \int_{-\pi}^{\pi} \big|f(x+t) - f(x)\big| dx, \tag{3.3.6}$$

the final member of (3.3.5) is equal to $\sigma_k[g](0)$; cf. (3.2.4). Now g is a continuous function (of period 2π), as one readily verifies by approximating f with piecewise constant functions s [cf. Section 2.1, 2.4]. Thus by Theorem 3.3.1,

$$\Delta_k \leq \sigma_k[g](0) \to g(0) = 0 \quad \text{as} \quad k \to \infty.$$

□

Exercises. 3.3.1. Let f be (periodic and) continuous at the point x and suppose that the Fourier series for f converges at x. Prove that the sum is equal to $f(x)$.

3.3.2. Prove that two integrable functions on $(-\pi, \pi)$ with the same Fourier series are equal at all points where they are continuous.

3.3.3. Let f be continuous on $[-\pi, \pi]$ with $f(\pi) \neq f(-\pi)$. Compute $\lim \sigma_k[f](\pi)$.

3.3.4. Let f be continuous on $[-\pi, \pi]$ and such that all *trigonometric moments* of f are equal to zero:

$$\int_{-\pi}^{\pi} f(x) \cos nx \, dx = \int_{-\pi}^{\pi} f(x) \sin nx \, dx = 0, \quad \forall\, n \in \mathbb{N}_0.$$

Prove that $f \equiv 0$.

3.3.5. Let $f \in \mathcal{C}_{2\pi}$ have $b_n[f] = 0$ for all $n \in \mathbb{N}$. Prove that f is even.

3.3.6. What can you say about $f \in \mathcal{C}[0, \pi]$ if $\int_0^\pi f(x) \cos nx \, dx = 0$ for all $n \geq 0$? What if $\int_0^\pi f(x) \cos nx \, dx = 0$ for all $n \geq 1$? What if $\int_0^\pi f(x) \cos nx \, dx = 0$ for all $n \geq 5$?

3.3.7. Let f be a bounded piecewise monotonic function on $[-\pi, \pi]$. Prove that the Fourier series for f is C-summable everywhere. Now use Exercises 2.1.8 and 3.1.9 to deduce that the Fourier series for f is everywhere convergent. Describe its sum function.

3.3.8. Prove part (ii) of Theorem 3.3.2.

3.3.9. Prove that the function g in (3.3.6) is continuous at the point $t = 0$.

3.3.10. Let f be integrable over $(-\pi, \pi)$ and such that $\int_{-\pi}^{\pi} f(x) e^{-inx} dx = 0$ for all $n \in \mathbb{Z}$. Prove that $\int_{-\pi}^{\pi} |f(x)| dx = 0$, so that $f(x) = 0$ almost everywhere on $(-\pi, \pi)$. [Cf. Integration Theory for the final step.]

3.4. Weierstrass theorem on polynomial approximation

Theorem 3.3.2 immediately implies

THEOREM 3.4.1. *Let f be continuous on $[-\pi, \pi]$ and such that $f(\pi) = f(-\pi)$. Then to every $\varepsilon > 0$ there is a trigonometric polynomial* (finite trigonometric sum) $S(x) = \alpha_0 + \sum_{n=1}^{k} (\alpha_n \cos nx + \beta_n \sin nx)$ *such that*

$$|f(x) - S(x)| < \varepsilon \quad \textit{for} \quad -\pi \leq x \leq \pi.$$

Indeed, f can be extended to a continuous function of period 2π (which we also call f), and the arithmetic means $\sigma_k[f] = \frac{1}{k}(s_0[f] + \cdots + s_{k-1}[f])$ of the partial sums of the Fourier series for f converge to f uniformly on $\mathbb{R}$.

As an application we will prove Weierstrass's theorem on uniform approximation by ordinary polynomials:

THEOREM 3.4.2. *Let f be continuous on the bounded closed interval $[a, b]$. Then to every $\varepsilon > 0$ there is a polynomial $p(x)$ such that*

$$|f(x) - p(x)| < \varepsilon \quad \textit{for} \;\; a \leq x \leq b.$$

PROOF. We may assume without loss of generality that $[a, b]$ is the interval $[-1, 1]$. Indeed, one can always carry out an initial transformation $x = \frac{1}{2}(a + b) + \frac{1}{2}(b - a)X$, so that $f(x)$ becomes a continous function $F(X)$ on $[-1, 1]$. If one has approximated the latter by a polynomial $P(X)$ on $[-1, 1]$, an approximating polynomial $p(x)$ for $f(x)$ on $[a, b]$ is obtained by setting

$$p(x) = P\left(\frac{2x - a - b}{b - a}\right).$$

Now starting with a continuous function f on $[-1, 1]$, we set

$$x = \cos t, \quad f(x) = f(\cos t) = g(t), \quad t \in \mathbb{R}.$$

Then g will be of class $\mathcal{C}_{2\pi}$ and even, so that the Fourier series for g contains only cosine terms. By Theorem 3.4.1,

$$\begin{aligned} f(\cos t) = g(t) &= \lim_{k\to\infty} \sigma_k[g](t) \\ &= \lim_{k\to\infty} \left\{ \frac{1}{2}\, a_0[g] + \sum_{n=1}^{k-1} \left(1 - \frac{n}{k}\right) a_n[g] \cos nt \right\}, \end{aligned} \tag{3.4.1}$$

uniformly on $\mathbb{R}$. [Of course, if g is sufficiently nice, also $s_k[g](t) \to g(t)$ uniform;y on $\mathbb{R}$.]

We finally express $\cos nt$ as a polynomial in $\cos t$, the *Chebyshev polynomial* $T_n(\cos t)$ (after Pafnuty Chebyshev, Russia, 1821–1894; [**18**]):

$$\begin{aligned}T_n(\cos t) = \cos nt &= \operatorname{Re} e^{int} = \operatorname{Re}(\cos t + i \sin t)^n \\ &= \operatorname{Re} \sum_{j=0}^{n} \binom{n}{j} (\cos^{n-j} t)(i \sin t)^j \\ &= \cos^n t - \binom{n}{2} (\cos^{n-2} t)(1-\cos^2 t) \\ &\quad + \binom{n}{4} (\cos^{n-4} t)(1-\cos^2 t)^2 - \cdots .\end{aligned}$$

Conclusion:

$$f(x) = \lim_{k\to\infty} \left\{ \frac{1}{2} a_0[g] + \sum_{n=1}^{k-1} \left(1 - \frac{n}{k}\right) a_n[g]\, T_n(x) \right\}, \tag{3.4.2}$$

uniformly on $[-1, 1]$. □

REMARK 3.4.3. Observe that the coefficient $A_{n,k}$ of the polynomial T_n in this approximation tends to the limit $A_n = a_n[g]$ as $k \to \infty$. In contrast the coefficient $b_{n,k}$ of x^n in the approximating polynomial $p(x) = p_k(x)$ behind the limit sign in (3.4.2) may vary a great deal as $k \to \infty$.

For later use we restate the definition of the Chebyshev polynomial T_n:

DEFINITION 3.4.4.

$$T_n(x) \stackrel{\text{def}}{=} \cos nt\Big|_{\cos t = x} = \sum_{0\le j\le n/2} (-1)^j \binom{n}{2j} x^{n-2j}(1-x^2)^j.$$

Exercises. 3.4.1. Let $f(x) = |x|$. Determine a sequence of polynomials which converges to f uniformly on $[-1, 1]$.

3.4.2. Let f be integrable over the finite interval (a, b). Prove that for every $\varepsilon > 0$ there is a polynomial p such that

$$\int_a^b |f(x) - p(x)| dx < \varepsilon.$$

3.4.3. Let f be continuous on the finite closed interval $[a, b]$ and such that all *power moments* of f are equal to zero:

$$\int_a^b f(x) x^n dx = 0, \quad \forall\, n \in \mathbb{N}_0.$$

Prove that $f \equiv 0$.

3.4.4. Prove that the Chebyshev polynomials T_0, T_1, T_2, $\cdots$ form an orthogonal system on $(-1,1)$ relative to the weight function $1/\sqrt{1-x^2}$:

$$\int_{-1}^{1} T_n(x)T_k(x)\frac{dx}{\sqrt{1-x^2}} = 0 \quad \text{whenever} \quad k \neq n.$$

3.4.5. Show that the coefficient of x^n in $T_n(x)$ is equal to

$$1+\binom{n}{2}+\binom{n}{4}+\cdots = \frac{1}{2}\left\{1+\binom{n}{1}+\binom{n}{2}+\cdots\right\} = 2^{n-1}.$$

Hint. Expand $(1\pm 1)^n$.

3.5. Abel summability. Poisson kernel

Let f be 2π-periodic and integrable over a period. Anticipating work with polar coordinates we write the Fourier series for f with variable θ:

$$f(\theta) \sim \frac{1}{2}a_0[f] + \sum_{n=1}^{\infty}(a_n[f]\cos n\theta + b_n[f]\sin n\theta).$$

The Abel means of the partial sums are given by

$$A_r[f](\theta) = \frac{1}{2}a_0[f] + \sum_{n=1}^{\infty}(a_n[f]\cos n\theta + b_n[f]\sin n\theta)r^n, \tag{3.5.1}$$

where $0 \leq r < 1$; cf. (3.1.5). We will express these means as integrals. Replacing the Fourier coefficients by their defining integrals, in which we use variable of integration s, one obtains

$$A_r[f](\theta) = \frac{1}{2}\,\frac{1}{\pi}\int_{-\pi}^{\pi} f(s)ds + \sum_{n=1}^{\infty}\frac{1}{\pi}\int_{-\pi}^{\pi} f(s)\cos n(s-\theta)\,ds\cdot r^n. \tag{3.5.2}$$

Here we may invert the order of summation and integration since $f(s)$ is integrable and the series $\sum_{n=1}^{\infty}\cos n(s-\theta)\cdot r^n$ converges uniformly in s. Thus

$$\begin{aligned} A_r[f](\theta) &= \int_{-\pi}^{\pi} f(s)\,\frac{1}{\pi}\left\{\frac{1}{2}+\sum_{n=1}^{\infty} r^n\cos n(s-\theta)\right\}ds \\ &= \int_{-\pi}^{\pi} f(\theta+t)\,\frac{1}{\pi}\left\{\frac{1}{2}+\sum_{n=1}^{\infty} r^n\cos nt\right\}dt, \quad 0\leq r<1. \end{aligned} \tag{3.5.3}$$

The kernel by which $f(\theta+t)$ is multiplied is called the *Poisson kernel*,

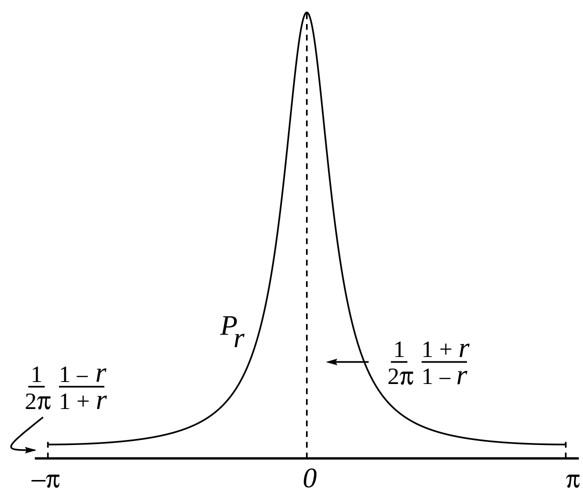

FIGURE 3.2

after the French mathematical physicist Siméon Denis Poisson (1781–1840; [**93**]); cf. [**94**] and see Figure 3.2. It may be expressed in closed form:

LEMMA 3.5.1. *For* $0 \le r < 1$ *and all* $t \in \mathbb{R}$,

$$P_r(t) \stackrel{\text{def}}{=} \frac{1}{\pi}\left(\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos nt\right) = \frac{1}{2\pi}\,\frac{1-r^2}{1-2r\cos t + r^2}. \tag{3.5.4}$$

PROOF. Writing $\cos nt = \operatorname{Re} e^{nit}$ one has

$$1 + 2\sum_{n=1}^{\infty} r^n \cos nt = \operatorname{Re}\left(1 + 2\sum_{n=1}^{\infty} r^n e^{nit}\right) = \operatorname{Re}\left(1 + 2\,\frac{re^{it}}{1-re^{it}}\right)$$
$$= \operatorname{Re}\frac{1+re^{it}}{1-re^{it}}\,\frac{1-re^{-it}}{1-re^{-it}} = \frac{1-r^2}{1-2r\cos t + r^2}.$$

□

The Poisson kernel has nice properties similar to those of the Fejér kernel:

LEMMA 3.5.2. P_r *is* nonnegative*, even and periodic with period* 2π*. As* $r \nearrow 1$*,* $P_r(t)$ *tends to* 0 uniformly *on the intervals* $[\delta, \pi]$ *and* $[-\pi, -\delta]$ *for any given* $\delta \in (0, \pi)$*, while*

$$\int_{-\pi}^{\pi} P_r(t)dt = 1, \quad \forall\, r \in [0, 1). \tag{3.5.5}$$

Indeed, $1-2r\cos t+r^2=(1-r)^2+4r\sin^2\frac{1}{2}t$, so that

$$0<P_r(t)\le\frac{1-r^2}{4r\sin^2\frac{1}{2}\delta}\quad\text{for}\quad\delta\le t\le\pi.$$

Also, for $f\equiv 1$, the Fourier series reduces to the constant 1, so that $\int_{-\pi}^{\pi}P_r=A_r[1]\equiv 1$; see (3.5.3).

Definition 3.5.3. For integrable f on $(-\pi,\pi)$, the integral

$$P[f](r,\theta)\stackrel{\text{def}}{=}\int_{-\pi}^{\pi}P_r(\theta-s)f(s)ds \tag{3.5.6}$$

is called the *Poisson integral* for f.

By the preceding one has

Proposition 3.5.4. *For periodic integrable f, the Abel mean $A_r[f]$ for the Fourier series is equal to the Poisson integral:*

$$A_r[f](\theta)=\int_{-\pi}^{\pi}f(\theta\pm t)P_r(t)dt=P[f](r,\theta).$$

Furthermore

$$A_r[f](\theta)-f(\theta)=\int_{-\pi}^{\pi}\{f(\theta\pm t)-f(\theta)\}P_r(t)dt,\quad 0\le r<1.$$

The methods of Section 3.3 may now be used to obtain the analogs of Theorems 3.3.1–3.3.3 for Abel summability. We only state some important aspects:

Theorem 3.5.5. (i) *For $f\in\mathcal{C}_{2\pi}$, the Abel means $A_r[f]$ of the partial sums of the Fourier series converge to f uniformly on $[-\pi,\pi]$ as $r\nearrow 1$.*

(ii) *For 2π-periodic f, piecewise continuous on $[-\pi,\pi]$, the Abel means $A_r[f]$ remain bounded as $r\nearrow 1$, and they converge to f uniformly on every closed subinterval $[\alpha,\beta]$ of an interval (a,b) where f is continuous.*

Exercises. 3.5.1. Justify the step from (3.5.2) to (3.5.3) by showing that the conditions "f integrable" and "$g_k\to g$ uniformly on (a,b)" imply that $\int_a^b fg_k\to\int_a^b fg$.

3.5.2. Use the Poisson integral for $A_r[f]$ to prove part (i) of Theorem 3.5.5.

3.5.3. Let $f(\theta)=\operatorname{sgn}\theta$, $|\theta|<\pi$ [cf. Exercise 1.2.5]. Determine $A_r[f]$ first as a series, then as an integral. Verify that $-1\le A_r[f]\le 1$ and show that $A_r[f]\to 1$ uniformly on $[\delta,\pi-\delta]$ as $r\nearrow 1$ (provided $0<\delta<\pi/2$).

3.6. Laplace equation: circular domains, Dirichlet problem

We will deal with (real) *harmonic functions*: solutions of Laplace's equation. A harmonic function u on a domain D in $\mathbb{R}^m$ is smooth (in fact, of class $\mathcal{C}^\infty$) and it cannot attain a maximum or minimum in D unless it is constant. As a consequence, if $\limsup u(x) \le M$ on all sequences of points in D that approach the boundary ∂D, then $u(x) \le M$ throughout D. It follows that harmonic functions on a bounded domain are uniquely determined by their boundary values whenever the latter form a continuous function on ∂D. [If the boundary function is only piecewise continuous one may impose a boundedness condition to ensure uniqueness.]

Here we consider circular domains D in the plane: annuli [ring-shaped domains] and discs, or the exterior of a disc. On such domains one will use polar coordinates r, θ with the origin at the center of the domain. Laplace's equation then takes the form

$$\Delta u \stackrel{\text{def}}{=} u_{xx} + u_{yy} \equiv u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta} = 0. \tag{3.6.1}$$

Solutions on the *annulus*

$$A(0,\rho,R) = \{(r,\theta) : \rho < r < R,\ \theta \in \mathbb{R}\}$$

have period 2π as functions of θ, hence, being smooth, they can be represented by Fourier series with coefficients depending on r:

$$u(r,\theta) = \frac{1}{2}\,A_0(r) + \sum_{n=1}^{\infty} \{A_n(r)\cos n\theta + B_n(r)\sin n\theta\}. \tag{3.6.2}$$

Here the coefficients $A_n(r) = \frac{1}{\pi}\int_{-\pi}^{\pi} u(r,\theta)\cos n\theta\, d\theta$ and $B_n(r)$ will be smooth functions of r. Also, for fixed r, the coefficients and their derivatives will form sequences that are $\mathcal{O}(n^{-p})$ for every p; cf. Lemma 2.1.2. We may then form Δu by termwise differentiation of the series in (3.6.2). [In problems of mathematical physics one should always *try* to carry out operations term by term, justification can wait till later!] Thus Laplace's equation becomes

$$\begin{aligned}\Delta u(r,\theta) = \frac{1}{2}&\left\{A_0''(r) + \frac{1}{r}\,A_0'(r)\right\}\\ &+ \sum_{n=1}^{\infty}\left\{\left[A_n''(r) + \frac{1}{r}\,A_n'(r) - \frac{n^2}{r^2}\,A_n(r)\right]\cos n\theta\right.\\ &+ \left.\left[B_n''(r) + \frac{1}{r}\,B_n'(r) - \frac{n^2}{r^2}\,B_n(r)\right]\sin n\theta\right\} = 0.\end{aligned}$$

Since Δu will be continuous, all coefficients in this Fourier series must be equal to zero. It follows that the functions $A_n(r)$ and $B_n(r)$ must satisfy the ordinary differential equation

$$\frac{d^2v(r)}{dr^2} + \frac{1}{r}\frac{dv(r)}{dr} - \frac{n^2}{r^2}v(r) = 0, \quad \rho < r < R. \tag{3.6.3}$$

Recall that "equidimensional equations" such as (3.6.3) have solutions of the form $v(r) = r^\alpha$. Substitution gives

$$\{\alpha(\alpha-1) + \alpha - n^2\}r^{\alpha-2} = 0, \quad \text{hence } \alpha = \pm n.$$

For $n = 0$ equation (3.6.3) has the additional solution $\log r$. Thus we find

$$\begin{aligned} &A_n(r) = a_n r^n + \tilde{a}_n r^{-n}, \quad B_n(r) = b_n r^n + \tilde{b}_n r^{-n} \quad \text{for } n \in \mathbb{N}, \\ &A_0(r) = a_0 + \tilde{a}_0 \log r, \end{aligned} \tag{3.6.4}$$

where a_n, $\tilde{a}_n$, b_n, $\tilde{b}_n$ are constants.

Dirichlet problem for the disc $B(0,1)$. The unit disc corresponds to an annulus with inner radius $\rho = 0$, but the origin has to be included in the domain. Therefore we have to reject the solutions r^{-n} and $\log r$ of equation (3.6.3): they would lead to solutions of Laplace's equation that become unbounded at the origin. Thus our candidate solution (3.6.2) for Laplace's equation in the disc takes the form

$$u(r,\theta) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\theta + b_n \sin n\theta)r^n, \quad 0 \le r < 1, \tag{3.6.5}$$

with constants a_n, b_n that make the series converge.

In the *Dirichlet problem* one prescribes the boundary function:

$$u(1,\theta) = f(\theta), \quad |\theta| \le \pi,$$

with a given function f. Ignoring questions of convergence on the boundary, we are thus led to expand $f(\theta) = u(1,\theta)$ in a Fourier series:

$$f(\theta) \sim \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\theta + b_n \sin n\theta), \quad |\theta| \le \pi;$$

cf. Figure 3.3. It is therefore natural to use the Fourier coefficients of f: $a_n = a_n[f]$ and $b_n = b_n[f]$, in our trial solution (3.6.5).

QUESTION 3.6.1. Does the function $u(r,\theta) = u_f(r,\theta)$ formed with these coefficients indeed have the correct boundary values?

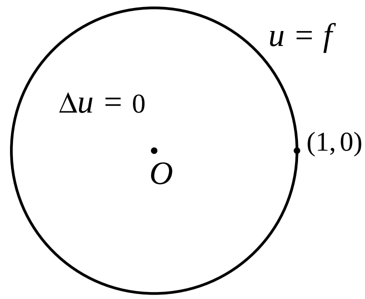

FIGURE 3.3

THEOREM 3.6.2. *For $f \in \mathcal{C}_{2\pi}$ the series* (or Abel means)

$$\begin{aligned} u_f(r,\theta) &= \frac{1}{2}a_0[f] + \sum_{n=1}^{\infty}(a_n[f]\cos n\theta + b_n[f]\sin n\theta)r^n \\ &= A_r[f](\theta) \end{aligned} \tag{3.6.6}$$

and the corresponding Poisson integral

$$\begin{aligned} P[f](r,\theta) &\stackrel{\text{def}}{=} \int_{-\pi}^{\pi} P_r(\theta - t)f(t)dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-r^2}{1-2r\cos(\theta-t)+r^2}\, f(t)dt \end{aligned} \tag{3.6.7}$$

both represent the (unique) *solution of Laplace's equation in the disc $B(0,1)$ with boundary values $f(\theta)$. That is, for every $\theta_0 \in \mathbb{R}$,*

$$\lim_{(r,\theta)\to(1,\theta_0)} u_f(r,\theta) = f(\theta_0). \tag{3.6.8}$$

Observe that we require more than just radial approach from inside B to the boundary ∂B.

PROOF. Every term in the series (3.6.6) satisfies Laplace's equation. Now the coefficients $a_n[f]$ and $b_n[f]$ form bounded sequences. It follows that the operator Δ may be applied to $u_f(r,\theta)$ term by term. Indeed, the differentiated series will be uniformly convergent for $0 \le r \le r_0 < 1$ and $\theta \in \mathbb{R}$. Hence also $\Delta u_f(r,\theta) = 0$.

For (3.6.8) we use the fact that $u_f(r,\theta)$ is equal to the Abel mean $A_r[f](\theta)$. By Theorem 3.5.5, $A_r[f](\theta) = P[f](r,\theta)$ converges to $f(\theta)$ uniformly in θ as $r \nearrow 1$. Thus for given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|u_f(r,\theta) - f(\theta)| < \varepsilon \quad \text{for } 1-\delta < r < 1 \text{ and all } \theta,$$
$$|f(\theta) - f(\theta_0)| < \varepsilon \quad \text{for } |\theta - \theta_0| < \delta.$$

As a result

$$|u_f(r,\theta) - f(\theta_0)| < 2\varepsilon \quad \text{for } 1-\delta < r < 1,\ |\theta - \theta_0| < \delta.$$

□

REMARK 3.6.3. Theorem 3.6.2 can be extended to the case of (bounded) piecewise continuous boundary functions f. At a point θ_0 where f is discontinuous, (3.6.8) may then be replaced by the condition that $u_f(r,\theta)$ must remain bounded as $(r,\theta) \to (1,\theta_0)$. Without such a condition there would be no uniqueness of the solution; cf. Exercise 3.6.6.

Exercises. 3.6.1. Prove directly that the Poisson integral $P[f](r,\theta)$ of a real integrable function satisfies Laplace's equation in the unit disc, either by differentiating under the integral sign, or by showing that $P[f]$ is the real part of an analytic function.

3.6.2. Use an infinite series to solve the Dirichlet problem for Laplace's equation in the disc $B(0,R)$ with boundary function $f(R,\theta)$. Then transform the series into the *Poisson integral for the disc* $B(0,R)$:

$$P_R[f](r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{R^2 - r^2}{R^2 - 2Rr\cos(\theta - t) + r^2}\, f(R,t)dt.$$

3.6.3. Solve the *Neumann problem* for Laplace's equation in $B = B(0,1)$:

$$\Delta u = 0 \text{ on } B, \quad \frac{\partial u}{\partial r}(1,\theta) = g(\theta), \quad |\theta| \le \pi.$$

[Here one has to require that $\int_{-\pi}^{\pi} g(\theta)d\theta = 0$.] Consider in particular the case where $g(\theta) = \operatorname{sgn}\theta$, $|\theta| < \pi$.

3.6.4. Use an infinite series to solve the Dirichlet problem for Laplace's equation on the exterior of the disc $\overline{B}(0,\rho)$, with boundary function $f(\rho,\theta)$. Here one requires that u remain bounded at infinity. Finally transform the series into a Poisson-type integral.

3.6.5. Determine the solution of Laplace's equation in the general annulus $A(0,\rho,R)$ with boundary function 1 on $C(0,R)$ and 0 on $C(0,\rho)$.

3.6.6. The Poisson kernel $P_r(\theta)$ represents a solution of Laplace's equation in the unit disc with boundary values 0 for $r = 1$, $0 < |\theta| \leq \pi$. Verify this, and investigate what happens when (r, θ) tends to the point $(1, 0)$.

3.6.7. Use an infinite series to solve the following boundary value problem for the semidisc $D = \{(r, \theta) : 0 < r < 1,\, 0 < \theta < \pi\}$:

$$\begin{aligned} &\Delta u = 0 \text{ in } D, \quad u(1, \theta) = 1 \text{ for } 0 < \theta < \pi, \\ &u(r, 0) = u(r, \pi) = 0 \text{ for } 0 < r < 1. \end{aligned}$$

Can you prove that your candidate solution has the correct boundary values?

3.6.8. Similarly for the sector $D = \{(r, \theta) : 0 < r < R,\, 0 < \theta < \alpha\}$:

$$\begin{aligned} &\Delta u = 0 \text{ on } D, \quad u(R, \theta) = 1 \text{ for } 0 < \theta < \alpha, \\ &u(r, 0) = u(r, \alpha) = 0 \text{ for } 0 < r < R. \end{aligned}$$

Show that for fixed (r, θ) and large R,

$$u(r, \theta) \approx \frac{4}{\pi} \left(\frac{r}{R}\right)^{\pi/\alpha} \sin \frac{\pi}{\alpha} \theta.$$

CHAPTER 4

Periodic distributions and Fourier series

Although an integrable function on $(-\pi, \pi)$ is essentially determined by its Fourier series, there is no simple convergence theory that applies to the Fourier series of every (Lebesgue) integrable function. Is it possible to do something about that? More important, for applications to boundary value problems one would like to have a theory in which there is no limitation on termwise differentiation of Fourier series; cf. Sections 1.3, 1.4 and 3.6. The difficulties can be overcome by the introduction of "convergence relative to test functions", and the extension of the class of integrable functions to the class of generalized functions or *distributions* in the sense of Laurent Schwartz; cf. [**110**], [**111**]). It will turn out that every periodic distribution can be considered as a generalized derivative, of some finite order, of a periodic integrable function.

4.1. The space L^1. Test functions

Integrable functions f_1 and f_2 on $(-\pi, \pi)$ that differ only on a set of measure zero have the same Fourier series:

$$c_n[f_1] = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(x)e^{-inx}dx = c_n[f_2] = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_2(x)e^{-inx}dx, \quad \forall\, n.$$

Indeed, since $f_1(x)e^{-inx}$ will be equal to $f_2(x)e^{-inx}$ almost everywhere [often abbreviated a.e.], that is, outside a set of measure zero, the two products will have the same integral. Conversely, integrable functions f_1 and f_2 with the same Fourier series will differ only on a set of measure zero; cf. Exercise 3.3.10. We will give another proof here.

THEOREM 4.1.1. *Let f be integrable on $(-\pi, \pi)$ and $c_n[f] = 0$ for all $n \in \mathbb{Z}$. Then $f(x) = 0$ almost everywhere, and in particular $f(x) = 0$ at every point x where the indefinite integral*

$$F(x) \stackrel{\text{def}}{=} c + \int_0^x f(t)dt \tag{4.1.1}$$

is differentiable and has derivative $F'(x)$ equal to $f(x)$.

PROOF. By Lebesgue's theory of integration (cf. [**77**], [**76**]), the indefinite integral $F(x)$ is differentiable outside a set of measure zero and $F'(x) = f(x)$ outside a (possibly larger) set of measure zero. With indefinite integrals one can carry out integration by parts:

$$\begin{aligned} 2\pi c_n[f] &= \int_{-\pi}^{\pi} f(x)e^{-inx}dx = \int_{-\pi}^{\pi} e^{-inx}dF(x) \\ &= \Big[F(x)e^{-inx}\Big]_{-\pi}^{\pi} + in\int_{-\pi}^{\pi} F(x)e^{-inx}dx. \end{aligned} \tag{4.1.2}$$

Here the integrated term will vanish since $e^{-in\pi} = e^{in\pi}$ and

$$F(\pi) - F(-\pi) = \int_{-\pi}^{\pi} f(t)dt = 2\pi c_0[f] = 0.$$

Thus $2\pi c_n[f] = in \cdot 2\pi c_n[F]$, and hence in our case $c_n[F] = 0$ for all $n \neq 0$. Subtracting from F its average $C = c_0[F]$ on $[-\pi, \pi]$, one will obtain a continuous (and a.e. differentiable) function with Fourier series 0:

$$c_n[F - C] = c_n[F] = 0, \ \ \forall\, n \neq 0, \ \ c_0[F - C] = c_0[F] - C = 0.$$

Conclusion: $F - C = \lim \sigma_k[F - C] = 0$, so that $F \equiv C$; cf. Theorem 3.3.1. Hence by Lebesgue's theory, $f(x) = F'(x) = 0$ almost everywhere. $\square$

REMARK 4.1.2. The linear space of integrable functions f on (a, b) is made into a *normed vector space*, called $L(a, b)$ or $L^1(a, b)$, by setting

$$\|f\| = \|f\|_1 = \int_a^b |f(x)|dx. \tag{4.1.3}$$

This is the L^1 *norm* or "L^1 length" of f. One now identifies functions which differ only on a set of measure 0. Note that $\|f_1 - f_2\| = 0$ if and only if $f_1(x) = f_2(x)$ a.e. Also, as is usual for a length,

$$\|\lambda f\| = |\lambda|\,\|f\| \quad \text{and} \quad \|f + g\| \leq \|f\| + \|g\|.$$

Convergence $f_k \to f$ in $L^1(a, b)$ means $\int_a^b |f - f_k| \to 0$.

Speaking precisely, the elements of $L^1(a, b)$ are not *functions*: they are *equivalence classes* of integrable functions on (a, b). For any given integrable function f, the class of all functions that are a.e. equal to it is sometimes denoted by $[f]$. The norm of the element or class $[f]$ is defined as $\|f\|$ and given by (4.1.3) for any element of the class. To every element of $L^1(-\pi, \pi)$ there is exactly one Fourier series. Different elements of $L^1(-\pi, \pi)$ have different Fourier series. We usually speak carelessly of the *functions* of $L^1(-\pi,\pi)$ instead of the elements.

DEFINITION 4.1.3. We say that integrable functions f_k converge to the integrable function f on (a,b) relative to the *test class* (class of test functions) A if

$$\int_a^b f_k\phi \to \int_a^b f\phi \ \text{ as } \ k\to\infty, \quad \forall\, \phi\in A.$$

In order to make it easy for f_k to converge to f we severely limit the test class A, but it must be large enough to make limits relative to A unique. For Fourier theory, we will use as test functions the infinitely differentiable functions ϕ of period 2π: the test class will be $\mathcal{C}^\infty_{2\pi}$.

EXAMPLE 4.1.4. By the Riemann–Lebesgue Lemma 2.1.1, the sequence $\{e^{ikx}\}$, $k=0,1,2,\cdots$ tends to 0 on $(-\pi,\pi)$ relative to the test class $\mathcal{C}^\infty_{2\pi}$. More surprisingly, a sequence such as $\{k^{100}e^{ikx}\}$, $k=0,1,2,\cdots$ also tends to 0 relative to this test class. Indeed, one has

$$\begin{aligned}&\int_{-\pi}^{\pi} k^{100}e^{ikx}\phi(x)dx = k^{100}\int_{-\pi}^{\pi}\phi(x)d\,\frac{e^{ikx}}{ik}\\ &= -\frac{k^{100}}{ik}\int_{-\pi}^{\pi} e^{ikx}\phi'(x)dx = \cdots = \frac{k^{100}}{(ik)^{100}}\int_{-\pi}^{\pi} e^{ikx}\phi^{(100)}(x)dx.\end{aligned}$$

The result tends to 0 as $k\to\infty$ since the function $\phi^{(100)}$ is continuous.

PROPOSITION 4.1.5. *In $L^1(-\pi,\pi)$, limits relative to the test class $\mathcal{C}^\infty_{2\pi}$ are unique.*

PROOF. Suppose that for integrable functions f_k, f, g on $(-\pi,\pi)$ one has

$$\int_{-\pi}^{\pi} f_k\phi \to \int_{-\pi}^{\pi} f\phi \ \text{ and also } \ \int_{-\pi}^{\pi} f_k\phi \to \int_{-\pi}^{\pi} g\phi$$

for all $\phi\in\mathcal{C}^\infty_{2\pi}$. Then $\int_{-\pi}^{\pi} f\phi = \int_{-\pi}^{\pi} g\phi$ for all ϕ. Thus, using the test functions e^{-inx}, one finds in particular that $c_n[f]=c_n[g]$ for all n. Hence by Theorem 4.1.1, $f=g$ almost everywhere, so that $[f]=[g]$ in $L^1(-\pi,\pi)$. □

Using the fact that for functions $\phi\in\mathcal{C}^\infty_{2\pi}$, the Fourier series is uniformly convergent to ϕ, we will show that for integrable functions f on $(-\pi,\pi)$, the Fourier series *converges to f relative to the test class $\mathcal{C}^\infty_{2\pi}$*.

LEMMA 4.1.6. *For integrable functions f and g on $(-\pi, \pi)$,*

$$\int_{-\pi}^{\pi} s_k[f]g = 2\pi \sum_{n=-k}^{k} c_n[f]c_{-n}[g] = \int_{-\pi}^{\pi} f s_k[g].$$

Indeed,

$$\int_{-\pi}^{\pi} s_k[f]g = \int_{-\pi}^{\pi} \sum_{n=-k}^{k} c_n[f]e^{inx}g(x)dx = \sum_{n=-k}^{k} c_n[f] \cdot 2\pi c_{-n}[g], \quad \text{etc.}$$

THEOREM 4.1.7. *For $f \in L^1(-\pi, \pi)$ and any $\phi \in \mathcal{C}_{2\pi}^{\infty}$,*

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} s_k[f]\phi = \int_{-\pi}^{\pi} f\phi = 2\pi \sum_{n=-\infty}^{\infty} c_n[f]c_{-n}[\phi]. \tag{4.1.4}$$

Indeed, one has

$$\int_{-\pi}^{\pi} s_k[f]\phi = \int_{-\pi}^{\pi} f s_k[\phi] \to \int_{-\pi}^{\pi} f\phi \quad \text{as } k \to \infty,$$

since

$$\left| \int_{-\pi}^{\pi} f(\phi - s_k[\phi]) \right| \leq \int_{-\pi}^{\pi} |f| \cdot \max_{[-\pi,\pi]} |\phi - s_k[\phi]| \to 0$$

by Theorem 2.4.3. The infinite series in (4.1.4) will be absolutely convergent; cf. Lemma 2.1.2 applied to ϕ instead of f.

DEFINITION 4.1.8. The *support* of a continuous function [or element of L^1, or generalized function] f on J is the smallest closed subset $E \subset J$ outside of which f is equal to zero [or equal to the zero element]. Notation: $\operatorname{supp} f$.

It is important to know that for any given finite closed interval $[\alpha, \beta]$, there are $\mathcal{C}^\infty$ functions on $\mathbb{R}$ with support $[\alpha, \beta]$.

EXAMPLES 4.1.9. The function

$$\psi(x) = \begin{cases} e^{-1/x} & \text{for } x > 0, \\ 0 & \text{for } x \leq 0 \end{cases}$$

is in $\mathcal{C}^\infty(\mathbb{R})$ and has support $[0, \infty)$.

For $\alpha < \beta$ the product $\psi(x - \alpha)\psi(\beta - x)$ is in $\mathcal{C}^\infty(\mathbb{R})$ and has support $[\alpha, \beta]$.

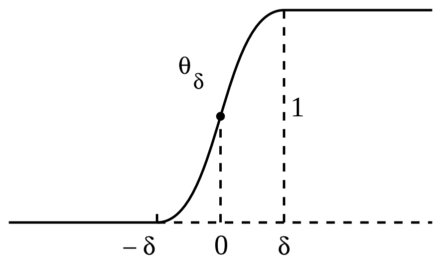

FIGURE 4.1

For any number $\delta > 0$, the function

$$\theta_\delta(x) = \begin{cases} \int_{-\delta}^{x} e^{-1/(\delta^2-t^2)}dt \big/ \int_{-\delta}^{\delta} e^{-1/(\delta^2-t^2)}dt & \text{for } -\delta \le x \le \delta, \\ 0 & \text{for } x \le -\delta, \\ 1 & \text{for } x \ge \delta \end{cases}$$

belongs to $\mathcal{C}^\infty(\mathbb{R})$; cf. Figure 4.1. Observe that $\theta_\delta(x) + \theta_\delta(-x) \equiv 1$.

For $0 < \delta < \frac{1}{4}(\beta - \alpha)$ the function

$$\omega(x) = \theta_\delta(x - \alpha - \delta)\, \theta_\delta(\beta - x - \delta)$$

is in $\mathcal{C}^\infty(\mathbb{R})$. It has support $[\alpha, \beta]$ and is equal to 1 on $[\alpha + 2\delta, \beta - 2\delta]$; cf. Figure 4.2.

Exercises. 4.1.1. Let f_k and f be integrable functions of period 2π. Prove the following implications:

$$\begin{aligned} &f_k \to f \quad \text{uniformly on } [-\pi, \pi] \\ &\Rightarrow f_k \to f \quad \text{in } L^1(-\pi, \pi) \\ &\Rightarrow f_k \to f \quad \text{relative to the test class } \mathcal{C}_{2\pi} \\ &\Rightarrow f_k \to f \quad \text{relative to the test class } \mathcal{C}_{2\pi}^\infty. \end{aligned}$$

4.1.2. (Continuation). Show that in each case, $c_n[f_k] \to c_n[f]$ as $k \to \infty$ for every n.

4.1.3. Let f_k, f, g be integrable on $(-\pi, \pi)$ and suppose that $f_k \to f$ relative to the test class $\mathcal{C}_{2\pi}^\infty$, while $f_k \to g$ uniformly on $(a, b) \subset (-\pi, \pi)$, or in the sense of $L^1(a, b)$. Prove that $f = g$ a.e. on (a, b).

Hint. Consider sequences $\{f_k \omega\}$ where ω has support $[a, b]$.

4.1.4. For indefinite integrals F on $[-\pi, \pi]$, the Fourier series converges to F everywhere on $(-\pi, \pi)$ (why?). Use this fact to show that for integrable functions f on $(-\pi, \pi)$, the series $\sum_{n=1}^\infty \frac{1}{n} b_n[f]$ is convergent.

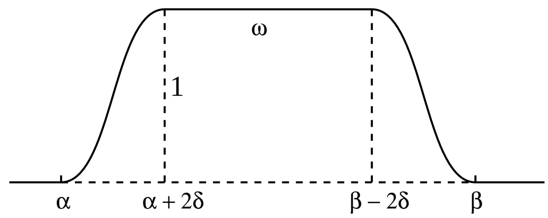

FIGURE 4.2

4.1.5. Prove that the series

$$\sum_{n=2}^{\infty} \frac{1}{\log n} \sin nx$$

converges for all x and that the sum function $g(x)$ is continuous on $(0, 2\pi)$. Show that the series cannot be the Fourier series of an integrable function on $(-\pi, \pi)$ or $(0, 2\pi)$.

4.1.6. Prove that the function $\psi(x)$ under Examples 4.1.9 is of class $\mathcal{C}^\infty(\mathbb{R})$. Derive that the other functions are also of class $\mathcal{C}^\infty(\mathbb{R})$.

4.1.7. Prove that an integrable function f on Γ is equal to a test function if and only if for every $p \in \mathbb{N}$, one has $c_n[f] = \mathcal{O}(|n|^{-p})$ as $|n| \to \infty$.

4.1.8. Construct an example to show that the Fourier series of a periodic L^1 function need not converge to f in the sense of L^1.

Hint. For every $k \in \mathbb{N}$ there is a function f_k with $|f_k(x)| \equiv 1$ such that $\|s_k[f_k]\|$ is close to $\|D_k\|$. Now form a suitable series $f = \sum a_j f_{p_j}$.

4.2. Periodic distributions: distributions on the unit circle

Functions on $\mathbb{R}$ of period 2π may be considered as functions on the unit circle (unit circumference) $\Gamma = \{z \in \mathbb{C} : |z| = 1\}$. Here we will not take $z = e^{ix}$ as our independent varaibale, but rather the (*signed*) *arc length* x from the point $z = 1$ to $z = e^{ix}$. Where one-to-one correspondence between the points of Γ and the values of x is important we may take $-\pi < x \leq \pi$ or $0 \leq x < 2\pi$. However, we sometimes speak of arcs on Γ of length $> 2\pi$, for example, the arc given by $-2\pi < x < 2\pi$. Integration over Γ relative to arc length shall be the same as integration over the interval $(-\pi, \pi]$ or $(-\pi, \pi)$. Differentiation with respect to arc length corresponds to differentiation on $\mathbb{R}$. The p times continuously differentiable functions on Γ correspond to the class $\mathcal{C}^p_{2\pi}$. By definition the test class $\mathcal{C}^\infty(\Gamma)$ on the unit circle corresponds to the test class $\mathcal{C}^\infty_{2\pi}$ on $\mathbb{R}$. Convergence of integrable functions $f_k \to f$ relative to the test class $\mathcal{C}^\infty(\Gamma)$ is the same as convergence on $(-\pi, \pi)$ relative to the test class $\mathcal{C}^\infty_{2\pi}$:

$$\int_\Gamma f_k \phi = \int_{-\pi}^{\pi} f_k(x)\phi(x)dx \to \int_\Gamma f\phi, \quad \forall\, \phi \in \mathcal{C}^\infty(\Gamma).$$

In the definition of convergence, integrable functions enter only through their action on test functions. The integrable function f appears in the form of the *linear functional* $T_f : \mathcal{C}^\infty(\Gamma) \Rightarrow \mathbb{C}$ given by

$$T_f(\phi) = < T_f, \phi > = \int_\Gamma f\phi, \quad \forall\, \phi \in \mathcal{C}^\infty(\Gamma). \tag{4.2.1}$$

Observe that the correspondence between f and the functional T_f is one to one if we identify functions that are equal almost everywhere, that is, (4.2.1) establishes a one-to-one correspondence between the elements $f \in L^1(\Gamma)$ and the associated functionals T_f.

We will now introduce more general linear functionals on our test class, and call these *generalized functions*. However, we will not allow *all* linear functionals – in order to get a good structure theory of generalized functions, we will impose a *continuity condition*. To this end we introduce a suitable concept of convergence for test functions. In order to make it easy for functionals to be continuous, we have to make it difficult for test functions to converge.

Definition 4.2.1. The *test space* $\mathcal{D}(\Gamma)$ consists of the $\mathcal{C}^\infty$ functions ϕ on Γ (henceforth called *test functions*), furnished with the following concept of convergence: $\phi_j \to \phi$ in $\mathcal{D}(\Gamma)$ if and only if

$$\begin{aligned}&\phi_j \to \phi \;\text{ uniformly on }\; \Gamma, \;\; \phi_j' \to \phi' \;\text{ uniformly on }\; \Gamma, \cdots, \\ &\phi_j^{(p)} \to \phi^{(p)} \;\text{ uniformly on }\; \Gamma, \cdots .\end{aligned}$$

Observe that convergence $\phi_j \to \phi$ in $\mathcal{D}(\Gamma)$ implies convergence $\phi_j' \to \phi'$ in $\mathcal{D}(\Gamma)$, etc.

Proposition 4.2.2. *For a test function ϕ on Γ, the Fourier series converges to ϕ in the strong sense of $\mathcal{D}(\Gamma)$:*

$$\begin{aligned}&s_j[\phi] \to \phi \;\textit{ uniformly on }\; \Gamma, \;\; s_j'[\phi] \to \phi' \;\textit{ uniformly on }\; \Gamma, \cdots, \\ &s_j^{(p)}[\phi] \to \phi^{(p)} \;\textit{ uniformly on }\; \Gamma, \cdots .\end{aligned}$$

PROOF. One has

$$s_j^{(p)}[\phi](x) = \left(\frac{d}{dx}\right)^p \sum_{n=-j}^{j} c_n[\phi]e^{inx} = \sum_{n=-j}^{j} (in)^p c_n[\phi]e^{inx}$$
$$= \sum_{n=-j}^{j} c_n[\phi^{(p)}]e^{inx} = s_j[\phi^{(p)}](x);$$

cf. Lemma 2.1.2. Furthermore, since $\phi^{(p)}$ is of class $\mathcal{C}^\infty(\Gamma)$ or $\mathcal{C}^\infty_{2\pi}$, the partial sum $s_j[\phi^{(p)}]$ converges to $\phi^{(p)}$ uniformly on Γ. $\square$

DEFINITION 4.2.3. A *distribution* (or generalized function) T on the unit circle Γ is a *continuous linear functional* on the test space $\mathcal{D}(\Gamma)$. Such a distribution can also be considered as a distribution on $\mathbb{R}$ of period 2π.

In order to make the definition more clear, we introduce different symbols for our continuous linear functionals. Besides T, we will use $T(\cdot)$ or $T(\phi)$, and also $< T, \phi >$. If there is no danger of confusion, one sometimes writes $T(x)$ in order to indicate the underlying independent variable x on Γ. However, a distribution need not have a value at the point x.

By the definition, a distribution T on Γ is a *map* $\mathcal{D}(\Gamma) \Rightarrow \mathbb{C}$ which is *linear*:

$$< T, \lambda_1\phi_1 + \lambda_2\phi_2 >= \lambda_1 < T, \phi_1 > +\lambda_2 < T, \phi_2 >$$

for all $\lambda_j \in \mathbb{C}$ and all $\phi_j \in \mathcal{D}(\Gamma)$, and *continuous*:

$$< T, \phi_j > \to < T, \phi > \quad \text{whenever } \ \phi_j \to \phi \ \text{ in } \ \mathcal{D}(\Gamma).$$

As linear functionals, distributions can be added and multiplied by scalars:

$$< \lambda_1 T_1 + \lambda_2 T_2, \phi >= \lambda_1 < T_1, \phi > +\lambda_2 < T_2, \phi >, \quad \forall\, \phi.$$

EXAMPLE 4.2.4. Every integrable function f on Γ defines a distribution T_f on Γ by formula (4.2.1). Indeed, $\int_\Gamma f\phi_j \to \int_\Gamma f\phi$ already if $\phi_j \to \phi$ uniformly on Γ. In the terminology of Chapter 5 we could say that f defines a continuous linear functional relative to the convergence in the space $C(\Gamma)$.

Since the correspondence $f \Leftrightarrow T_f$ is one to one for $f \in L^1(\Gamma)$ and preserves linear combinations, we may identify T_f with f and also write $< f, \phi >$ instead of $< T_f, \phi >$. Thus (periodic) integrable functions (more precisely, equivalence classes of integrable functions) become special cases of (periodic) distributions. Distributions on Γ form a generalization of integrable functions. Some important distributions correspond to special non-integrable functions; cf. Example 4.2.6 below.

EXAMPLES 4.2.5. The *delta distribution* on the circle, notation δ_Γ, is defined by the formula

$$< \delta_\Gamma, \phi > = \phi(0), \quad \forall\, \phi \in \mathcal{D}(\Gamma). \tag{4.2.2}$$

In physics, a distribution that assigns the value $\phi(0)$ to test functions ϕ is usually called a *Dirac delta function*, after the British physicist Paul Dirac (1902–1984; [**22**]). It is given symbolically by the formula

$$\int_{-\pi}^{\pi} \delta_\Gamma(x)\phi(x)dx = \phi(0);$$

cf. [**24**]. However, δ_Γ cannot be identified with an integrable function as we will see below. [Incidentally, one may consider δ_Γ also as a 2π-periodic distribution on $\mathbb{R}$; in that case we use the notation $\delta_{2\pi}^{\mathrm{per}}$.]

*The distribution δ_Γ actually defines a continuous linear functional relative to the less demanding convergence in $C(\Gamma)$: if $\phi_j \to \phi$ uniformly on Γ, then $< \delta_\Gamma, \phi_j > \to < \delta_\Gamma, \phi >$. By a representation theorem of Frigyes Riesz (Hungary, 1880–1956; [**100**]), cf. [**101**]), a continuous linear functional on $C(\Gamma)$ can be identified with a (real or complex) Borel measure. (Such measures are named after Émile Borel, France, 1871–1956; [**9**]; cf. [**10**].) Thus the delta distribution is an example of a measure.

For any nonnegative integer p, the formula

$$< T, \phi > = \phi^{(p)}(0), \quad \forall\, \phi \in \mathcal{D}(\Gamma)$$

defines a distribution on Γ. Indeed, if $\phi_j \to \phi$ in $\mathcal{D}(\Gamma)$, then $\phi_j^{(p)}(0) \to \phi^{(p)}(0)$.

EXAMPLE 4.2.6. We recall the definition of a *principal value integral*. Let $a < c < b$ and suppose that a function f is integrable over $(a, c - \varepsilon)$ and over $(c + \varepsilon, b)$ for all small $\varepsilon > 0$, but not necessarily over (a, b) itself. Then f has a principal value integral over (a, b) [relative to the point c] if the following limit exists:

$$\lim_{\varepsilon \searrow 0} \int_{(a,b)\setminus(c-\varepsilon,c+\varepsilon)} f(x)dx. \tag{4.2.3}$$

It is important that the omitted interval be *symmetric* with respect to c. The principal value p.v. $\int_a^b f(x)dx$ is defined by the value of the limit in (4.2.3).

Suppose now that $f(x)$ has the form $(1/x)\phi(x)$, where ϕ is of class $\mathcal{C}^1$ on a finite interval $[a, b]$ with $a < 0 < b$. Then the principal value p.v. $\int_a^b f(x)dx$

[relative to 0] exists and may be obtained through integration by parts. Indeed, writing $(1/x)dx = d\log|x|$ one finds

$$\left(\int_a^{-\varepsilon} + \int_\varepsilon^b\right)\frac{1}{x}\,\phi(x)dx = \Big[\phi(x)\log|x|\Big]_a^{-\varepsilon} + \Big[\phi(x)\log|x|\Big]_\varepsilon^b$$
$$- \left(\int_a^{-\varepsilon} + \int_\varepsilon^b\right)(\log|x|)\,\phi'(x)dx.$$

The second member has a finite limit as $\varepsilon \searrow 0$ since $(\log|x|)\,\phi'(x)$ is integrable over (a,b) and $\{\phi(\varepsilon) - \phi(-\varepsilon)\}\log\varepsilon \to 0$. Thus

$$\text{p.v.}\int_a^b \frac{1}{x}\,\phi(x)dx = \phi(b)\log b - \phi(a)\log|a| - \int_a^b (\log|x|)\phi'(x)dx.$$

The *principal value distribution* on Γ corresponding to the nonintegrable function $1/x$ is defined by

$$\left\langle \text{pv}_\Gamma\,\frac{1}{x}, \phi(x)\right\rangle \overset{\text{def}}{=} \text{p.v.}\int_{-\pi}^{\pi}\frac{1}{x}\,\phi(x)dx.$$

Here integration by parts will show that

$$\left\langle \text{pv}_\Gamma\,\frac{1}{x}, \phi(x)\right\rangle = -\int_{-\pi}^{\pi}\log\frac{|x|}{\pi}\,\phi'(x)dx, \quad \forall\,\phi\in\mathcal{D}(\Gamma).$$

The continuity of the functional follows from the uniform convergence of ϕ_j' to ϕ' when $\phi_j \to \phi$ in $\mathcal{D}(\Gamma)$.

Definition 4.2.7. (Simple operations) The *translate* $T_c(x) = T(x-c)$, the *reflection* $T_R(x) = T(-x)$ and the *product* Tg of T with a test function g are defined as if $<T,\phi>$ is an integral, just like $<T_f,\phi>=<f,\phi>= \int_\Gamma f\phi$:

$$\begin{aligned}
<T(x-c),\phi(x)> &= \int_\Gamma T(x-c)\phi(x)dx \\
&= \int_\Gamma T(x)\phi(x+c)dx = <T(x),\phi(x+c)>, \\
<T(-x),\phi(x)> &= \int_\Gamma T(-x)\phi(x)dx \\
&= \int_\Gamma T(x)\phi(-x)dx = <T(x),\phi(-x)>,
\end{aligned}$$

$$<Tg,\phi> \;=\; <gT,\phi> = \int_\Gamma Tg\phi = <T,g\phi>\,.$$

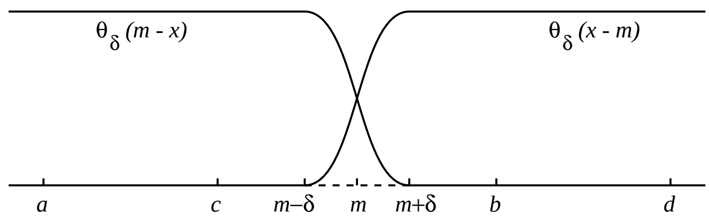

FIGURE 4.3

Here the integral signs have been used symbolically. For given T, products Tg may often be defined for less regular functions g than test functions; cf. the Exercises and Section 4.7.

An important notion is the concept of *local equality.*

DEFINITION 4.2.8. One says that $T = 0$ on the open set $\Omega \subset \Gamma$ if $< T, \phi > = 0$ for all test functions ϕ with support in Ω.

By the preceding the distribution δ_Γ is even: $\delta_\Gamma(-x) = \delta_\Gamma(x)$. Furthermore

$$\delta_\Gamma(x) = 0 \text{ on } 0 < x < 2\pi \text{ (and on } -2\pi < x < 0). \tag{4.2.4}$$

Indeed, $< \delta_\Gamma, \phi > = \phi(0) = 0$ whenever $\operatorname{supp}\phi \subset (0, 2\pi)$. The distribution δ_Γ has the single point $\{0\}$ as its "support" on Γ. It follows that δ_Γ cannot be equal to an integrable function f on Γ: if f has support 0 then $< f, \phi >$ $= \int_\Gamma f\phi = 0$ for all ϕ. [The distribution or measure δ_Γ corresponds to the "mass distribution" that consists of a single point mass 1 at the point 0.]

We will need the following property: If $T = 0$ on Ω_1 and on Ω_2, then $T = 0$ on the union $\Omega_1 \cup \Omega_2$. Since open sets on Γ are unions of disjoint open intervals, it is sufficient to show that "$T = 0$ on (a, b)" and "$T = 0$ on (c, d)", with $a < c < b < d$, implies "$T = 0$ on (a, d)". In order to prove the latter, we decompose a given test function ϕ with support in (a, d) as $\phi_1 + \phi_2$, with $\operatorname{supp}\phi_1 \subset (a, b)$ and $\operatorname{supp}\phi_2 \subset (c, d)$. Such a decomposition may be obtained by setting

$$\phi_1(x) = \theta_\delta(m - x)\phi(x), \quad \phi_2(x) = \theta_\delta(x - m)\phi(x),$$

where $m = \frac{1}{2}(b + c)$, $\delta = \frac{1}{2}(b - c)$ and θ_δ is as in Examples 4.1.9; cf. Figure 4.3. It now follows that

$$< T, \phi > = < T, \phi_1 > + < T, \phi_2 > = 0$$

whenever $T = 0$ on (a, b) and (c, d).

By the preceding, there is a *maximal* open subset $\Omega \subset \Gamma$ on which a given distribution T is equal to zero. [Ω may of course be empty.] The complement of Ω in Γ is called the *support* of T.

Exercises. 4.2.1. Prove that a trigonometric series $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ represents a $\mathcal{C}^\infty$ function ϕ on Γ if (and only if) for every $p \in \mathbb{N}$, there is a constant B_p such that $|c_n| \le B_p/|n|^p$ for all $n \ne 0$.

4.2.2. Compute $< \delta_\Gamma, e^{-inx} >$ for each n. Use the result to verify that there can be no integrable function f such that $\delta_\Gamma = f$ on Γ.

4.2.3. Show that the formula

$$T(\phi) = a_0\phi(0) + a_1\phi'(0) + \cdots + a_m\phi^{(m)}(0), \quad \forall\, \phi \in \mathcal{D}(\Gamma)$$

defines a distribution T on Γ. Determine $\operatorname{supp} T$.

4.2.4. Let f be an integrable function on Γ. Prove that $f = 0$ on $(a,b) \subset \Gamma$ in the sense of distributions if and only if $f(x) = 0$ a.e. on (a,b).

4.2.5. Prove that the functionals $T(x-c)$, $T(-x)$ and Tg in Definition 4.2.7 are continuous on $\mathcal{D}(\Gamma)$.

4.2.6. For $T \in \mathcal{D}(\Gamma)$ and $g \in \mathcal{C}^\infty(a,b)$ one may define a product Tg on every interval (α,β) with $[\alpha,\beta] \subset (a,b)$ and $\beta-\alpha < 2\pi$ in the following way. Extending the restriction of g to $[\alpha,\beta]$ to a test function h on Γ with support in a subinterval $[\alpha-\varepsilon,\beta+\varepsilon]$ of (a,b) of length $< 2\pi$, one sets $Tg = Th$ on (α,β). Show that this definition is independent of the extension h.

4.2.7. Prove that $\delta_\Gamma g = g(0)\delta_\Gamma$ for every test function g. Also show that $x \cdot \delta_\Gamma(x) = 0$ on $(-\pi,\pi)$.

4.2.8. Prove that $\mathrm{pv}_\Gamma\,(1/x) = 1/x$ on $(-\pi,0)$ and on $(0,\pi)$. Also show that $x \cdot \mathrm{pv}_\Gamma\,(1/x) = 1$ on $(-\pi,\pi)$.

4.2.9. Show that

$$\{\delta_\Gamma(x)\,x\}\,\mathrm{pv}_\Gamma\,\frac{1}{x} \ne \delta_\Gamma(x)\left\{x \cdot \mathrm{pv}_\Gamma\,\frac{1}{x}\right\} \quad \text{on} \;\; (-\pi,\pi).$$

Thus distributional multiplication is not associative in general.

4.2.10. Verify that for $T \in \mathcal{D}_\Gamma$, the definition of the translate $T(x-c)$ under Definition 4.2.7 implies that indeed $T(x+2\pi) = T(x)$, as it reasonably should.

4.3. Distributional convergence

For the time being we deal only with distributions on Γ.

Definition 4.3.1. One says that distributions T_k (which may be equal to integrable functions) converge to a distribution T on Γ if $T_k \to T$ relative to the test class $\mathcal{D}(\Gamma)$:

$$< T_k, \phi > \to < T, \phi > \quad \text{as} \quad k \to \infty, \quad \forall\, \phi \in \mathcal{D}(\Gamma).$$

With this definition of convergence the distributions on Γ form the space $\mathcal{D}'(\Gamma)$, the so-called *dual space* of $\mathcal{D}(\Gamma)$.

We use a corresponding definition for "directed families" T_λ. Here λ runs over a real or complex index set Λ, and tends to a limit λ_0, which may be infinity. Similarly for convergence $T_k \to T$ on $(a, b) \subset \Gamma$. Distributional limits are unique: if $T_k \to T$ and also $T_k \to \tilde{T}$, then $\tilde{T} = T$. For integrable functions f_k and f on Γ, distributional convergence $f_k \to f$ is the same as the earlier convergence relative to the test class $\mathcal{C}^\infty(\Gamma)$.

Examples 4.3.2. A sequence or more general directed family of integrable functions which converges to the delta distribution on Γ is called a *delta sequence* or *delta family* on Γ. Concrete examples are provided by

(i) the Dirichlet kernels $D_k,\ k = 0, 1, 2, \cdots$ [Section 2.2],

(ii) the Fejér kernels $F_k,\ k = 1, 2, \cdots$ [Section 3.2],

(iii) the family given by the Poisson kernel $P_r,\ 0 \le r \nearrow 1$ [Section 3.5],

(iv) any family $g_\varepsilon(x) = (1/\varepsilon) g(x/\varepsilon),\ 1 \ge \varepsilon \searrow 0$,
generated by an integrable function g with support in $[-1, 1]$
and such that $\displaystyle\int_{-1}^{1} g(x)dx = 1$.

It is easy to verify that $D_k \to \delta_\Gamma$ on Γ. Indeed,

$$< D_k, \phi > = \int_{-\pi}^{\pi} D_k(t)\phi(t)dt = s_k[\phi](0) \to \phi(0) = < \delta_\Gamma, \phi > \tag{4.3.1}$$

for all $\phi \in \mathcal{D}(\Gamma)$ since the Fourier series for ϕ converges to ϕ. The proofs in the other cases are not difficult either. Delta families occur in many problems of approximation.

We next consider somewhat different examples.

Example 4.3.3. Suppose that the trigonometric series $\sum_{n=-\infty}^{\infty} d_n e^{inx}$ converges to a distribution T on Γ, that is, $s_k = \sum_{n=-k}^{k} d_n e^{inx} \to T$ as $k \to \infty$. Then in particular

$$< T, e^{-inx} > = \lim < s_k, e^{-inx} > = \lim \int_{-\pi}^{\pi} s_k(x) e^{-inx} dx = 2\pi d_n. \tag{4.3.2}$$

This formula will motivate the definition of distributional Fourier series.

EXAMPLE 4.3.4. We will show that for every distribution T on Γ, there is a distribution S such that

$$\lim_{h\to 0}\frac{T(x+h)-T(x)}{h}=S(x). \tag{4.3.3}$$

S is called the (distributional) derivative of T. In Section 4.5 we will introduce the derivative in a somewhat different manner.

**Existence of the limit* in (4.3.3). We show first that for every test function ϕ,

$$\begin{aligned}\left\langle\frac{T(x+h)-T(x)}{h},\phi(x)\right\rangle &= \left\langle T(x),\frac{\phi(x-h)-\phi(x)}{h}\right\rangle\\ &\to < T(x),-\phi'(x) > \quad \text{as } h\to 0.\end{aligned} \tag{4.3.4}$$

Indeed [replacing $-h$ by h], the difference

$$\begin{aligned}\frac{\phi(x+h)-\phi(x)}{h}-\phi'(x) &= \frac{1}{h}\int_0^h\{\phi'(x+t)-\phi'(x)\}dt\\ &= \int_0^1\{\phi'(x+hv)-\phi'(x)\}dv\end{aligned} \tag{4.3.5}$$

will converge to zero uniformly in x as $h\to 0$, because the periodic function ϕ' is uniformly continuous. The pth order derivative of the difference in (4.3.5) also converges uniformly to zero:

$$\begin{aligned}&\frac{\phi^{(p)}(x+h)-\phi^{(p)}(x)}{h}-\phi^{(p+1)}(x)\\ &=\int_0^1\{\phi^{(p+1)}(x+hv)-\phi^{(p+1)}(x)\}dv\to 0\end{aligned}$$

uniformly in x. Thus the test functions given by (4.3.5) converge to 0 in the sense of $\mathcal{D}(\Gamma)$ as $h\to 0$. Going back to $-h$, we find that

$$\frac{\phi(x-h)-\phi(x)}{h}\to -\phi'(x) \quad \text{in } \mathcal{D}(\Gamma).$$

Since T is continuous on $\mathcal{D}(\Gamma)$ relation (4.3.4) follows.

We now define a linear functional S on $\mathcal{D}(\Gamma)$ by

$$< S,\phi > = < T,-\phi' > = - < T,\phi' > . \tag{4.3.6}$$

The functional S is continuous: if $\phi_j \to \phi$ in $\mathcal{D}(\Gamma)$, then $\phi_j' \to \phi'$ in $\mathcal{D}(\Gamma)$, hence $< T, \phi_j' > \to < T, \phi' >$. Thus S is a distribution on Γ. Combining (4.3.4) and (4.3.6) we obtain

$$\left\langle \frac{T(x+h) - T(x)}{h}, \phi(x) \right\rangle \to < S, \phi >, \quad \forall\, \phi.$$

This proves relation (4.3.3).

Exercises. 4.3.1. Prove that all directed families in Examples 4.3.2 are delta families.

4.3.2. Let $\{g_\varepsilon\}$ be as in Examples 4.3.2. Prove that for $\varepsilon \searrow 0$,

$$\int_{-\pi}^{\pi} g_\varepsilon(t) f(x-t)dt \to f(x)$$

uniformly on Γ for every function $f \in \mathcal{C}(\Gamma)$.

4.3.3. Use Definition 4.3.1 to show that the series $\sum_{n=1}^{\infty} n^{100} \cos nx$ converges in $\mathcal{D}'(\Gamma)$.

4.4. Fourier series

Example 4.3.3 motivates the following

Definition 4.4.1. The (complex) Fourier series for the distribution T on Γ is the series

$$T \sim \sum_{n=-\infty}^{\infty} c_n[T] e^{inx} \text{ with } c_n[T] \stackrel{\text{def}}{=} \frac{1}{2\pi} < T, e^{-inx} >, \;\; \forall\, n \in \mathbb{Z}.$$

One can of course also introduce the "real" Fourier series. The partial sums of the real series and the symmetric partial sums $\sum_{-k}^{k}$ of the complex series are denoted by $s_k[T]$.

Example 4.4.2. The Fourier series for the delta distribution on Γ and its partial sums are

$$\delta_\Gamma \sim \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{inx} = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{n=1}^{\infty} \cos nx \right),$$

$$s_k[\delta_\Gamma] = \frac{1}{2\pi} \sum_{n=-k}^{k} e^{inx} = D_k(x) \quad \text{(Dirichlet kernel).} \tag{4.4.1}$$

The Fourier series for δ_Γ converges to δ_Γ. Indeed, $s_k[\delta_\Gamma] = D_k \to \delta_\Gamma$ by (4.3.1).

PROPOSITION 4.4.3. *For $T \in \mathcal{D}'(\Gamma)$ and $\phi \in \mathcal{D}(\Gamma)$,*

$$
\begin{aligned}
< T, \phi > &= 2\pi \lim_{k\to\infty} \sum_{n=-k}^{k} c_n[T]c_{-n}[\phi] \\
&= \sum_{n=-\infty}^{\infty} c_n[T]c_{-n}[\phi]. \qquad (4.4.2)
\end{aligned}
$$

PROOF. By Proposition 4.2.2 the Fourier series for ϕ converges to ϕ in $\mathcal{D}(\Gamma)$. Hence by the continuity of T,

$$
\begin{aligned}
< T, \phi > &= \lim < T, s_k[\phi] > = \lim \sum_{n=-k}^{k} c_n[\phi] < T, e^{inx} > \\
&= \lim 2\pi \sum_{n=-k}^{k} c_n[\phi]c_{-n}[T], \quad \text{etc.}
\end{aligned}
$$

□

One may derive from Section 4.6 below that the series in (4.4.2) is absolutely convergent.

THEOREM 4.4.4. *For every distribution T on Γ the Fourier series converges to T.*

PROOF. [Cf. Lemma 4.1.6 and Theorem 4.1.7] By Proposition 4.4.3,

$$
\begin{aligned}
< s_k[T], \phi > &= \sum_{n=-k}^{k} c_n[T] < e^{inx}, \phi > = \sum_{n=-k}^{k} c_n[T] \int_{-\pi}^{\pi} e^{inx}\phi(x)dx \\
&= 2\pi \sum_{n=-k}^{k} c_n[T]c_{-n}[\phi] \to < T, \phi >, \quad \forall\, \phi \in \mathcal{D}(\Gamma).
\end{aligned}
$$

□

COROLLARY 4.4.5. *A distribution T on Γ is determined by its Fourier series: if $c_n[T] = 0$ for all n, then $T = 0$.*

Exercises. 4.4.1. Suppose that $T_k \to T$ in $\mathcal{D}'(\Gamma)$ as $k \to \infty$. Prove that $c_n[T_k] \to c_n[T]$ for every n.

4.4.2. Express the Fourier series for $T(x - c)$ and $T(-x)$ in terms of the Fourier series for $T(x)$.

4.4.3. Prove that a distribution T on Γ is equal to a test function if and only if for every $p \in \mathbb{N}$, one has $c_n[T] = \mathcal{O}(|n|^{-p})$ as $|n| \to \infty$.

4.4.4. Compute the distributional sums of the series

$$\cos x + \cos 2x + \cos 3x + \cdots , \quad \cos x - \cos 2x + \cos 3x - \cdots ,$$
$$\cos x + \cos 3x + \cos 5x + \cdots .$$

4.4.5. It is more difficult to determine the distributional sum $T(x)$ of the series

$$\sin x + \sin 2x + \sin 3x + \cdots .$$

Denoting the partial sum of order k by s_k, prove that for all $\phi \in \mathcal{D}(\Gamma)$,

$$\begin{aligned} < T, \phi > &= \lim < s_k, \phi > = \int_0^\pi \frac{1}{2}\left(\cot \frac{1}{2}x\right)\{\phi(x) - \phi(-x)\}dx \\ &= \lim_{\varepsilon \searrow 0}\left(\int_{-\pi}^{-\varepsilon} + \int_{\varepsilon}^{\pi}\right)\frac{1}{2}\left(\cot \frac{1}{2}x\right)\phi(x)dx \\ &= \text{p.v.} \int_{-\pi}^{\pi} \frac{1}{2}\left(\cot \frac{1}{2}x\right)\phi(x)dx; \end{aligned} \tag{4.4.3}$$

cf. Exercise 3.2.3. The final expression defines the distribution pv $\frac{1}{2}\cot\frac{1}{2}x$. More in Exercise 4.5.12.

4.5. Derivatives of distributions

Let F be a function in $\mathcal{C}^1_{2\pi} = \mathcal{C}^1(\Gamma)$, or more generally, a function in $\mathcal{C}(\Gamma)$ that can be written as an indefinite integral (4.1.1). Then for every test function ϕ on Γ, integration by parts gives

$$< F', \phi > = \int_{-\pi}^{\pi} F'\phi = \Big[F\phi\Big]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} F\phi' = - < F, \phi' > .$$

We extend this formula to distributions:

Definition 4.5.1. Let T be any distribution on Γ. Then the *distributional derivative* DT of T is the distribution on Γ obtained by formal integration by parts:

$$< DT, \phi > \stackrel{\text{def}}{=} - < T, \phi' >, \quad \forall\, \phi \in \mathcal{D}(\Gamma).$$

This formula indeed defines $S = DT$ as a continuous linear functional on $\mathcal{D}(\Gamma)$, hence as a distribution; cf. the lines following (4.3.6).

Whereas the derivative F' of a $\mathcal{C}^1$ function F is defined at every point, the derivative DT of a distribution T need not have a value at any point;

it does not represent pointwise rate of change. Nevertheless distributional differentiation has a *local character* in the following sense:

$$\begin{aligned}&\text{if } T_1 = T_2 \text{ on } (a,b) \in \Gamma, \text{ then } DT_1 = DT_2 \text{ on } (a,b);\\ &\text{if } T = F \text{ on } (a,b), \text{ where } F \text{ is a } \mathcal{C}^1 \text{ function on } \Gamma\\ &\text{or just on } (a,b), \text{ then } DT = F' \text{ on } (a,b); \text{ in particular:}\\ &\text{if } T = C \text{ (a constant) on } (a,b), \text{ then } DT = 0 \text{ on } (a,b).\end{aligned} \tag{4.5.1}$$

Indeed, if $< T_1 - T_2, \phi >= 0$ for all test functions ϕ with support in (a,b), then $< D(T_1 - T_2), \phi >= - < T_1 - T_2, \phi' >= 0$ for all such functions ϕ, since ϕ' also is a test function with support in (a,b).

The final part of (4.5.1) has a converse which is *fundamental* for the distributional theory of *differential equations*:

$$\begin{aligned}&\text{if } DT = 0 \text{ on } \Gamma \text{ [or on } (a,b) \in \Gamma], \text{ then } T = C,\\ &\text{a constant function, on } \Gamma \text{ [or on } (a,b), \text{ respectively].}\end{aligned} \tag{4.5.2}$$

The first part is a nice application of Fourier series. Indeed,

$$\begin{aligned}2\pi c_n[DT] &=< DT, e^{-inx} >= - < T, (e^{-inx})' >\\ &= in < T, e^{-inx} >= 2\pi inc_n[T],\end{aligned} \tag{4.5.3}$$

hence if $DT = 0$ on Γ, then $inc_n[T] = c_n[DT] = 0$ for all n, so that $c_n[T] = 0$ for all $n \neq 0$. Thus

$$T = \sum_{n=-\infty}^{\infty} c_n[T]e^{inx} = c_0[T] \quad \text{on } \Gamma.$$

For the "local part" of (4.5.2), see Exercises 4.5.6, 4.5.7.

All distributions T on Γ will have derivatives of every order. In particular, integrable functions f acquire (distributional) derivatives of every order. Of course the derivative Df is not a function unless f is equal to an indefinite integral.

PROPERTY 4.5.2. (Product Rule) For $T \in \mathcal{D}'(\Gamma)$ and $g \in \mathcal{D}(\Gamma)$,

$$D(Tg) = DT \cdot g + Tg' \quad \text{on } \Gamma. \tag{4.5.4}$$

Indeed, for any test function ϕ, by Definition 4.2.7,

$$\begin{aligned}&< D(Tg), \phi >= - < Tg, \phi' >= - < T, g\phi' >\\ &= - < T, (g\phi)' - g'\phi >=< DT, g\phi > + < T, g'\phi >\\ &=< DT \cdot g, \phi > + < Tg', \phi > .\end{aligned}$$

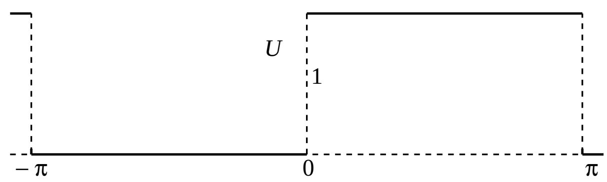

FIGURE 4.4

EXAMPLES 4.5.3. Let U be the *unit step function*, here restricted to the interval $(-\pi, \pi)$:

$$U(x) = 1_+(x) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{on } (-\pi, 0), \\ 1 & \text{on } (0, \pi); \end{cases}$$

we usually set $U(0) = 0$, but the value at 0 is irrelevant for our integrals. The 2π-periodic function U^{per} defined by U (cf. Figure 4.4) also has jumps at $\pm\pi$, etc, but we only wish to study $DU^{\text{per}} = DU$ on $(-\pi, \pi)$. Hence let ϕ be any test function with support on $[a, b] \subset (-\pi, \pi)$ where $a < 0 < b$. Then

$$\begin{aligned} < DU^{\text{per}}, \phi > &= - < U^{\text{per}}, \phi' > = - \int_a^b U^{\text{per}} \phi' = - \int_0^b \phi' \\ &= -\phi(b) + \phi(0) = \phi(0) = < \delta_\Gamma, \phi > . \end{aligned}$$

Thus

$$DU = DU^{\text{per}} = \delta_\Gamma \quad \text{on } (-\pi, \pi). \tag{4.5.5}$$

Observe that the distributional derivative is equal to the ordinary derivative - zero - on $(-\pi, 0)$ and on $(0, \pi)$, as it should be; cf. (4.5.1). The *delta distribution* in the answer *reflects the jump* 1 of U at the origin.

Since $xU(x)$ is an indefinite integral of U on $(-\pi, \pi)$, one has $D(xU) = U$ there. Application of (4.5.4) thus shows that $x\delta_\Gamma = xDU = 0$ on $(-\pi, \pi)$; cf. Exercise 4.2.7.

By Example 4.2.6 one has $< \text{pv}_\Gamma\,(1/x), \phi > = - < \log(|x|/\pi), \phi' >$ for all test functions ϕ, hence

$$\text{pv}_\Gamma \frac{1}{x} = D \log \frac{|x|}{\pi} = D \log |x| \quad \text{on } (-\pi, \pi).$$

THEOREM 4.5.4. (Continuity of distributional differentiation) *Suppose that $T_k \to T$ on* Γ [or on $(a, b) \subset \Gamma$]. *Then $DT_k \to DT$ on* Γ [or on $(a, b) \subset \Gamma$, respectively].

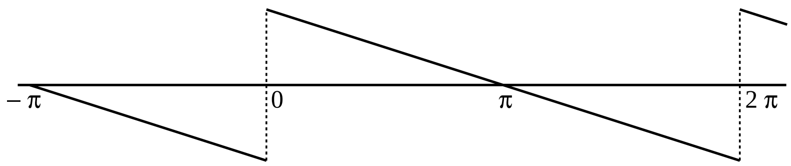

FIGURE 4.5

Indeed, the derivative of a test function [with support in (a,b)] is also a test function [with support in (a,b)], hence

$$< DT_k, \phi > = - < T_k, \phi' > \to \; - < T, \phi > = < DT, \phi > .$$

COROLLARY 4.5.5. (Termwise differentiation) *Every distributionally convergent series on* Γ [or on $(a,b) \subset \Gamma$] *may be differentiated term by term:*

$$\textit{if } T = \sum_{n=0}^{\infty} U_n, \quad \textit{then } DT = \sum_{n=0}^{\infty} DU_n.$$

In particular Fourier series of distributions may be differentiated term by term:

$$\text{if } T = \sum_{n=-\infty}^{\infty} c_n e^{inx} \left(= \lim_{k\to\infty} \sum_{n=-k}^{k} \right),$$
$$\text{then } DT = \sum_{n=-\infty}^{\infty} inc_n e^{inx}.$$

[We knew already that $c_n[DT] = inc_n[T]$; see (4.5.3).]

COROLLARY 4.5.6. *Let* $\sum_{n=0}^{\infty} g_n$ *be a uniformly or* L^1 *convergent series of integrable functions on* Γ *with sum* f. *Then*

$$\sum_{n=0}^{\infty} D^p g_n = D^p f \quad \textit{on } \Gamma, \quad \forall\, p \in \mathbb{N}.$$

EXAMPLE 4.5.7. Let f be the 2π-periodic function given by

$$f(x) = \frac{\pi - x}{2} \quad \text{on } (0, 2\pi), \quad \text{or } f(x) = -\frac{x}{2} + \frac{\pi}{2}\operatorname{sgn} x \quad \text{on } (-\pi, \pi);$$

cf. Figure 4.5. The Fourier series of this integrable function is $\sum_{n=1}^{\infty} \frac{1}{n} \sin nx$. It converges to f in L^1, hence distributionally; cf. Exercises 1.2.1, 1.2.2 or

Section 2.6. Thus it may be differentiated term by term to give

$$\sum_{n=1}^{\infty} \cos nx = Df(x).$$

On $(0, 2\pi)$ the function f is of class $\mathcal{C}^1$, hence $Df = f' = -\frac{1}{2}$. At 0 our f has a jump equal to π; on $(-\pi, \pi)$, the difference $f - \pi U$ is equal to the $\mathcal{C}^1$ function $-\frac{1}{2}x - \frac{1}{2}\pi$ when properly defined at 0. Thus $D(f - \pi U) = Df - \pi\delta_\Gamma = -\frac{1}{2}$ on $(-\pi, \pi)$. Conclusion:

$$\sum_{n=1}^{\infty} \cos nx = Df(x) = -\frac{1}{2} + \pi\delta_\Gamma \text{ on } \Gamma,$$

in accordance with the known Fourier series for δ_Γ in Example 4.4.2.

We could also have *started* with the Fourier series for δ_Γ. From it, we could have computed the sum of the series $\sum_{n=1}^{\infty} \frac{1}{n} \sin nx$; cf. Exercise 4.5.12.

Exercises. 4.5.1. Let f be an integrable function on Γ. Show that the distributional derivative $D^p f$ may be represented by the formula

$$< D^p f, \phi > = (-1)^p \int_\Gamma f\phi^{(p)}, \quad \forall\, \phi \in \mathcal{D}(\Gamma).$$

Verify that this formula gives $D^p f$ as a continuous linear functional on $\mathcal{D}(\Gamma)$.

4.5.2. Let T be a distribution on Γ such that $T = F$ on $(a, b) \subset \Gamma$, where F is a $\mathcal{C}^p$ function on (a, b). Starting with $p = 1$, prove that $D^p T = F^{(p)}$ on (a, b).

4.5.3. Compute $D^p(x^q U)$ on $(-\pi, \pi)$, (i) if $p \le q$, (ii) if $p > q$.

4.5.4. Compute $\sum_{n=-\infty}^{\infty} n^p e^{inx}$ on Γ for all $p \in \mathbb{N}$.

4.5.5. Show that a distribution T on Γ has a distributional antiderivative on Γ if and only if $c_0[T] = 0$.

4.5.6. Let ϕ be any test function on Γ with support in $(a, b) \subset \Gamma$, and let ω be a fixed test function with support in (a, b) and $\int_\Gamma \omega = 1$. Prove that $\phi - c\omega$ will be the derivative ψ' of a test function ψ [with support in (a, b)] if and only if $c = \int_\Gamma \phi = < 1, \phi >$.

4.5.7. (Continuation) Let T be a distribution on Γ such that $DT = 0$ on $(a, b) \subset \Gamma$. Prove that $T = C$ on (a, b).

Hint. Take ϕ and ω as above and form $< T, \phi - c\omega >$.

4.5.8. Discuss the distributional differential equation $(D - a)u = 0$, (i) on Γ; (ii) on $(-\pi, \pi)$.

Hint. $(D - a)u = e^{ax} D(e^{-ax} u)$ on $(-\pi, \pi)$.

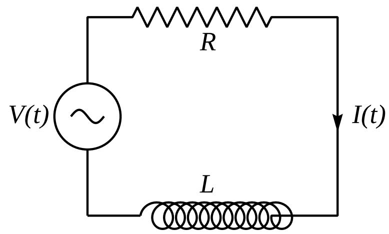

FIGURE 4.6

4.5.9. Let f be an integrable function on $(-\pi, \pi)$. Show that all distributional solutions u [in $\mathcal{D}'(\Gamma)$] of the following differential equations are equal to ordinary functions on $(-\pi, \pi)$:

$$\text{(i) } (D-a)u = f \ \text{ on } \ (-\pi,\pi); \quad \text{(ii) } (D-a)u = Df \ \text{ on } \ (-\pi,\pi).$$

4.5.10. Consider an electric circuit containing a resistance R and an inductance L in series with a generator that supplies a voltage $V(t)$ (Figure 4.6). Here the current $I(t)$ will satisfy the differential equation

$$L\frac{dI}{dt} + RI = V(t).$$

Determine the current over a time interval $-b < t < b$ when $V(t)$ is a unit voltage impulse at $t = 0$, while $I(t) = 0$ for $t < 0$.

Hint. Taking $b = \pi$, a unit voltage impulse at $t = 0$ may be represented by $V(t) = \delta_\Gamma(t)$.

4.5.11. Let f be 2π-periodic, continuous on $[-\pi, \pi]$ except for a jump at the point $c \in (-\pi, \pi)$ and such that the restriction of f to $[-\pi, c)$ could be extended to a $\mathcal{C}^1$ function on the closure of this interval, and similarly for the restriction of f to $(c, \pi]$. Prove that

$$Df(x) = f'(x) + \{f(c+) - f(c-)\}\delta_\Gamma(x-c) \quad \text{on } \Gamma.$$

[*Conclusion:* the distributional derivative contains more information than the ordinary derivative!]

4.5.12. In Exercise 4.4.5 it was found that $\sum_{n=1}^\infty \sin nx = \mathrm{pv}\,\frac{1}{2}(\cot\frac{1}{2}x)$ on Γ. Prove that for all $\phi \in \mathcal{D}(\Gamma)$,

$$\begin{aligned}\left\langle \mathrm{pv}\,\frac{1}{2}\cot\frac{1}{2}x, \phi(x)\right\rangle &\stackrel{\text{def}}{=} \text{p.v.} \int_{-\pi}^{\pi} \frac{1}{2}\left(\cot\frac{1}{2}x\right)\phi(x)dx \\ &= -\int_{-\pi}^{\pi} \log|\sin(x/2)| \cdot \phi'(x)dx.\end{aligned}$$

Put into words what this means. Finally compute $\sum_{n=1}^{\infty} \frac{1}{n} \cos nx$ on Γ.

4.5.13. Let U^{per} be the 2π-periodic extension of the unit step function U on $(-\pi, \pi)$. Compute the real Fourier series for U^{per} and express DU^{per} on Γ [not just on $(-\pi, \pi)$] in terms of delta distributions.

4.5.14. Let f be an integrable function on $(-\pi, \pi)$. Prove that the series, obtained by differentiation of the Fourier series for f, is the Fourier series for Df^{per}, where f^{per} is the 2π-periodic extension of f.

4.6. Structure of periodic distributions

We begin with a characterization of the class of distributionally convergent trigonometric series.

PROPOSITION 4.6.1. *A trigonometric series*

$$\sum_{n=-\infty}^{\infty} d_n e^{inx} \tag{4.6.1}$$

converges in $\mathcal{D}'(\Gamma)$, that is, $\lim_{k\to\infty} \sum_{-k}^{k} d_n e^{inx}$ exists as a distribution, if and only if there are constants B and β such that

$$|d_n| \le B|n|^{\beta}, \quad \forall\, n \neq 0. \tag{4.6.2}$$

PROOF. (i) Suppose that the numbers d_n satisfy the inequalities (4.6.2) and let s be the smallest nonnegative integer greater than $\beta + 1$. Then the inequalities

$$\left| \frac{d_n}{(in)^s} \right| \le \frac{B}{|n|^{s-\beta}} \quad (\text{where } \; s - \beta > 1)$$

show that the series

$$\sum_{n\neq 0} \frac{d_n}{(in)^s}\, e^{inx}$$

is [absolutely and] uniformly convergent on Γ. The sum function $f = f(x)$ of that series is continuous, and by differentiation [as in Corollary 4.5.6] we find that the series (4.6.1) converges to the distribution

$$T = d_0 + D^s f \quad \text{on } \; \Gamma. \tag{4.6.3}$$

(ii) Suppose now that the series (4.6.1) converges to a distribution T on Γ, that is, for $k \to \infty$,

$$\left\langle \sum_{n=-k}^{k} d_n e^{inx}, \phi \right\rangle \to < T, \phi >, \quad \forall\, \phi \in \mathcal{D}(\Gamma).$$

Thus the series

$$\sum_{n=-\infty}^{\infty} d_n c_{-n}[\phi] \tag{4.6.4}$$

must converge for every test function ϕ; cf. Proposition 4.4.3. We will use this fact to obtain an indirect proof for the validity of inequalities of the form (4.6.2).

Suppose to the contrary that the sequence

$$\{n^{-j} d_n\}, \quad n = \pm 1, \pm 2, \cdots \tag{4.6.5}$$

is unbounded for every positive integer j. Then, starting with $j = 1$, there must be an integer n_1 of (smallest) absolute value $|n_1| \geq 1$ such that $|n_1^{-1} d_{n_1}| > 1$. Taking $j = 2$, there must be an integer n_2 of (smallest) absolute value $|n_2| > |n_1|$ such that $|n_2^{-2} d_{n_2}| > 1$. In general, there exists an integer n_j of (smallest) absolute value $|n_j| > |n_{j-1}|$ such that $|n_j^{-j} d_{n_j}| > 1$. Clearly $|n_j| \geq j$ for all $j \in \mathbb{N}$. Now define

$$\phi(x) = \sum_{j=1}^{\infty} n_j^{-j} e^{-in_j x}.$$

Since $|n_j^{-j}| \leq j^{-j} \leq j^{-2}$ for every $j \geq 2$, the series for ϕ is (absolutely and) uniformly convergent, hence ϕ is well-defined and continuous. The function ϕ will actually be of class $\mathcal{C}_{2\pi}^{\infty}$. Indeed, for every positive integer p, the p times differentiated series

$$\sum_{j=1}^{\infty} (-i)^p n_j^{p-j} e^{-in_j x}$$

is also (absolutely and) uniformly convergent, since $|n_j^{p-j}| \leq j^{p-j} \leq j^{-2}$ as soon as $j \geq p+2$.

However, for our special test function ϕ, the series in (4.6.4) will be divergent. Indeed, $c_{-n}[\phi] = n_j^{-j}$ for $n = n_j$ and $c_{-n}[\phi] = 0$ when n does not have the form n_k for some k. Hence

$$\sum_{n=-|n_k|}^{|n_k|} d_n c_{-n}[\phi] = \sum_{j=1}^{k} d_{n_j} n_j^{-j},$$

and the latter sums are the partial sums of an infinite series, all of whose terms have absolute value greater than one.

This contradiction proves that there must be a positive integer j for which the sequence (4.6.5) is bounded. $\square$

THEOREM 4.6.2. (Structure of the distributions on Γ) *Let T be an arbitrary distribution on Γ. Then there exist a continuous function f on Γ, a nonnegative integer s and a constant d_0 ($= c_0[T]$) such that T has the representation $T = d_0 + D^s f$ given in* (4.6.3)*.*

PROOF. The Fourier series for T converges to T in $\mathcal{D}'(\Gamma)$; see Theorem 4.4.4. Hence by Proposition 4.6.1, there are constants B and β such that the Fourier coefficients $c_n[T] = d_n$ satisfy the inequalities (4.6.2). The first part of the proof of Proposition 4.6.1 now shows that T has a representation (4.6.3). $\square$

Order of a distribution. The smallest nonnegative integer s for which T has a representation (4.6.3) with a continuous or integrable function f on Γ may be called the *order* of T relative to the continuous or integrable functions on Γ.

*However, Laurent Schwartz used a somewhat different definition; cf. [**110**]. Observe that the representation (4.6.3) implies that $| < T, \phi > |$ can be bounded in terms of $\sup|\phi|$ and $\sup|\phi^{(s)}|$. The *Schwartz order* of T is the smallest nonnegative integer m such that $| < T, \phi > |$ can be bounded in terms of $\sup|\phi|$ and $\sup|\phi^{(m)}|$. Thus δ_Γ is a distribution of order zero.

*By Theorem 3.3.2, every continuous function on Γ is a uniform limit of trigonometric polynomials. Anticipating the terminology of normed spaces [see Chapter 5], we may conclude that every function $\phi \in C(\Gamma)$ is a limit of test functions under the distance derived from the norm $\|\phi\| = \sup|\phi|$. It follows that a distribution T of order zero can be extended by continuity to a continuous linear functional on $C(\Gamma)$. Hence by Riesz's representation theorem, every distribution of order zero can be identified with a measure; cf. Examples 4.2.5.

The following refinement of Proposition 4.6.1 provides a *characterization of distributional convergence $T_k \to T$.*

THEOREM 4.6.3. *The following three statements about distributions T_k, $k = 1, 2, \cdots$ and T on Γ are equivalent:*

(*i*) $T_k \to T$ *in* $\mathcal{D}'(\Gamma)$, *that is,* $< T_k, \phi > \to < T, \phi >$, $\forall\, \phi \in \mathcal{D}(\Gamma)$;

(*ii*) $c_n[T_k] \to c_n[T]$ *as* $k \to \infty$, $\forall\, n \in \mathbb{Z}$,
and there are constants B *and* β *such that*
$|c_n[T_k]| \le B|n|^\beta, \ \ \forall\, n \ne 0, \ \forall\, k \in \mathbb{N};$

(*iii*) *There are continuous functions* f_k *and* f *on* Γ,
a nonnegative integer s *and constants* d_{k0} *and* d_0 *such that*
$T_k = d_{k0} + D^s f_k, \quad T = d_0 + D^s f,$
while $f_k \to f$ *uniformly on* Γ *and* $d_{k0} \to d_0$.

Instead of sequences $\{T_k\}$ one may consider more general directed families $\{T_\lambda\}$. Convergence $T_k \to T$ according to (*i*) is sometimes called *weak* convergence, while convergence according to (*ii*) or (*iii*) may be called *strong* convergence. For distributions, weak and strong convergence are equivalent.

The difficult part in the proof of Theorem 4.6.3 is the implication (*i*) $\Rightarrow$ (*ii*). It may be derived from

PROPOSITION 4.6.4. *Let* $a_n^{(\lambda)}$, $n = 1, 2, \cdots$, $\lambda \in \Lambda$ *be a family of sequences with the following property. For every sequence* $b = \{b_n\}$ *such that* $b_n = \mathcal{O}(n^{-p})$ *for every* p, *the associated sums*

$$\sigma(\lambda) = \sigma(\lambda, b) = \sum_{n=1}^{\infty} a_n^{(\lambda)} b_n$$

are well-defined and form a bounded set $E = E(b)$ *as* λ *runs over* Λ. *Then there are constants* A *and* α *such that*

$$M_n \stackrel{\text{def}}{=} \sup_{\lambda \in \Lambda} |a_n^{(\lambda)}| \le A n^\alpha, \quad \forall\, n \in \mathbb{N}. \tag{4.6.6}$$

A proof will be sketched in Exercise 4.6.13. The proposition also implies the completeness of the space $\mathcal{D}'(\Gamma)$, cf. Exercise 4.6.11:

THEOREM 4.6.5. *Let* $\{T_k\}$ *be a Cauchy sequence in* $\mathcal{D}'(\Gamma)$, *that is to say,* $< T_j - T_k, \phi > \to 0$ *as* j, $k \to \infty$ *for every test function* ϕ *on* Γ. *Then the sequence* $\{T_k\}$ *converges to a distribution* T *on* Γ.

REMARK 4.6.6. In Section 5.2 we will discuss a general construction of *completion* of metric spaces. A similar construction can be used to complete the space $\mathcal{L}(\Gamma)$ of the integrable functions on Γ under the concept of convergence relative to test functions. This provides another way to arrive at the distribution space $\mathcal{D}'(\Gamma)$; cf. [**68**].

Exercises. 4.6.1. Prove that a series $\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$ is the (real) Fourier series of a distribution T on Γ if and only if there are constants B and β such that $|a_n| + |b_n| \le Bn^\beta$ for all $n \in \mathbb{N}$.

4.6.2. Prove that the series $2\pi \sum_{n=-\infty}^{\infty} c_n[T]c_{-n}[\phi]$ for $< T, \phi >$ in Proposition 4.4.3 is absolutely convergent.

4.6.3. (*Antiderivative*) Prove that a distribution T on Γ has an antiderivative in $\mathcal{D}'(\Gamma)$ if and only if $c_0[T] = 0$. What can you say about the order of the antiderivative(s)?

4.6.4. Suppose $T = d_0 + D^s f$ on Γ with f integrable. Prove that there are constants B_0 and B_s (depending on T) such that

$$| < T, \phi > | \le B_0 \sup |\phi| + B_s \sup |\phi^{(s)}|, \quad \forall\, \phi \in \mathcal{D}(\Gamma).$$

4.6.5. Represent δ_Γ in the form (4.6.3), (i) with $s = 1$ and f integrable; (ii) with $s = 2$ and f continuous. Show that δ_Γ is a distribution of (Schwartz) order zero and that $D^m\delta_\Gamma$ has order m.

4.6.6. (*Characterization of distributions with point support*) Let T be a distribution on Γ whose support is the point 0. Prove that on $(-\pi, \pi)$, $T = D^s F$ for some continuous function F and $s \ge 1$. Using the fact that $D^s F$ must vanish on $(-\pi, 0)$ and on $(0, \pi)$, what can you say about F on those intervals?

Derive that on $(-\pi, \pi)$, one can write $T = D^s(PU)$, where P is a polynomial of degree $< s$ and U is the unit step function. Finally show that on $(-\pi, \pi)$ [and in fact, on Γ], T can be represented in the form

$$a_0\delta_\Gamma + a_1 D\delta_\Gamma + \cdots + a_m D^m\delta_\Gamma, \quad \text{with} \;\; a_m \ne 0,$$

where m is the Schwartz order of T.

4.6.7. A distribution T is called positive if $T(\psi) \ge 0$ for all test functions $\psi \ge 0$. Prove that a positive distribution on Γ has Schwartz order zero. [Hence it can be identified with a measure.]

Hint: if ϕ is an arbitrary real test function and $\sup |\phi| = \gamma$, then the functions $\gamma \cdot 1 \pm \phi$ are nonnegative test functions.

4.6.8. Show that distributions of Schwartz order m can be extended to continuous linear functionals on the space $\mathcal{D}^m(\Gamma)$, obtained from $\mathcal{C}^m(\Gamma)$ by imposing the norm $\|\phi\| = \sup |\phi| + \sup |\phi^{(m)}|$.

[Convergence $\phi_j \to \phi$ in $\mathcal{D}^m(\Gamma)$ is equivalent to uniform convergence $\phi_j \to \phi$, $\phi_j' \to \phi'$, $\cdots$, $\phi_j^{(m)} \to \phi^{(m)}$.]

4.6.9. Let T be a distribution on Γ. Prove the existence of, and compute, $\lim_{h\to 0}\{T(x+h)-T(x)\}/h$ with the aid of Fourier series.

4.6.10. Prove that statement (ii) in Theorem 4.6.3 implies statement (iii), and that (iii) implies (i).

4.6.11. Use Proposition 4.6.4 to prove the implication (i)$\Rightarrow$(ii) in Theorem 4.6.3 and also to prove Theorem 4.6.5.

Hint. If $c_n[\phi]=0$ for all $n\geq 0$ one has

$$\sigma(k)=\sigma(k,\phi)=<T_k,\phi>=2\pi\sum_{n=1}^{\infty}c_n[T_k]c_{-n}[\phi].$$

4.6.12. Suppose that the sequence $\{a_n\}$, $n=1,2,\cdots$ has the following property: the series $\sum_{n=1}^{\infty}a_nb_n$ converges for every sequence $\{b_n\}$ such that $b_n=\mathcal{O}(n^{-p})$ for every p. Show that there must be constants β and n_0 such that $|a_n|\leq n^{\beta}$ for all $n\geq n_0$.

*4.6.13. Fill in the details in the following sketch of a proof for Proposition 4.6.4. For every $\lambda\in\Lambda$ there will be constants $\beta(\lambda)$ and $\nu(\lambda)$ such that $|a_n^{(\lambda)}|\leq n^{\beta(\lambda)}$ for all $n\geq\nu(\lambda)$. Supposing now that (4.6.6) fails for all pairs (A,α), there exist, for each pair (A_j,α_j), an arbitrarily large integer n_j and a parameter value λ_j such that

$$\left|a_{n_j}^{(\lambda_j)}\right|>A_jn_j^{\alpha_j}. \tag{4.6.7}$$

Setting $b_{n_j}=n_j^{-\alpha_j}$ and $b_n=0$ for n different from all n_k, one may inductively determine A_j, $\alpha_j\nearrow\infty$, $n_j\nearrow\infty$ and $\lambda_j\in\Lambda$ as follows. Start with $A_1=\alpha_1=1$ and select n_1 and λ_1 in accordance with (4.6.7). For $j\geq 2$, choose A_j such that the final inequality in (4.6.8) below is satisfied, and take $\alpha_j=\max\{\alpha_{j-1}+1,\beta(\lambda_{j-1})+2\}$. Finally choose $n_j\geq\max\{n_{j-1}+1,\nu(\lambda_{j-1})\}$ and λ_j such that (4.6.7) holds. As a result one has

$$\begin{aligned}|\sigma(\lambda_j,b)|&=\left|\sum_{n=1}^{\infty}a_n^{(\lambda_j)}b_n\right|\\&\geq\left|a_{n_j}^{(\lambda_j)}b_{n_j}\right|-\sum_{k<j}\left|a_{n_k}^{(\lambda_j)}b_{n_k}\right|-\sum_{k>j}\left|a_{n_k}^{(\lambda_j)}b_{n_k}\right|\\&>A_j-\sum_{k<j}M_{n_k}n_k^{-\alpha_k}-\sum_{k>j}n^{\beta(\lambda_j)-\alpha_k}\\&>A_j-\sum_{k<j}M_{n_k}n_k^{-\alpha_k}-\sum_{k>j}k^{-2}>j,\quad\forall j\geq 2,\end{aligned} \tag{4.6.8}$$

which contradicts the boundedness of the family $\{\sigma(\lambda, b)\}$ for our b.

4.7. Product and convolution of distributions

The unlimited differentiability of distributions has a price: within the class $\mathcal{D}(\Gamma)$, multiplication is not generally possible. This is not too surprising since distributions are generalizations of integrable functions. The product of two integrable functions need not be integrable, and there is no general method to associate a distribution with a nonintegrable function. Where multiplication of distributions *is* defined, it need not be associative; cf. Exercise 4.2.9.

The product of a given distribution T and a test function g is always defined; see Definition 4.2.7. What other products Tg can be formed depends on the order of T: the higher its order, the smoother g must be. This becomes plausible through formal multiplication of the Fourier series:

$$Tg(x) = \sum_k c_k[T]e^{ikx} \sum_l c_l[g]e^{ilx} = \sum_n \left\{ \sum_{k+l=n} c_k[T]c_l[g] \right\} e^{inx},$$

hence one would like to be able to define

$$c_n[Tg] = \sum_{k=-\infty}^{\infty} c_k[T]c_{n-k}[g]. \tag{4.7.1}$$

EXAMPLE 4.7.1. Using Example 4.4.2 for $c_k[\delta_\Gamma]$, formula (4.7.1) gives

$$c_n[\delta_\Gamma g] = \sum_k c_k[\delta_\Gamma]c_{n-k}[g] = \frac{1}{2\pi}\sum_k c_{n-k}[g] = \frac{1}{2\pi}g(0) = g(0)c_n[\delta_\Gamma].$$

This implies that

$$\delta_\Gamma g = g(0)\delta_\Gamma,$$

at least for all $\mathcal{C}^1$ functions g.

Actually, a distribution of (Schwartz) order m can be multiplied by any $\mathcal{C}^m$ function g; cf. Exercise 4.7.2.

The sequence $\{c_n[Tg]\}$ given by (4.7.1) is called the convolution of the sequences $\{c_n[T]\}$ and $\{c_n[g]\}$; convolution of sequences is not always possible. However, the dual operation, where one multiplies corresponding Fourier coefficients, always leads to another disttribution.

DEFINITION 4.7.2. The *convolution* $S * T$ of distributions S and T on Γ is the distribution given by

$$S * T = T * S = 2\pi \sum_{n=-\infty}^{\infty} c_n[S]c_n[T]e^{inx}. \tag{4.7.2}$$

The convolution $S * T$ is well-defined: if $c_n[S] = \mathcal{O}(|n|^\alpha)$ and $c_n[T] = \mathcal{O}(|n|^\beta)$ as $|n| \to \infty$, then $c_n[S * T] = \mathcal{O}(|n|^{\alpha+\beta})$, hence the series for $S * T$ is distributionally convergent; see Proposition 4.6.1. The factor 2π in (4.7.2) is necessary to obtain the standard convolution in the case of functions; cf. (4.7.3) below.

PROPERTIES 4.7.3. The distribution δ_Γ is the unit element relative to convolution in $\mathcal{D}'(\Gamma)$:

$$\delta_\Gamma * T = 2\pi \sum c_n[\delta_\Gamma]c_n[T]e^{inx} = \sum c_n[T]e^{inx} = T.$$

For the derivative of a convolution one has

$$D(S * T) = 2\pi \sum inc_n[S]c_n[T]e^{inx} = DS * T = S * DT.$$

LEMMA 4.7.4. *If g is a test function, $T * g$ is also a test function, and*

$$(T * g)(x) = < T(y), g(x - y) > .$$

Indeed, by (4.7.2), for $n \neq 0$,

$$\begin{aligned} c_n[T * g] &= \mathcal{O}(|n|^\beta |n|^{-p}) \quad \text{(for some } \beta \text{ and all } p \in \mathbb{N}) \\ &= \mathcal{O}(|n|^{-q}) \quad \text{for all } q \in \mathbb{N}. \end{aligned}$$

Hence $T * g$ is equal to a $\mathcal{C}^\infty$ function; cf. Exercise 4.2.1. Furthermore, by the continuity of T,

$$\begin{aligned} (T * g)(x) &= \sum \; < T(y), e^{-iny} > \; c_n[g]e^{inx} \\ &= \left\langle T(y), \sum c_n[g]e^{in(x-y)} \right\rangle = < T(y), g(x - y) > . \end{aligned}$$

When T is equal to an integrable function f we find

$$\begin{aligned} (f * g)(x) &= < f(y), g(x - y) > = \int_\Gamma f(y)g(x - y)dy \\ &= \int_\Gamma f(x - y)g(y)dy. \end{aligned} \tag{4.7.3}$$

This formula will make sense for any two integrable functions f and g.

PROPOSITION 4.7.5. *For integrable functions f and g on Γ, the convolution integral* (4.7.3) *exists for almost all $x \in \Gamma$. It defines an integrable function h on Γ, and*

$$\int_\Gamma h = \int_\Gamma f \int_\Gamma g, \quad \int_\Gamma h(x)e^{-inx}dx = \int_\Gamma f(y)e^{-iny}dy \int_\Gamma g(z)e^{-inz}dz,$$

in accordance with (4.7.2).

*PROOF. By Fubini's theorem [of Integration Theory] for positive functions, one has the following equalities for repeated integrals involving absolute values:

$$\begin{aligned}\int_\Gamma dy \int_\Gamma |f(y)g(x-y)|dx &= \int_\Gamma |f(y)|dy \int_\Gamma |g(x-y)|dx \\ &= \int_\Gamma |f(y)|dy \int_\Gamma |g(z)|dz.\end{aligned}$$

The finiteness of the product on the right implies that the double integral of $f(y)g(x-y)$ over $\Gamma \times \Gamma$ exists. Still by Fubini, it follows that the repeated integral

$$\int_\Gamma dx \int_\Gamma f(y)g(x-y)dy$$

exists, in the sense that the inner integral exists for almost all x, and that it thereby defines an integrable function $h(x)$. Furthermore, the *order* of integration in the final integral may be *inverted*:

$$\begin{aligned}\int_\Gamma h(x)dx &= \int_\Gamma dx \int_\Gamma f(y)g(x-y)dy = \int_\Gamma f(y)dy \int_\Gamma g(x-y)dx \\ &= \int_\Gamma f(y)dy \int_\Gamma g(z)dz.\end{aligned}$$

The second formula in the Proposition may be proved in the same way. □

We return now to the general case $S*T$ of Definition 4.7.2 and compute the action of $S*T$ on a test function ϕ, using Proposition 4.4.3:

$$\begin{aligned}< S*T, \phi > &= (2\pi)^2 \sum c_n[S]c_n[T]c_{-n}[\phi] = (2\pi)^2 \sum c_n[S]c_{-n}[T_R]c_{-n}[\phi] \\ &= 2\pi \sum c_n[S]c_{-n}[T_R * \phi] \ = < S, T_R * \phi > .\end{aligned}$$

Indeed, by Lemma 4.7.4, $T_R * \phi$ is a test function.

Exercises. 4.7.1. Solve the distributional equation $(e^{ix}-1)T = 0$ on Γ.

4.7.2. Let T be a distribution on Γ of order m, considered as a continuous linear functional on the space $\mathcal{D}^m(\Gamma)$ [that is, $\mathcal{C}^m(\Gamma)$ supplied with an appropriate concept of convergence; cf. Exercise 4.6.8.] Show that the rule

$$< Tg, \phi > = < T, g\phi >, \quad \forall\, \phi \in \mathcal{C}^m(\Gamma)$$

defines Tg as another distribution of order m.

4.7.3. Compute $\delta_\Gamma g$ for all continuous functions g, and $D\delta_\Gamma \cdot g$ for all $\mathcal{C}^1$ functions g.

4.7.4. Show that for integrable functions f and $\mathcal{C}^1$ functions g on Γ,

$$Df \cdot g = D(fg) - fg'.$$

4.7.5. Write $S * T$ in the standard form $d_0 + D^s f$ with integrable f if $S = a_0 + D^p g$ and $T = b_0 + D^q h$, where g and h are integrable functions on Γ with average zero.

4.7.6. Prove that distributional convergence $S_k \to S$ and $T_k \to T$ implies that $S_k * T_k \to S * T$.

4.7.7. Show that for delta families $\{f_\lambda\}$, $\lambda \to \lambda_0$ on Γ [cf. Examples 4.3.2], one has

$$f_\lambda * T \to \delta_\Gamma * T = T, \quad \forall\, T \in \mathcal{D}'(\Gamma).$$

4.7.8. Prove the partial sum formula $s_k[T] = D_k * T$, where D_k is the Dirichlet kernel. Use it to show [once again] that $s_k[T] \to T$, $\forall\, T \in \mathcal{D}'(\Gamma)$.

4.7.9. Let f be a distribution on Γ and $c_n = c_n[f]$. Describe the distributions which have the following Fourier coefficients:

$$\text{(i) } nc_n; \;\; \text{(ii) } c_n/n, \;\; (n \neq 0); \;\; \text{(iii) } c_{-n}; \;\; \text{(iv) } \overline{c_n}; \;\; \text{(v) } c_n^2; \;\; \text{(vi) } |c_n|^2.$$

4.7.10. Let f be continuous or piecewise continuous on $[-\pi, \pi]$, $c_n = c_n[f]$. Prove that the series $\sum |c_n|^2 e^{inx}$ is Cesàro summable and derive that $\sum |c_n|^2$ converges. Express the sum in terms of an integral. [In this exercise, square integrability of f would suffice.]

CHAPTER 5

Metric, normed and inner product spaces

In approximation problems, the degree of approximation is usually measured with the aid of a metric or distance concept. Many kinds of convergence of functions, such as uniform convergence and convergence in the mean, correspond to metrics on linear spaces of functions. In an arbitrary metric space the geometry may be so strange that it is of little help in solving problems. The situation is better in normed linear spaces, where every element or vector has a norm or length, and where the distance $d(u, v)$ is the length of $u - v$. The geometry is particularly nice – essentially Euclidean – in scalar product spaces, where for the vectors there are not only lengths, but also angles. Particularly useful for applications is the concept of orthogonality.

5.1. Metrics

Let X be an arbitrary set of elements which we call points.

DEFINITION 5.1.1. A function $d(u, v)$ defined for all points u, v in X is called a *distance function* or *metric* if

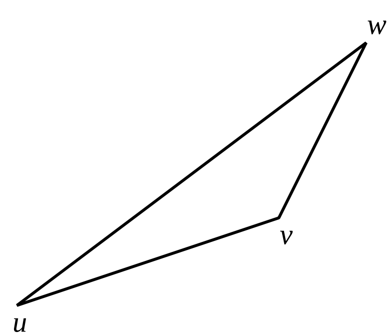

FIGURE 5.1

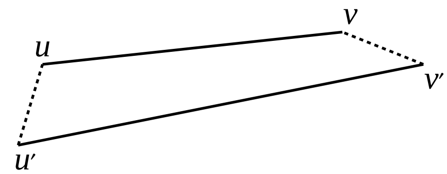

FIGURE 5.2

$$
\begin{aligned}
&\text{(i) } 0 \le d(u,v) < \infty, \quad \forall\, u,\, v\, \in X;\\
&\text{(ii) } d(u,v) = 0 \ \text{ if and only of } u = v;\\
&\text{(iii) } d(v,u) = d(u,v), \quad \forall\, u,\, v\, \in X;\\
&\text{(iv) } d(u,w) \le d(u,v) + d(v,w), \quad \forall\, u,\, v,\, w\, \in X\\
&\qquad (\textit{triangle inequality}; \text{ cf. Figure 5.1}).
\end{aligned}
$$

A set X with a distance function d is called a *metric space*, sometimes denoted by (X, d). In $X = (X, d)$ one defines *convergence* as follows:

$$u_k \to u \quad \text{or} \quad \lim u_k = u \iff d(u, u_k) \to 0 \ \text{ as } \ k \to \infty.$$

If a sequence $\{u_k\}$ converges to u, every subsequence also converges to u. A sequence in a metric space has at most one limit. It follows from the triangle inequality that

$$|d(u,v) - d(u',v')| \le d(u,u') + d(v,v');$$

cf. Figure 5.2. Thus the distance function $d(u,v)$ is continuous on $X \times X$.

A *subspace* Y of a metric space X is simply a subset, equipped with the metric provided by X.

EXAMPLES 5.1.2. Let $\mathbb{R}^n$ as usual be the real linear space of the vectors $x = (x_1, \cdots, x_n)$: ordered n-tuples of real numbers. Addition and multiplication by scalars λ (here real numbers) are carried out componentwise.

The metric space $\mathbb{E}^n$, *Euclidean* (coordinate) *space*, is obtained from $\mathbb{R}^n$ by imposing the Euclidean distance d_2:

$$d_2(x,y) = \left\{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2\right\}^{\frac{1}{2}}.$$

[Here one takes the nonnegative square root.] Thus one may write $\mathbb{E}^n = (\mathbb{R}^n, d_2)$. Many other distance functions are possible on $\mathbb{R}^n$, for example,

$$\begin{aligned} d_1(x,y) &= |x_1 - y_1| + \cdots + |x_n - y_n|, \\ d_\infty(x,y) &= \max\{|x_1 - y_1|, \cdots, |x_n - y_n|\}, \\ \tilde{d}(x,y) &= \min\{d_2(x,y), 1\}. \end{aligned}$$

Analogous definitions may be used on $\mathbb{C}^n$, the complex linear space of the ordered n-tuples $z = (z_1, \cdots, z_n)$ of complex numbers, with the complex numbers as scalars. The distance function d_2:

$$d_2(z,w) = \left\{|z_1 - w_1|^2 + \cdots + |z_n - w_n|^2\right\}^{\frac{1}{2}}$$

now leads to *unitary space* $\mathbb{U}^n = (\mathbb{C}^n, d_2)$.

There are corresponding distance functions on linear spaces whose elements are infinite sequences of real or complex numbers; cf. Examples 5.3.5. We first discuss the corresponding metrics on linear spaces of functions.

EXAMPLES 5.1.3. Let $[a,b]$ be a finite closed interval, $\mathcal{C}[a,b]$ the (complex) linear space of the continuous functions on $[a,b]$. Here the sum $f+g$ and the scalar multiple λf are defined in the usual way. If we only consider real-valued functions and real scalars, we will write $\mathcal{C}_{\mathrm{re}}[a,b]$.

Uniform convergence $f_k \to f$ on $[a,b]$ can be derived from the metric

$$d_\infty(f,g) = \sup_{a \le x \le b} |f(x) - g(x)|.$$

The metric space $(\mathcal{C}[a,b], d_\infty)$ will be denoted by $C[a,b]$. Other standard distance functions on $\mathcal{C}[a,b]$ are

$$\begin{aligned} d_1(f,g) &= \int_a^b |f(x) - g(x)| dx, \\ d_2(f,g) &= \left(\int_a^b |f(x) - g(x)|^2 dx\right)^{\frac{1}{2}}. \end{aligned}$$

In Section 5.6, the triangle inequality for d_2 will be derived from the general 'Cauchy–Schwarz inequality'.

Analogous definitions may be used on $\mathcal{C}(K)$, the (complex) linear space of the continuous functions on an arbitrary bounded closed set K in $\mathbb{E}^n$.

EXAMPLE 5.1.4. In formulating concepts and theorems, it is useful to keep in mind the somewhat pathological *discrete metric*. For any set X it is defined as follows:

$$d(u,v) = 1, \;\; \forall\, u,\, v \in X \;\text{ with } u \neq v; \;\; d(u.u) = 0, \;\; \forall\, u \in X.$$

5.2. Metric spaces: general results

In a metric space X one introduces the (open) *ball* $B(a,r)$ as the subset $\{u \in X : d(a,u) < r\}$. For a subset $E \subset X$ one may next define interior points x_0 (E contains a ball $B(x_0,\rho)$ of X) and the interior E^0. A point $c \in X$ is called a limit point of (or for) E if every ball $B(c,r)$ in X contains infinitely many points of E. The closure $\overline{E} = \operatorname{clos} E$ consists of E together with its limit points. A point $b \in X$ is called a boundary point of (or for) E if every ball $B(b,r)$ in X contains a point of E and a point not in E. The boundary ∂E is given by $\overline{E} \setminus E^0$. The boundary of $B(a,r)$ will be the *sphere* $S(a,r)$. A set $E \subset X$ may be open ($E = E^0$), closed ($E = \overline{E}$), or neither.

$E \subset X$ is called *dense* in X if $\overline{E} = X$; in this case, every point of X is the limit of a sequence of elements of E.

The subspace $\mathbb{E}^2_{\mathrm{rat}}$ of $\mathbb{E}^2$, consisting of the points with rational coordinates, is dense in $\mathbb{E}^2$. The set of all trigonometric polynomials is dense in $C(\Gamma)$. The set of all polynomials in x, restricted to the finite closed interval $[a,b]$, is dense in the space $C[a,b]$; cf. Section 3.4.

A metric space with a countable dense subset is called *separable*.

$E \subset X$ is called *bounded* if E is contained in a ball $B(a,r) \subset X$.

In $\mathbb{E}^n$, every bounded infinite set has a limit point, every bounded infinite sequence, a convergent subsequence. However, most metric spaces do not have these properties. Just think of $\mathbb{R}^2$ with the discrete metric, of $\mathbb{E}^2_{\mathrm{rat}}$, or of $(\mathbb{R}^2, \tilde{d})$ as in Examples 5.1.2. Other examples are $C[0,1]$, cf. Exercise 5.2.6, and the sequence $\{e^{int}\}$ in $(\mathcal{C}(\Gamma), d_2)$.

$E \subset X$ is called *compact* if every infinite sequence in E has a convergent subsequence with limit in E, or equivalently, if every covering of E by open subsets contains a finite subcovering. Compact sets are bounded and closed; in $\mathbb{E}^n$, the converse in also true.

Let T be a map from a metric space X to a metric space Y. One says that T is *continuous* at $u \in X$ if for every sequence $\{u_k\}$ in X with limit u, one has $Tu_k \to Tu$ in Y. The map T is called continuous on $E \subset X$ if it is continuous at every point of E. In the special case where $d(Tu,Tv) = d(u,v)$ for all $u,\, v \in X$, the map is called an *isometry*. The spaces X and $Y = TX$ are then called *isometric*. The complex plane (relative to ordinary distance) is isometric with $\mathbb{E}^2$.

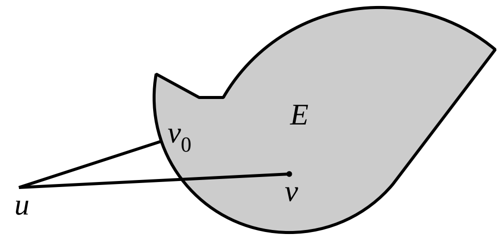

FIGURE 5.3

THEOREM 5.2.1. *Let X and Y be metric spaces, $E \subset X$ compact, $T : E \mapsto Y$ continuous. Then the image $TE \subset Y$ is also compact. In particular, a continuous real valued function on a* (nonempty) *compact set $E \subset X$ is bounded, and assumes a maximum and a minimum value on E.*

PROOF. For compact E and continuous T, any sequence $\{Tu_k\}$ with $\{u_k\} \subset E$ will have a convergent subsequence with limit in TE. Indeed, let $\{u_{n_k}\}$ be any subsequence of $\{u_k\}$ with a limit $u \in E$. Then $Tu_{n_k} \to Tu$. As to the second part, any (nonempty) bounded closed subset of $\mathbb{E}^1$, the real numbers under ordinary distance, has a largest and a smallest element. □

APPLICATION 5.2.2. For u in X and a nonempty compact subset E of X, the distance

$$d(u, E) \stackrel{\text{def}}{=} \inf_{v \in E} d(u, v) \tag{5.2.1}$$

is attained for some element v_0 in E: $d(u, E) = d(u, v_0)$.

Cf. Figure 5.3. One may say that there is an element v_0 in E that provides an optimal approximation to u.

Completeness. We will now discuss the important concept of completeness. Let $X = (X, d)$ be a metric space. A sequence $\{x_k\}$ in X is called a *Cauchy sequence* or fundamental sequence if

$$d(x_j, x_k) \to 0 \quad \text{as } j,\ k \to \infty. \tag{5.2.2}$$

Every convergent sequence is a Cauchy sequence, but in many metric spaces X there are Cauchy sequences which do not converge to a point of X. Such spaces are called incomplete.

DEFINITION 5.2.3. A metric space X is called *complete* if all Cauchy sequences in X converge to a point of X.

Every Cauchy sequence is bounded: if $d(x_j, x_k) < \varepsilon$ for all $j,\ k \geq p$, then $x_k \in B(x_p, \varepsilon)$ for all $k \geq p$. Similarly, a Cauchy sequence for which there is a limit point $c \in X$ must converge to c. Thus a metric space in which every bounded sequence has a limit point is complete.

EXAMPLES 5.2.4. The Euclidean plane $\mathbb{E}^2$ and the complex plane $(\mathbb{C}, d_2)$ are complete. The subspace $\mathbb{E}^2_{\text{rat}}$ of $\mathbb{E}^2$ is incomplete: the Cauchy sequence

$$x_1 = (1, 1.4),\ x_2 = (1, 1.41),\ \cdots,\ x_k = (1, r_k),\ \cdots,$$

where r_k is the largest decimal number $1.d_1 \cdots d_k$ with $r_k^2 < 2$, does not converge to a rational point. The metric spaces $(\mathbb{R}^n, d_2) = \mathbb{E}^n$, $(\mathbb{R}^n, d_1)$ and $(\mathbb{R}^n, d_\infty)$ [Examples 5.1.2] are complete. Indeed, every Cauchy sequence in these spaces is componentwise convergent, and hence convergent, to an element of the space.

The space $C[a, b]$ is *complete*. Indeed, let $\{f_k\}$ be an arbitrary Cauchy sequence in $C[a, b]$. Then for every point $x_0 \in [a, b]$, the sequence of complex numbers $\{f_k(x_0)\}$ is a Cauchy sequence, and hence convergent. Let f be the pointwise limit function of the sequence $\{f_k\}$, that is, $f(x) = \lim f_k(x)$ for every $x \in [a, b]$. For given $\varepsilon > 0$, we now take p so large that

$$d_\infty(f_j, f_k) = \max_{a \leq x \leq b} |f_j(x) - f_k(x)| < \varepsilon, \ \ \forall\, j,\ k \geq p.$$

Letting $j \to \infty$, this inequality implies that for all $x \in [a, b]$,

$$|f(x) - f_k(x)| \leq \varepsilon, \ \ \forall\, k \geq p.$$

Since $\varepsilon > 0$ was arbitrary, the conclusion is that $d_\infty(f, f_k) \to 0$ as $k \to \infty$: the sequence $\{f_k\}$ converges uniformly to f. It follows that f is continuous, hence $f \in C[a, b]$.

EXAMPLE 5.2.5. The space $X = (\mathcal{C}[a, b], d_1)$, with d_1 as in Examples 5.1.3, is *incomplete*. For a proof we take $a = -1$, $b = 1$, and define

$$f_k(x) = \begin{cases} 0 & \text{for } x \leq 0, \\ kx & \text{for } 0 \leq x \leq 1/k, \\ 1 & \text{for } x \geq 1/k. \end{cases}$$

The sequence $\{f_k\}$ is a Cauchy sequence in X: for $j,\ k \geq p$,

$$d_1(f_j, f_k) = \int_{-1}^{1} |f_j - f_k| \leq \int_0^{1/p} |f_j - f_k| \leq 1/p;$$

cf. Figure 5.4. The sequence $\{f_k\}$ converges at every point x; the pointwise limit function $U(x)$ is equal to 0 for $x \leq 0$ and equal to 1 for $x > 0$. Also,

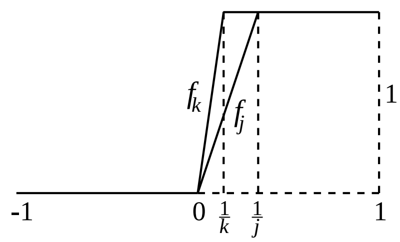

FIGURE 5.4

$\int_{-1}^{1} |U - f_k| \to 0$ as $k \to \infty$. However, there can be no continuous function f such that $d_1(f, f_k) \to 0$ as $k \to \infty$.

Indeed, suppose that there would be such an f. Then

$$\int_{-1}^{1} |f - U| = \int_{-1}^{1} |f - f_k + f_k - U|$$
$$\leq \int_{-1}^{1} |f - f_k| + \int_{-1}^{1} |f_k - U| \to 0 \quad \text{as} \;\; k \to \infty.$$

Hence the (constant) left-hand side would be equal to 0. But then the nonnegative function $|f(x) - U(x)|$ would have to be zero at every point where it is continuous, that is, for all $x \neq 0$. Thus one would have $f(x) = 0$ for $x < 0$, and $f(x) = 1$ for $x > 0$. But this would contradict the postulated continuity of f.

Every *incomplete* metric space can be *completed* by a standard abstract construction (see below).

DEFINITION 5.2.6. A metric space $\hat{X} = (\hat{X}, \hat{d})$ is called a *completion* of the metric space $X = (X, d)$ if $\hat{X}$ is complete, and X is [or can be considered as] a dense subspace of $\hat{X}$. That is, X lies dense in $\hat{X}$ and $\hat{d} = d$ on X.

Completions $\hat{X}$ of a given space X are unique up to isometry. In practice one can often indicate a concrete completion of a given space X.

Thus the space $\mathbb{E}^2_{\text{rat}}$ has as its completion the Euclidean plane $\mathbb{E}^2$. The space $X = (\mathcal{C}[a, b], d_1)$ has as its completion the space $L(a, b) = L^1(a, b)$ of the Lebesgue integrable functions on (a, b), where $\hat{d}_1(f, g) = \int_a^b |f - g|$. [It is understood that in $L^1(a, b)$ one identifies functions that agree outside a set of measure zero; cf. Remark 4.1.2.] The completeness of $L^1(a, b)$ follows from the Riesz–Fischer theorem of Integration Theory, which was named

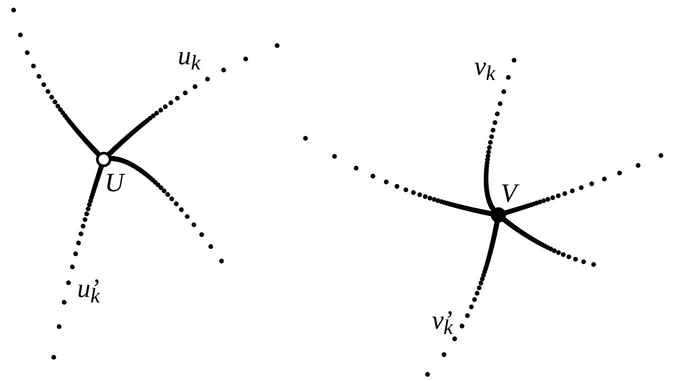

FIGURE 5.5

after F. Riesz (Section 4.2) and the Austrian mathematician Ernst Fischer (1875–1954; [**31**]). One form of the theorem says that for any sequence $\{f_k\}$ of integrable functions such that $\int_a^b |f_j - f_k| \to 0$, there is an integrable function f such that $\int_a^b |f - f_k| \to 0$; cf. [**102**], [**68**]. Furthermore, to given $f \in L^1(a,b)$ and any number $\varepsilon > 0$, there is a step function s such that $\hat{d}_1(f,s) < \varepsilon$. From this one may derive that there is a continuous function g on $[a,b]$ such that $\hat{d}_1(f,g) < 2\varepsilon$.

The general construction of completion. An abstract completion $\hat{X} = (\hat{X}, \hat{d})$ of a given (incomplete) metric space $X = (X, d)$ can be obtained as follows. Every nonconvergent Cauchy sequence $\{u_k\}$ in X identifies a "missing point", a "hole", in X. Think of $\mathbb{E}^2_{\text{rat}}$, where the "holes" are given by points that have at least one irrational coordinate. For completion of X, every hole has to be filled by a sort of "generalized limit" of the Cauchy sequence that defines it. Of course, different Cauchy sequences that "belong to the same hole" must be assigned the same generalized limit. Mathematically, a "hole" may be described as an equivalence class of (nonconvergent) Cauchy sequences. We will say that Cauchy sequences $\{u_k\}$ and $\{\tilde{u}_k\}$ are equivalent, notation $\{u_k\} \sim \{\tilde{u}_k\}$, if $d(u_k, \tilde{u}_k) \to 0$ as $k \to \infty$.

We now define a new metric space $\hat{X} = (\hat{X}, \hat{d})$ as follows. The elements $U, V, \cdots$ of $\hat{X}$ are equivalence classes of (nonconvergent or convergent) Cauchy sequences in X, cf. Figure 5.5, and

$$\hat{d}(U,V) = \lim d(u_k, v_k) \quad \text{if} \quad \{u_k\} \in U, \ \{v_k\} \in V. \tag{5.2.3}$$

It is easy to verify that the function $\hat{d}(U,V)$ is well-defined and that $(\hat{X},\hat{d})$ is a metric space.

Every element $u \in X$ is represented in $\hat{X}$ by the special equivalence class u^* of the Cauchy sequences in X that converge to u. An example of such a sequence is the "constant sequence" $\{u,\, u,\, u,\, \cdots\}$. Clearly

$$\hat{d}(u^*,v^*) = \lim\{d(u,v),\, d(u,v),\, \cdots\} = d(u,v).$$

The special elements u^*, v^*, $\cdots$ form a subspace of $\hat{X}$ isometric with X. Identifying u^* with u, v^* with v, etc, X becomes a subspace of $\hat{X}$. It is easy to see that X is dense in $\hat{X}$. Indeed, if $U \in \hat{X}$ and $\{u_k\}$ is one of its Cauchy sequences, then $u_k = u_k^*$ converges to U in $\hat{X}$. For if $d(u_j,u_k) < \varepsilon$ for all $j,\, k \geq p$, then by (5.2.3)

$$\begin{aligned}\hat{d}(U,u_k^*) &= \lim\{d(u_1,u_k),\, d(u_2,u_k),\, \cdots\}\\ &= \lim_{j\to\infty} d(u_j,u_k)) \leq \varepsilon, \quad \forall\, k \geq p.\end{aligned}$$

Finally, $\hat{X}$ will be complete. If $\{U_k\}$ is a Cauchy sequence in $\hat{X}$, then for each k we can choose an element $u_k \in X$ such that $\hat{d}(U_k,u_k) < 1/k$. By the triangle inequality, the sequence $\{u_k\}$ will be a Cauchy sequence in X. Its generalized limit U, located in $\hat{X}$, will also be the limit of the sequence $\{U_k\}$ in $\hat{X}$.

Remark 5.2.7. The space $\hat{X}$ contains two kinds of elements: equivalence classes of nonconvergent Cauchy sequences, and equivalence classes of convergent Cauchy sequences. The book Mathematical Methods vol. 1 [**68**] speaks figuratively of "spiders with a hole" and "spiders with a heart"; cf. Figure 5.5.

Exercises. 5.2.1. For an arbitrary subset E of a metric space X and for $u \in X$, the distance $d(u,E)$ is defined as in (5.2.1). Prove that

$$|d(u,E) - d(v,E)| \leq d(u,v), \quad \forall\, u,\, v \in X.$$

Thus the distance $d(u,E)$ is continuous on X.

5.2.2. Let $\mathbb{R}^{\mathbb{N}}$ be the linear space of all infinite sequences $x = (x_1,\, x_2,\, \cdots)$ of real numbers. Prove that the formula

$$d^*(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min\{|x_n - y_n|, 1\}$$

defines a metric on $\mathbb{R}^{\mathbb{N}}$. Show that convergence $x^{(k)} \to x$ in $(\mathbb{R}^{\mathbb{N}}, d^*)$ is exactly the same as "componentwise convergence": $x_n^{(k)} \to x_n$ for each $n \in \mathbb{N}$.

5.2.3. Show that convergence $\phi_k \to \phi$ in the test space $\mathcal{D}(\Gamma)$ [Definition 4.2.1] may be derived from a metric.

5.2.4. Prove that the piecewise constant functions (step functions) form a dense subspace of $C[a, b]$.

5.2.5. Show that the spaces $C[a, b]$ and $L^1(a, b)$ are separable.

5.2.6. Prove that the bounded sequence $f_k(x) = x^k$, $k = 1, 2, \cdots$ in $C[0, 1]$ does not have a uniformly convergent subsequence on $[0, 1]$.

5.2.7. Show that a closed subspace Y of a complete metric space X is complete.

5.3. Norms on linear spaces

In this section, V denotes a linear space or vector space. In analysis, the scalars are (almost) always the real or the complex numbers. Accordingly, we speak of real or complex linear spaces. We begin by reviewing some terminology concerning linear spaces.

A (linear) *subspace* W of V is a subset which is also a linear space under the given addition and multiplication by scalars in V. Examples in the case of $V = \mathcal{C}[a, b]$: the subspace $\mathcal{P}$ of the polynomials in x (restricted to $[a, b]$), the subspace $\mathcal{P}_n$ of the polynomials of degree $\leq n$.

For a subset A of V, the (linear) *span* $S(A)$ is the subspace "spanned" or "generated" by A. It consists of all finite linear combinations $\lambda_1 u_1 + \cdots + \lambda_k u_k$ of elements u_j in A. Thus in $\mathcal{C}[a, b]$, the subspace $\mathcal{P}$ is the span of the subset $\{1, x, x^2, \cdots\}$. A subset $A \subset V$ is called *linearly independent* if a finite linear combination $\lambda_1 u_1 + \cdots + \lambda_k u_k$ of elements of A is equal to zero only when $\lambda_1 = \cdots = \lambda_k = 0$. The set $\{1, x, x^2, \cdots\}$ is linearly independent in $\mathcal{C}[a, b]$.

A subset A of V is called a *basis* (or algebraic basis) for V if every element u in V can be represented in exactly one way as a finite linear combination $c_1 u_1 + \cdots + c_k u_k$ of elements $u_j \in A$. A basis is the same as a linearly independent spanning set for V. All algebraic bases of V have the same number of elements, or more precisely, the same "cardinal number". [In other words, one can set up a one-to-one correspondence between the elements of any two algebraic bases.] This cardinal number [which could, for example, be "countably infinite"] is called the (algebraic) *dimension* of V.

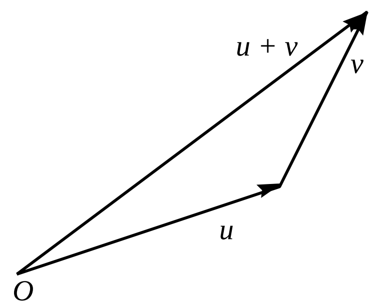

FIGURE 5.6

For subspaces W_1 and W_2 of V one can form the (vector) *sum* $W_1 + W_2$, that is, the subspace of all vectors $w = w_1 + w_2$ with $w_j \in W_j$. If $W_1 \cap W_2 = \{0\}$, the zero element, the representation $w = w_1 + w_2$ is unique. In this case one speaks of the *direct sum* of W_1 and W_2, notation $W_1 \oplus W_2$. If $V = W_1 \oplus W_2$ one calls W_1 and W_2 *complementary subspaces* of V. In this case one can define a codimension: $\operatorname{codim} W_1 = \dim W_2$. If $\dim V$ is finite, $\operatorname{codim} W_1 = \dim V - \dim W_1$.

DEFINITION 5.3.1. A function $\|\cdot\|$ on V is called a length or *norm* if

(i) $0 \le \|u\| < \infty, \quad \forall\, u \in V$;
(ii) $\|u\| = 0$ if and only if $u = 0$;
(iii) $\|\lambda u\| = |\lambda|\, \|u\|, \quad \forall\, u \in V, \forall\, \text{scalars}\, \lambda$;
(iv) $\|u + v\| \le \|u\| + \|v\|, \quad \forall\, u,\, v \in V$
(*triangle inequality*; cf. Figure 5.6).

From the norm one may derive a distance function d by setting

$$d(u, v) \stackrel{\text{def}}{=} \|u - v\|; \tag{5.3.1}$$

cf. Figure 5.7. As usual, the metric d implies a notion of convergence:

$$u_k \to u \quad \text{if and only if} \quad d(u, u_k) = \|u - u_k\| \to 0. \tag{5.3.2}$$

DEFINITION 5.3.2. A normed linear space $V = (V, \|\cdot\|)$ is a linear space V with a norm function $\|\cdot\|$ and the associated distance (5.3.1) and convergence (5.3.2).

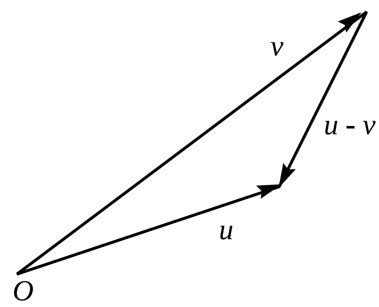

FIGURE 5.7

The norm function is continuous: if $u_k \to u$, then $\|u_k\| \to \|u\|$. By definition, a subspace W of a normed linear space V is a linear subspace, equipped with the norm provided by V.

EXAMPLES 5.3.3. On $\mathbb{R}$ and $\mathbb{C}$, the absolute value is a norm. On $\mathbb{R}^n$ and $\mathbb{C}^n$, with elements denoted by $x = (x_1, \cdots, x_n)$, one has the norms

$$\|x\|_\infty = \max_{1 \le \nu \le n} |x_\nu|; \quad (\mathbb{C}^n, \|\cdot\|_\infty) \text{ is also called } l^\infty(n);$$

$$\|x\|_1 = |x_1| + \cdots + |x_n|; \quad (\mathbb{C}^n, \|\cdot\|_1) \text{ is also called } l^1(n);$$

$$\|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}; \quad (\mathbb{C}^n, \|\cdot\|_2) \text{ is also called } l^2(n).$$

The corresponding distances are d_∞, d_1 and d_2 as in Examples 5.1.2. On $\mathbb{R}^n$, the third norm is the Euclidean norm or length; we also denote the normed space $(\mathbb{R}^n, d_2)$ by $\mathbb{E}^n$. Similarly $l^2(n)$ and unitary space $\mathbb{U}^n$ are identified. Convergence under all these norms is the same as componentwise convergence.

EXAMPLES 5.3.4. On $\mathcal{C}[a,b]$ (where $[a,b]$ denotes a bounded closed interval), the formula

$$\|f\|_\infty = \sup_{a \le x \le b} |f(x)| = \max_{a \le x \le b} |f(x)| \tag{5.3.3}$$

defines a norm, usually referred to as the *supremum norm*. The corresponding distance is

$$d_\infty(f,g) = \max_{a \le x \le b} |f(x) - g(x)|.$$

The corresponding convergence is uniform convergence; cf. Examples 5.1.3. From now on we will use the notation $C[a,b]$ for the *normed* linear space

$(\mathcal{C}[a,b], \|\cdot\|_\infty)$. If we restrict ourselves to real functions and scalars we may write $C_{\mathrm{re}}[a,b]$.

Another important norm on $\mathcal{C}[a,b]$ is

$$\|f\|_1 = \int_a^b |f(x)|dx. \tag{5.3.4}$$

This formula makes sense for all (Lebesgue) integrable functions f on (a,b), even if the interval (a,b) is unbounded. It defines a norm on the space $\mathcal{L}^1(a,b)$, provided we identify functions that differ only on a set of (Lebesgue) measure zero. The corresponding distance is the L^1-distance,

$$d_1(f,g) = \int_a^b |f(x) - g(x)|dx.$$

From now on we will use the notation $L^1(a,b)$ for the *normed* linear space of the integrable functions on (a,b) with the norm (5.3.4) [identifying almost equal functions]. For bounded intervals (a,b), L^1-convergence is the same as "convergence in the mean" on (a,b). Indeed, $f_k \to f$ in $L^1(a,b)$ if and only if

$$\frac{1}{b-a}\int_a^b |f(x) - f_k(x)|dx \to 0;$$

the mean or average deviation $|f(x) - f_k(x)|$ must tend to zero. Note that convergence in $C[a,b]$ means that the maximum deviation tends to zero.

The distance $d_2(f,g)$ on $\mathcal{C}[a,b]$ (Examples 5.1.3) may be derived from the norm

$$\|f\|_2 = \left\{\int_a^b |f(x)|^2 dx\right\}^{1/2}; \tag{5.3.5}$$

see Section 5.6 for a proof of the triangle inequality.

The definitions in Examples 5.3.3 cannot be extended to *all infinite* sequences $x = (x_1, x_2, \cdots)$ of complex numbers. One has to impose appropriate restrictions:

EXAMPLES 5.3.5. For the bounded infinite sequences $x = (x_1, x_2, \cdots)$, the definition

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n| \quad \text{gives the space } \ l^\infty = l^\infty(\mathbb{N});$$

for the infinite sequences $x = (x_1, x_2, \cdots)$ such that the series $\sum_{n=1}^{\infty} |x_n|$ converges, the definition

$$\|x\|_1 = \sum_{n=1}^{\infty} |x_n| \quad \text{gives the space } \; l^1 = l^1(\mathbb{N});$$

for the infinite sequences $x = (x_1, x_2, \cdots)$ such that the series $\sum_{n=1}^{\infty} |x_n|^2$ converges, the definition

$$\|x\|_2 = \left\{ \sum_{n=1}^{\infty} |x_n|^2 \right\}^{1/2} \qquad \text{gives the space } \; l^2 = l^2(\mathbb{N}).$$

For a proof of the triangle inequality in the case of "little el two", see Section 5.6.

Exercises. 5.3.1. Prove that the set of powers $\{1, x, x^2, \cdots\}$ (restricted to the interval $[0,1]$) is linearly independent in the space $\mathcal{C}[0,1]$. Can you also prove that every set of pairwise different exponential functions $\{e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots\}$ is linearly independent in $\mathcal{C}[0,1]$?

5.3.2. In $V = \mathcal{C}[-1,1]$, let W_1 and W_2 be the linear subspaces of the odd, and the even, continuous functions, respectively. Prove that V can be written as the direct sum $W_1 \oplus W_2$.

5.3.3. Let V be a linear space with a norm function $\|\cdot\|$. Verify that the formula $d(u,v) = \|u - v\|$ defines a metric on V. Show that $\|u - v\| \geq \|u\| - \|v\|$, and deduce that the norm function is continuous.

5.3.4. Use the ordinary coordinate plane to draw pictures of the "unit sphere" $S(0,1)$ [set of vectors of length 1] in each of the (real) spaces $l^{\infty}_{\mathrm{re}}(2)$, $l^1_{\mathrm{re}}(2)$, and $l^2_{\mathrm{re}}(2) = \mathbb{E}^2$.

5.3.5. Prove that the unit ball $B(0,1)$ [set of vectors of length < 1] in a normed linear space is always *convex*. [Given u_0, $u_1 \in B(0,1)$, show that $u_\lambda = (1-\lambda)u_0 + \lambda u_1$ is in $B(0,1)$ for every $\lambda \in (0,1)$.]

5.3.6. Verify that $\|f\|_\infty$ and $\|f\|_1$ as in (5.3.3), (5.3.4) are norms on $\mathcal{C}[a,b]$. Prove an inequality between these norms. Deduce that uniform convergence (on a bounded interval) implies L^1-convergence.

5.3.7. Consider the sequence of functions $f_k(x) = k^\alpha x^k$, $k = 1, 2, \cdots$, in $\mathcal{C}[0,1]$. For which real numbers α will the sequence converge to the zero function under the norms (i) $\|\cdot\|_\infty$; (ii) $\|\cdot\|_1$; (iii) $\|\cdot\|_2$?

5.3.8. Which of the following formulas define a norm on $\mathcal{C}^1[a,b]$:

$$(i)\ \|f\| = \max_{a\le x\le b} |f(x)|;\quad (ii)\ \|f\| = \max|f(x)| + \max|f'(x)|;$$

$$(iii)\ \|f\| = \max|f(x)+f'(x)|;\quad (iv)\ \|f\| = |f(a)| + \int_a^b |f'(x)|dx.$$

5.3.9. Consider the infinite sequences $x = (x_1, x_2, \cdots)$ of complex numbers for which the series $\sum_{n=1}^{\infty} |x_n|$ converges. Verify that they form a linear space V, and that the formula $\|x\|_1 = \sum_{n=1}^{\infty} |x_n|$ defines a norm on V.

5.3.10. Construct a sequence $\{f_k\}$ of piecewise constant functions on $[0,1]$ which converges to zero in the mean $[\int_0^1 |f_k - 0| \to 0]$, but which fails to converge to zero at *every* point of $[0,1]$ [for every $x \in [0,1]$, $f_k(x) \not\to 0$].

5.3.11. For $x \in \mathbb{C}^n$ and any $p \ge 1$ one may define

$$\|x\|_p = \left\{|x_1|^p + \cdots + |x_n|^p\right\}^{1/p}.$$

Prove the following relation which explains the notation $\|x\|_\infty$ for the supremum norm:

$$\lim_{p\to\infty} \|x\|_p = \max_{1\le\nu\le n} |x_\nu| = \|x\|_\infty.$$

5.4. Normed linear spaces: general results

In this section, V will denote a real or complex normed linear space. We begin by considering

Finite dimensional V. Setting $\dim V = n$, we choose a basis

$$B = \{u_1,\, u_2,\, \cdots,\, u_n\} \tag{5.4.1}$$

for V, so that every element of $u \in V$ has a unique representation

$$u = c_1u_1 + \cdots + c_nu_n, \quad \text{with}\ \ c_j = c_j(u). \tag{5.4.2}$$

We will compare the norm of u with the norm of $c = (c_1, \cdots, c_n) = c(u)$ as an element of $\mathbb{E}^n$ (if V is a real space) or $\mathbb{U}^n$ (if V is a complex space).

LEMMA 5.4.1. *There are positive constants m_B and M_B* (depending on V) *such that*

$$\begin{aligned} m_B\big\{|c_1|^2 + \cdots + |c_n|^2\big\}^{1/2} &\le \|c_1u_1 + \cdots + c_nu_n\|_V \\ &\le M_B\big\{|c_1|^2 + \cdots + |c_n|^2\big\}^{1/2}, \quad \forall\, c = (c_1, \cdots, c_n). \end{aligned} \tag{5.4.3}$$

PROOF. It is enough to consider the real case, the complex case being similar. If $c = 0$ the inequalities (5.4.3) are satisfied no matter what constants m_B and M_B we use. Supposing $c \neq 0$, we may take $\|c\|_2 = 1$ [by homogeneity, one may in (5.4.3) replace c by λc]. Thus we may assume that $c \in S(0,1)$, the unit sphere in $\mathbb{E}^n$. Introducing the function

$$f(c) = \|c_1 u_1 + \cdots + c_n u_n\|_V, \quad c \in S(0,1) \subset \mathbb{E}^n,$$

we have to show that f has a positive lower bound m_B and a finite upper bound M_B. For this we need two facts:

(i) f is continuous. Indeed, if $c' \to c$ in $\mathbb{E}^n$, then $c'_\nu \to c_\nu$ for $\nu = 1, \cdots, n$. Hence $c'_1 u_1 + \cdots + c'_n u_n \to c_1 u_1 + \cdots + c_n u_n$ in V (by the triangle inequality), and thus $f(c') \to f(c)$ by the continuity of the norm function.

(ii) $S(0,1)$ is compact: it is a bounded closed set in $\mathbb{E}^n$.

Conclusion: f assumes a minimum value m_B and a maximum value M_B on $S(0,1)$; cf. Theorem 5.2.1. Since $0 < f(c) < \infty$ for every $c \in S(0,1)$, we have $0 < m_B \le M_B < \infty$. □

THEOREM 5.4.2. *Let V be a finite dimensional* (real or complex) *normed linear space. Then*

(*i*) *Every bounded sequence in V has a convergent subsequence*;

(*ii*) *Every bounded closed set $E \subset V$ is compact*;

(*iii*) *V is complete.*

PROOF. We set $\dim V = n$ and choose a basis $\{u_1, \cdots, u_n\}$.

(i) Let $u^{(k)} = c_1^{(k)} u_1 + \cdots + c_n^{(k)} u_n$, $k = 1, 2, \cdots$, be a bounded sequence in V. Then by Lemma 5.4.1, the sequence $\{c^{(k)}\}$ in $\mathbb{E}^n$ is bounded, hence the coefficient sequences $\{c_1^{(k)}\}, \cdots, \{c_n^{(k)}\}$ are bounded. Taking a suitable subsequence $\{k_p\}$ of the sequence of positive integers $\{k\}$, we obtain convergent coefficient (sub)sequences. Denoting the limits by $c_1, \cdots, c_n$, respectively, we conclude that for $k = k_p \to \infty$,

$$u^{(k)} = c_1^{(k)} u_1 + \cdots + c_n^{(k)} u_n \to u = c_1 u_1 + \cdots + c_n u_n.$$

(ii) This is clear from the definition of compactness.

(iii) Every Cauchy sequence is bounded; cf. Section 5.2. Thus by part (i), every Cauchy sequence $\{u^{(k)}\}$ in V has a convergent subsequence. Calling its limit u, the whole Cauchy sequence $\{u^{(k)}\}$ will converge to u; cf. Section 5.2. □

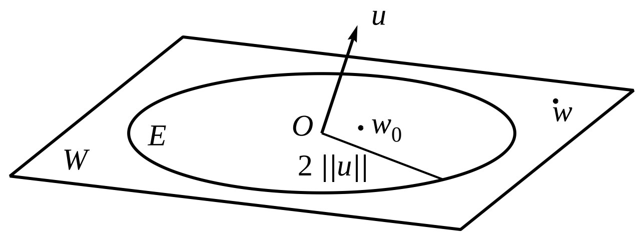

FIGURE 5.8

APPLICATION 5.4.3. **Existence of optimal approximations**. Let V be an arbitrary (real or complex) normed linear space, W a finite-dimensional subspace. Let u be an arbitrary given element of V. Then among the elements of W there is an element w_0 which provides an optimal approximation to u:

$$d(u, w_0) = d(u, W) \stackrel{\text{def}}{=} \inf_{w \in W} d(u, w).$$

PROOF. In looking for an optimal approximation to u, we may restrict ourselves to vectors w in the closed ball $E = \overline{B}(0, 2\|u\|)$ in W (Figure 5.8). Indeed, if $\|w\| > 2\|u\|$, then $0 \in W$ is a better approximation to u than w would be:

$$d(u, w) = \|w - u\| \geq \|w\| - \|u\| > \|u\| = d(u, 0).$$

Now by Theorem 5.4.2, the closed ball E in W is compact, hence by Application 5.2.2, there is an element $w_0 \in E$ which provides an optimal approximation to u. □

Let us consider the special case $V = C[a, b]$ and $W = \mathcal{P}_n$, the subspace of the polynomials in x of degree $\leq n$ (restricted to the interval $[a, b]$). Here we obtain the following

COROLLARY 5.4.4. *For every function $f \in C[a, b]$ and every n, there is a polynomial p_0 of degree $\leq n$ which provides an optimal approximation to f from the class $\mathcal{P}_n$:*

$$\|f - p_0\|_\infty = \min_{p \in \mathcal{P}_n} \|f - p\|.$$

Banach spaces. A *complete* normed linear space is called a Banach space, after the Polish mathematician Stefan Banach (1892–1945; [**5**]). Examples are: the finite dimensional (real or complex) normed linear spaces [Theorem 5.4.2]; the space $C[a,b]$ [cf. Examples 5.3.4, 5.2.4]; the space $L^1(a,b)$ [Examples 5.3.4 and Section 5.2]; the spaces l^∞ and l^1 [cf. Exercise 5.4.9].

In a finite dimensional normed linear space, the unit sphere $S(0,1)$ is compact, but in an infinite dimensional normed linear space, it never is. For example, in $C[0,1]$, the sequence of unit vectors $f_k(x) = x^k$, $k = 1, 2, \cdots$, fails to have a convergent subsequence; cf. Exercise 5.2.6. In l^∞, the sequence of unit vectors

$$e_1 = (1,0,0,0,\cdots),\ e_2 = (0,1,0,0,\cdots),\ e_3 = (0,0,1,0,\cdots),\cdots \tag{5.4.4}$$

fails to have a convergent subsequence. Indeed, $\|e_j - e_k\|_\infty = 1$ whenever $j \neq k$. For the general case, cf. Exercise 5.4.10.

In a Banach space, every "norm convergent' series converges to an element of the space:

THEOREM 5.4.5. *Let V be a complete normed linear space, $u_1 + u_2 + \cdots$ an infinite series in V such that*

$$\sum_{n=1}^{\infty} \|u_n\| \text{ converges.} \tag{5.4.5}$$

Then the series $\sum_{n=1}^{\infty} u_n$ converges to an element s in V.

PROOF. Writing $u_1 + \cdots + u_k = s_k$, it will follow from (5.4.5) that $\{s_k\}$ is a Cauchy sequence in V. Indeed, taking $k > j$ as we may,

$$\|s_k - s_j\| = \|u_{j+1} + \cdots + u_k\| \leq \|u_{j+1}\| + \cdots + \|u_k\|.$$

For any given $\varepsilon > 0$, there will be an index k_0 such that the final sum is $< \varepsilon$ whenever $j,\ k > k_0$.

Since V is complete, the Cauchy sequence $\{s_k\}$ has a limit s in V. □

EXAMPLES 5.4.6. (i) $V = \mathbb{R}$ and $V = \mathbb{C}$, with the absolute value of a number as norm. Every absolutely convergent series of real or complex numbers is convergent.

(ii) $V = C[a,b]$. Every infinite series $\sum_{n=1}^{\infty} g_n$, consisting of continuous functions on the finite closed interval $[a,b]$, and such that the series

$$\sum_{n=1}^{\infty} \|g_n\|_\infty = \sum_{n=1}^{\infty} \max_{a \leq x \leq b} |g_n(x)| \quad \text{converges,}$$

is uniformly convergent on $[a,b]$. This is essentially *Weierstrass's test* for uniform convergence which says the following. If there are numbers M_n

such that

$$|g_n(x)| \leq M_n \text{ on } [a,b], \text{ while } \sum_{n=1}^{\infty} M_n \text{ converges,}$$

then the series $\sum_{n=1}^{\infty} g_n(x)$ converges uniformly on $[a,b]$.

(iii) $V = L^1(a,b)$. Every series $\sum_{n=1}^{\infty} g_n$ of Lebesgue integrable functions on (a,b), such that the series

$$\sum_{n=1}^{\infty} \|g_n\|_1 = \sum_{n=1}^{\infty} \int_a^b |g_n(x)|dx \text{ converges,} \tag{5.4.6}$$

will be convergent on (a,b) in L^1-sense. That is, the partial sums $f_k = g_1 + \cdots + g_k$ will converge to an integrable function f on (a,b) in the sense that $\int_a^b |f - f_k| \to 0$. It will follow that

$$\int_a^b f = \lim \int_a^b f_k = \lim \left\{ \int_a^b g_1 + \cdots + \int_a^b g_k \right\}.$$

In other words,

$$\int_a^b f = \int_a^b \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty} \int_a^b g_n.$$

In Integration Theory it is shown that under condition (5.4.6), the series $\sum_{n=1}^{\infty} g_n$ is pointwise convergent on (a,b) outside a set of measure zero (which may be empty). Denoting the pointwise sum function by f, one also has $f_k \to f$ in L^1-sense, so that the series for f may be integrated term by term. The result is sometimes called Levi's theorem, after Beppo Levi (Italy, 1875–1961; [**80**]).

Normed space bases. Let V be an infinite dimensional normed linear space. A sequence $\{u_n\}$ in V is called a normed space basis or *Schauder basis* for V, after the Polish mathematician Juliusz Schauder (1899–1943; [**105**]), if every element u in V has a unique representation as the sum of a series $\sum_{n=1}^{\infty} c_n u_n$. For example, the sequence of unit vectors (5.4.4) is a normed space basis for l^1. The standard (separable) normed linear spaces all possess a Schauder basis; cf. [**106**].

Exercises. 5.4.1. Let $\{u_1, \cdots, u_n\}$ be a basis for the normed linear space V. Prove that elements $u^{(k)} = c_1^{(k)} u_1 + \cdots + c_n^{(k)} u_n$ converge to $u = c_1 u_1 + \cdots + c_n u_n$ in V if and only if they converge componentwise, that is, $c_\nu^{(k)} \to c_\nu$ for every ν.

5.4.2. Prove that every finite dimensional subspace W of a normed linear space V is closed.

5.4.3. Determine the constant (real) functions which optimally approximate $f(x) = x^2$ on $[0, 1]$ relative to the norms $\|\cdot\|_\infty$ and $\|\cdot\|_1$.

5.4.4. Same question for the unit step function $U(x)$ on $[-1, 1]$. Are the optimally approximating functions unique in each case?

5.4.5. Prove that for each n there is a constant $K = K_n$ such that

$$\max_{0 \le x \le 1} |p(x)| \le K \int_0^1 |p(x)| dx$$

for all polynomials p of degree $\le n$.

5.4.6. Let $e_1, e_2, e_3, \cdots$ denote the unit vectors (5.4.4) in l^1, l^2 or l^∞. Determine $d(e_j, e_k)$ in each of these spaces.

5.4.7. Verify that the unit vectors $e_1, e_2, e_3, \cdots$ form a normed space basis for l^1. [They do not form such a basis for l^∞.]

5.4.8. Show that l^1 is separable. [l^∞ is not.]

5.4.9. Prove that l^1 is complete.

Hint. A Cauchy sequence $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots)$ in l^1 will be componentwise convergent, $x_n^{(k)} \to y_n$, say, $\forall\, n$. It will also be bounded, $\|x^{(k)}\| = \sum_{n=1}^\infty |x_n^{(k)}| \le M$, $\forall\, n$. Deduce that $\sum_{n=1}^N |y_n| \le M$, $\forall\, N$, so that $\ldots$. Finally show that $\|y - x^{(k)}\| \to 0$.

5.4.10. Let V be an infinite dimensional normed linear space. Construct a sequence of unit vectors $e_1, e_2, \cdots$ in V such that

$$d\{e_{k+1}, S(e_1, \cdots, e_k)\} = 1, \quad k = 1, 2, \cdots .$$

Hint. Choose any u in V outside $W_k = S(e_1, \cdots, e_k)$ and consider an optimal approximation u_k for u in W_k. Compute $d(u - u_k, W_k)$, etc.

5.4.11. For complex rectangular matrices $A = [\alpha_{ij}]$ we define

$$\|A\| = \|A\|_1 = \sum_{i,j} |\alpha_{ij}|.$$

Prove that $\|AB\| \le \|A\| \cdot \|B\|$ whenever the product AB makes sense.

5.4.12. Let $\sum A_\nu = \sum [\alpha_{ij}^{(\nu)}]$ be a series of $k \times n$ matrices such that $\sum \|A_\nu\|$ converges. Prove that the series $\sum A_\nu$ is elementwise convergent, that is, the series $\sum_\nu \alpha_{ij}^{(\nu)}$ converges for each i, j.

5.4.13. Let A be an $n \times n$ matrix such that $\|A\| < 1$. Prove that the series $\sum_{\nu=0}^\infty A^\nu$ is elementwise convergent to a matrix C, and that $C = (I_n - A)^{-1}$. Here $A^0 = I_n = [\delta_{ij}]$.

5.4.14. Let A be any $n \times n$ matrix. Prove that the series $\sum_{\nu=0}^{\infty} A^\nu/\nu!$ is elementwise convergent. [The matrix sum of the series is called e^A.]

5.5. Inner products on linear spaces

This time we begin with examples.

Example 5.5.1. *The Euclidean space* $\mathbb{E}^n$. The ordinary scalar product (inner product, dot product) (x, y) or $x \cdot y$ of the vectors x, y in $\mathbb{R}^n$ is given by

$$(x, y) = x_1 y_1 + \cdots + x_n y_n. \tag{5.5.1}$$

This inner product is symmetric, and linear in each "factor". Observe that $(x, x) \geq 0$, $\forall\, x$, and that $(x, x) = 0$ if and only if $x = 0$. The Euclidean length of x can be expressed in terms of the inner product by the formula

$$\|x\| = (x, x)^{1/2} \quad \text{(nonnegative square root)}. \tag{5.5.2}$$

When both vectors x and y are $\neq 0$, the angle θ from x to y is given by

$$\cos\theta = \frac{(x, y)}{\|x\|\,\|y\|}.$$

In particular, x is perpendicular or orthogonal to y if $\cos\theta = 0$:

$$x \perp y \quad \text{if and only if} \ \ (x, y) = 0. \tag{5.5.3}$$

Adopting the convention that the zero vector is orthogonal to every vector, (5.5.3) holds in full generality.

Relation (5.5.2) is often expressed as follows: the norm function of $\mathbb{E}^n$ can be derived from an inner product. From here on, we will consider $\mathbb{E}^n$ as the space $\mathbb{R}^n$ furnished with the inner product (5.5.1), and the associated concepts of norm, distance, convergence, angle and orthogonality.

Observe that one could also have started with the notions of length and angle. As in the case of $\mathbb{R}^2$ and $\mathbb{R}^3$, one could then define the scalar product (x, y) as $\|x\|\,\|y\|\cos\theta$.

Example 5.5.2. *The unitary space* $\mathbb{U}^n$. On $\mathbb{C}^n$, formula (5.5.1) does not define a good inner product: the associated lengths $(x, x)^{1/2}$ would not always be nonnegative real numbers. One therefore uses the definition

$$(x, y) = x_1 \overline{y}_1 + \cdots + x_n \overline{y}_n. \tag{5.5.4}$$

[Actually, one could just as well define $(x, y) = \overline{x}_1 y_1 + \cdots + \overline{x}_n y_n$, as is common in physics.] Now (x, x) is real and ≥ 0 for all $x \in \mathbb{C}^n$, and $(x, x) = 0$ only if $x = 0$. One next defines $\|x\|$ by (5.5.2) and "$x \perp y$" by (5.5.3).

The result is the unitary space $\mathbb{U}^n$ with an "inner product" that gives the standard norm. Observe that the present inner product is "conjugate symmetric": $(y,x) = \overline{(x,y)}$. Our inner product is linear in the first factor, conjugate linear in the second.

From here on, we will consider $\mathbb{U}^n$ as the space $\mathbb{C}^n$ furnished with the inner product (5.5.4) and the associated concepts.

In mathematical analysis, the limit case $n \to \infty$ is important:

EXAMPLE 5.5.3. *The space $l^2 = l^2(\mathbb{N})$: "little el two".* The elements of l^2 are the infinite sequences $x = (x_1, x_2, \cdots)$ of complex numbers such that the series $\sum_{n=1}^{\infty} |x_n|^2$ converges. They form a linear space on which one can form the "inner product"

$$(x,y) = \sum_{n=1}^{\infty} x_n \overline{y}_n. \tag{5.5.5}$$

The series will be absolutely convergent since $2|x_n\overline{y}_n| \le |x_n|^2 + |y_n|^2$. Via (5.5.2) this inner product gives the l^2 norm $\|x\| = \left\{ \sum_{n=1}^{\infty} |x_n|^2 \right\}^{1/2}$ [Examples 5.3.5]. From here on, we will consider l^2 as a space with the inner product (5.5.5).

EXAMPLE 5.5.4. *The space $L^2(J)$: "big el two".* Let $J \subset \mathbb{R}$ be any finite or infinite interval. We consider the linear space $\mathcal{L}^2(J)$ of the "square-integrable functions" f on J. That is, f itself is supposed to be Lebesgue integrable over every finite subinterval of J, while $|f|^2$ must be integrable over all of J. For example, $f(x) = (1+x^2)^{-1/2}$ belongs to $\mathcal{L}^2(\mathbb{R})$. On $\mathcal{L}^2(J)$ it makes sense to define

$$(f,g) = \int_J f(x)\overline{g(x)}dx. \tag{5.5.6}$$

The integral exists because $2|f\overline{g}| \le |f|^2 + |g|^2$. Observe that $(f,f) = \int_J |f|^2$ is real and ≥ 0 for all $f \in \mathcal{L}^2(J)$. Since we want $(f,f) = 0$ only if $f = 0$, we must identify functions that are equal on J outside a set of measure zero. Now the definition

$$\|f\| = \|f\|_2 = (f,f)^{1/2} = \left\{ \int_J |f(x)|^2 dx \right\}^{1/2} \tag{5.5.7}$$

will give a true norm, the so-called L^2 norm, cf. (5.3.5). Orthogonality is [again] defined as follows:

$$f \perp y \quad \text{if and only if} \quad (f,g) = 0. \tag{5.5.8}$$

The resulting space, consisting of the square-integrable functions on J with the usual identification of almost equal functions, and with the inner product (5.5.6), as well as the associated norm etc, is called $L^2(J)$. Convergence in this space is so-called mean square convergence: when J is finite, $f_k \to f$ is the same as saying that

$$\frac{1}{\text{length } J} \int_J |f(x) - f_k(x)|^2 dx \to 0. \tag{5.5.9}$$

If J is a finite closed interval $[a, b]$ and we restrict ourselves to the continuous functions on $[a, b]$, we speak of the space $L^2\mathcal{C}[a, b]$: the continuous functions on $[a, b]$ equipped with the L^2 norm.

It is shown in Integration Theory that the space $L^2(J)$ is *complete* (Riesz–Fischer theorem; cf. [**102**]). The step functions (piecewise constant functions) on J with bounded support lie dense in $L^2(J)$. For finite (a, b), $L^2(a, b)$ is also the completion of $L^2\mathcal{C}[a, b]$.

We are now ready to give an abstract definition of inner products:

Definition 5.5.5. Let V be a real or complex linear space. A function (u, v) on $V \times V$ is called an *inner product function* if the following conditions are satisfied:

(i) The values (u, v) are scalars (hence real numbers if V is real, real or complex numbers if V is complex);

(ii) $(v, u) = \overline{(u, v)}$ for all u, $v \in V$ (conjugate symmetry);

(iii) The function (u, v) is linear in the first "factor", and hence, by (ii), conjugate linear in the second factor:

$$(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 (u_1, v) + \lambda_2 (u_2, v),$$
$$(u, \lambda_1 v_1 + \lambda_2 v_2) = \overline{\lambda}_1 (u, v_1) + \overline{\lambda}_2 (u, v_2),$$

for all scalars λ_1, λ_2 and all elements u_1, u_2, v, u, v_1, v_2 of V;

(iv) (u, u), which is real because of (ii), is nonnegative for all $u \in V$ and $(u, u) = 0$ if and only if $u = 0$.

Supposing now that V is a linear space with an inner product function $(\cdot, \cdot)$, one defines on V:

$$\begin{cases} \|u\| = (u, u)^{1/2}, \;\; d(u, v) = \|u - v\| = (u - v, u - v)^{1/2}, \\ u_k \to u \;\text{ if and only if }\; d^2(u, u_k) = (u - u_k, u - u_k) \to 0, \\ u \perp v \;\text{ if and only if }\; (u, v) = 0. \end{cases} \tag{5.5.10}$$

The function $\| \cdot \|$ will then have the properties of a norm (see Section 5.6 for the triangle inequality), and hence d will be a metric.

DEFINITION 5.5.6. An *inner product space* $V = \{V, (\cdot,\cdot)\}$ is a linear space V with an inner product function $(\cdot,\cdot)$, and the associated norm, distance, convergence and orthogonality (5.5.10).

From here on we suppose that V is an inner product space. By a *subspace* W of V we then mean a linear subspace furnished with the inner product of V. If E is a subset of V, one says that u in V is *orthogonal* to E, notation $u \perp E$, if u is orthogonal to all elements of E. The set of all elements of V that are orthogonal to E is called the *orthogonal complement* of E, notation $E^{\perp}$, ("E perp"). The orthogonal complement will be a closed linear subspace of V.

The definition below was proposed by John von Neumann (Hungary–USA, 1903–1957; [**87**]) in honor of David Hilbert; cf. the end of Section 1.6.

DEFINITION 5.5.7. A complete inner product space is called a *Hilbert space*; cf. [**49**].

For us the most important Hilbert spaces are $L^2(J)$ where J is an interval, $L^2(E)$ where E is a more general subset of a space $\mathcal{R}^n$ (unit circle, unit disc, etc.), and the related spaces where the inner product involves a weight function. Other examples are the "model space" l^2 and, of course, $\mathbb{E}^n$ and $\mathbb{U}^n$. Every inner product space can be completed to a Hilbert space.

Exercises. 5.5.1. Show that the formula $(x, y) = xy$ defines an inner product on $\mathbb{R}$. What is the corresponding inner product on $\mathbb{C}$?

5.5.2. Verify that formula (5.5.4) defines an inner product on $\mathbb{C}^n$ with the properties required in Definition 5.5.5.

5.5.3. Every inner product function on $\mathbb{R}^2$ must be of the form

$$\begin{aligned}(x, y) &= (x_1e_1 + x_2e_2, y_1e_1 + y_2e_2)\\ &= ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2.\end{aligned}$$

Under what conditions on a, b, c is this an inner product function?

5.5.4. (Continuation) What sort of curve is the "unit sphere" $S(0, 1)$ in the general inner product space $\{\mathbb{R}^2, (\cdot,\cdot)\}$, relative to Cartesian coordinates?

5.5.5. Characterize the matrices $A = [\alpha_{ij}]$ for which the formula $(x, y) = \sum_{i,j=1}^{n} \alpha_{ij}x_iy_j$ defines an inner product function on $\mathbb{R}^n$.

5.5.6. What can you say about an element u in an inner product space V that is orthogonal to all elements of V?

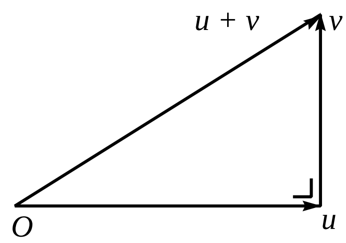

FIGURE 5.9

5.5.7. Let (u, v) be an inner product function on V. Verify that the function $(u, u)^{1/2}$ satisfies conditions (i)-(iii) for a norm in Definition 5.3.1.

5.5.8. Verify that the formula $(f, g) = \int_a^b f(x)g(x)dx$ defines an inner product on $\mathcal{C}_{\rm re}[a, b]$ in the sense of Definition 5.5.5.

5.5.9. Under what conditions on the weight function w will the formula $(f, g) = \int_a^b f(x)\overline{g(x)}w(x)dx$ define an inner product on $\mathcal{C}[a, b]$?

5.5.10. Let D be a bounded connected open set in $\mathbb{E}^2$. Verify that the formula $(u, v) = \int_D (u_x v_x + u_y v_y)dxdy$ defines an inner product on $\mathcal{C}^1_{\rm re}(\overline{D})$, provided one identifies functions that differ only by a constant.

5.5.11. Verify that formula (5.5.5) defines an inner product on the linear space of the complex sequences $x = (x_1, x_2, \cdots)$ such that $\sum_{n=1}^{\infty} |x_n|^2$ converges.

5.5.12. Prove that l^2 is complete. [Cf. Exercise 5.4.9.]

5.6. Inner product spaces: general results

In this section, V denotes an inner product space. A basic result is the "Pythagorean theorem", cf. Figure 5.9.

THEOREM 5.6.1. ("Pythagoras") *If* $u \perp v$ *in* V, *then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2. \tag{5.6.1}$$

PROOF. By Definition 5.5.5,

$$\begin{aligned}(u + v, u + v) &= (u, u) + (u, v) + (v, u) + (v, v)\\ &= (u, u) + (u, v) + \overline{(u, v)} + (v, v).\end{aligned} \tag{5.6.2}$$

If $(u, v) = 0$, the result equals $(u.u) + (v.v)$. □

Corollary 5.6.2. *For pairwise orthogonal vectors* $u_1, u_2, \cdots, u_k$, *induction will show that*

$$\|u_1 + u_2 + \cdots + u_k\|^2 = \|u_1\|^2 + \|u_2\|^2 + \cdots + \|u_k\|^2.$$

A series $\sum_1^\infty u_n$ in V whose terms are pairwise orthogonal is called an *orthogonal series*. For such series we have the important

Theorem 5.6.3. *Let* $\sum_1^\infty u_n$ *be an orthogonal series in* V. *Then*

(i) The partial sums $s_k = \sum_1^k u_n$ *form a Cauchy sequence in* V *if and only if the numerical series* $\sum_1^\infty \|u_n\|^2$ *converges;*

(ii) If V *is a Hilbert space, the orthogonal series* $\sum_1^\infty u_n$ *converges in* V *if and only if the numerical series* $\sum_1^\infty \|u_n\|^2$ *converges;*

(iii) If $\sum_1^\infty u_n = u$ *in* V, *then* $\sum_1^\infty \|u_n\|^2 = \|u\|^2$.

Proof. We write $\sum_1^k \|u_n\|^2 = \sigma_k$. Then by Pythagoras for $k > j$,

$$\begin{aligned}\|s_k - s_j\|^2 &= \|u_{j+1} + \cdots + u_k\|^2 \\ &= \|u_{j+1}\|^2 + \cdots + \|u_k\|^2 = \sigma_k - \sigma_j.\end{aligned}$$

Thus $\{s_k\}$ is a Cauchy sequence in V if and only if $\{\sigma_k\}$ is a Cauchy sequence of real numbers, or equivalently, a convergent sequence of reals. This observation proves (i) and it implies (ii). Part (iii) follows from the fact that $\sigma_k = \|s_k\|^2 \to \|u\|^2$ when $s_k \to u$ in V. Here we have anticipated the result that the function $\|u\| = (u, u)^{1/2}$ is continuous on V. This will follow from the triangle inequality which will be proved below (Applications 5.6.6). □

Another important consequence of Theorem 5.6.1 is the general *Cauchy–Schwarz inequality*, named after Cauchy (Section 1.2) and Hermann Schwarz (Germany, 1843–1921; [**112**]); cf. [**15**].

Theorem 5.6.4. (Cauchy–Schwarz) *For all vectors* u, v *in the inner product space* V *one has*

$$|(u, v)| \leq \|u\| \, \|v\|. \tag{5.6.3}$$

Proof. We may assume $u \neq 0$, $v \neq 0$. As motivation for the proof we start with the case of $\mathbb{E}^n$. There $(u, v) = \|u\| \, \|v\| \cos\theta$ (Figure 5.10), so that the desired inequality is obvious. However, the proof for $\mathbb{E}^n$ may also be based on another geometric interpretation of (u, v), one that has general validity. Let λv be the component of u in the direction of v. More precisely, let λv be the orthogonal projection of u onto the 1-dimensional subspace formed by the scalar multiples of v. By definition, λv is the orthogonal

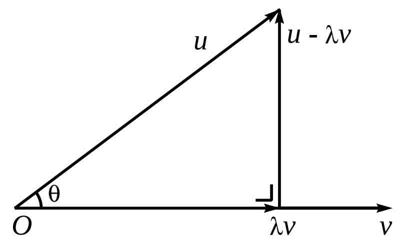

FIGURE 5.10

projection of u if $u - \lambda v \perp v$. Thus $(u - \lambda v, v) = 0$ or $(u, v) = \lambda(v, v)$, so that

$$|(u, v)| = |\lambda| \, \|v\| \, \|v\| = \|\lambda v\| \, \|v\|. \tag{5.6.4}$$

In $\mathbb{E}^n$ it is clear that $\|\lambda v\| \leq \|v\|$. In the general case we appeal to Pythagoras:

$$\|\lambda v\|^2 + \|u - \lambda v\|^2 = \|u\|^2,$$

hence $\|\lambda v\| \leq \|u\|$, so that (5.6.3) follows from (5.6.4). $\square$

REMARK 5.6.5. If $v \neq 0$, the equal sign holds in "Cauchy–Schwarz" if and only if u is a scalar multiple of v.

APPLICATIONS 5.6.6. (i) The *triangle inequality* for the norm in V. By (5.6.2) and Cauchy–Schwarz,

$$\begin{aligned} \|u + v\|^2 &= (u, u) + 2\mathrm{Re}\,(u, v) + (v, v) \\ &\leq \|u\|^2 + 2\|u\| \, \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2. \end{aligned}$$

(ii) The *continuity* of (u, v) in the first factor (or the second, or in both factors jointly):

$$|(u, v) - (u', v)| = |(u - u', v)| \leq \|u - u'\| \, \|v\|, \text{ etc.}$$

(iii) The *classical Cauchy inequality* for a sum of products. Taking $V = \mathbb{U}^n$, one has

$$\begin{aligned} \left| \sum_1^n x_k y_k \right| &= |(x, \overline{y})| \leq \|x\| \, \|\overline{y}\| \\ &= \left(\sum_1^n |x_k|^2 \right)^{1/2} \left(\sum_1^n |y_k|^2 \right)^{1/2}. \end{aligned}$$

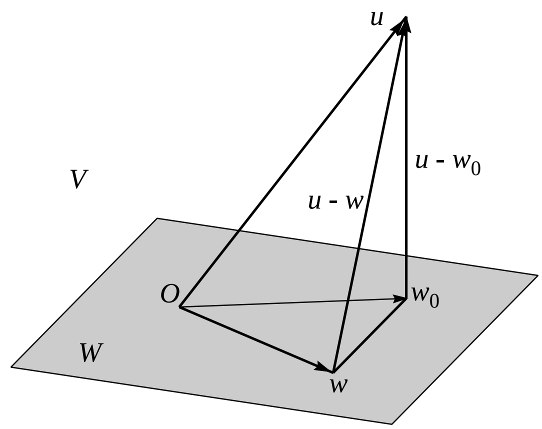

FIGURE 5.11

(iv) The *classical Schwarz inequality* for the integral of a product. For f, g in $V = L^2(J)$,

$$\left|\int_J f(x)g(x)dx\right| = |(f, \overline{g})| \leq \|f\|\,\|\overline{g}\|$$
$$= \left(\int_J |f(x)|^2 dx\right)^{1/2} \left(\int_J |g(x)|^2 dx\right)^{1/2}.$$

Optimal approximation and orthogonal projection. From here on let W be a subspace of V, and u an arbitrary element of V.

LEMMA 5.6.7. *There is at most one element $w_0 \in V$ such that $u - w_0 \perp W$.*

Indeed, if there is such an element w_0, then for any other element $w \in W$ we have $u - w_0 \perp w - w_0$, hence by Pythagoras,

$$\|u - w\|^2 = \|u - w_0\|^2 + \|w - w_0\|^2 > \|u - w_0\|^2. \tag{5.6.5}$$

Thus $u - w$ cannot be $\perp W$, for otherwise one would also have $\|u - w_0\|^2 > \|u - w\|^2$! Cf. Figure 5.11.

DEFINITION 5.6.8. If there is an element $w_0 \in W$ such that $u - w_0 \perp W$, then w_0 is called the (*orthogonal*) *projection* of u on W, notation $w_0 = Pu = P_W u$.

Inequality (5.6.5) implies the following result on approximation: if the orthogonal projection $w_0 = P_W u$ exists, it is the (unique) element of W at minimal distance from u. There is also a converse result:

THEOREM 5.6.9. *Let W be a subspace of the inner product space V and let u be an arbitrary element of V. Then the following statements about an element w_0 in W are equivalent:*

(i) w_0 is the orthogonal projection of u on W;

(ii) w_0 is an (the) *element of W that optimally approximates u relative to the metric of V.*

PROOF. In view of the preceding we need only prove that (ii) implies (i). Accordingly, let w_0 be as in (ii):

$$d(u, w_0) \le d(u, w), \quad \forall\, w \in W.$$

We have to show that $u - w_0$ is orthogonal to W.

Let us consider $w = w_0 + \varepsilon z$, where $z \in W$ is arbitrary and $\varepsilon > 0$. Then

$$\begin{aligned}\|u - w_0\|^2 \le \|u - w\|^2 &= \|u - w_0 - \varepsilon z\|^2 \\ &= \|u - w_0\|^2 - 2\varepsilon \operatorname{Re}(u - w_0, z) + \varepsilon^2 \|z\|^2.\end{aligned}$$

Hence

$$2\mathrm{Re}(u - w_0, z) \le \varepsilon \|z\|^2.$$

Letting ε go to zero, it follows that

$$\mathrm{Re}(u - w_0, z) \le 0, \quad \forall\, z \in W. \tag{5.6.6}$$

Applying (5.6.6) also to $-z \in W$ instead of z, one finds that

$$\mathrm{Re}(u - w_0, z) = 0, \quad \forall\, z \in W. \tag{5.6.7}$$

If V is a real space, (5.6.7) shows that $u - w_0 \perp W$. In the complex case one may apply (5.6.6) also to $iz \in W$. Thus one finds that $\mathrm{Im}(u - w_0, z) = 0$, $\forall\, z \in W$. The conclusion is that $(u - w_0, z) = 0$ for all $z \in W$, as had to be proved. □

If W is a *finite* dimensional subspace of V, we know that there always is an optimal approximation $w_0 \in W$ to a given $u \in V$; see Application 5.4.3. Hence for finite dimensional W, the orthogonal projection $P_W u$ exists for every $u \in V$. It will be seen in Chapter 6 that this orthogonal projection is easy to compute if we know an orthogonal basis for W.

Exercises. In these exercises V always denotes an inner product space.

5.6.1. Let u be the sum of a convergent series $\sum_1^\infty u_n$ in V. Prove that $(u, v) = \sum_1^\infty (u_n, v)$, $\forall\, v \in V$.

5.6.2. Let E be a subset of V. Prove that the orthogonal complement $E^\perp$ is a closed linear subspace of V.

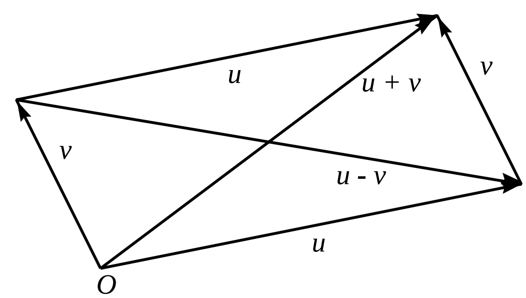

FIGURE 5.12

5.6.3. Let W be the subspace of the even functions in $V = L^2\mathcal{C}_{2\pi}$. Determine the orthogonal complement $W^\perp$.

5.6.4. Express the inner product function of V in terms of norms, (a) in the case where V is a real space, (b) in the complex case. [Cf. (5.6.2).]

5.6.5. Prove that the inner product function (u, v) is continuous on $V \times V$.

5.6.6. For incomplete V, let $\hat{V}$ denote the completion as in Section 5.2. Prove that the inner product function of V can be extended to an inner product function on $\hat{V}$ by setting $(U, U') = \lim (u_k, u'_k)$, where $\{u_k\}$ and $\{u'_k\}$ are arbitrary Cauchy sequences in V that belong to U, and U', respectively. Thus $\hat{V}$ becomes a (complete) inner product space which contains V as a dense subspace.

5.6.7. Deduce the Cauchy–Schwarz inequality for (u, v) in a real space V from the inequality

$$0 \leq (\lambda u + v, \lambda u + v) = \lambda^2\|u\|^2 + 2\lambda(u, v) + \|v\|^2, \quad \forall\, \lambda \in \mathbb{R}.$$

5.6.8. (*Parallelogram identity*) Prove that for any two vectors u, $v \in V$,

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

In words: for a parallelogram in an inner product space, the sum of the squares of the lengths of the diagonals is equal to the sum of the squares of the lengths of the four sides; cf. Figure 5.12.

[One can actually prove the following. If the identity holds for all elements u, v of a normed linear space V, the norm of V can be derived from an inner product function as in (5.5.10). (This is the Jordan–von Neumann theorem, after Pascual Jordan, Germany, 1902–1980, [**59**] and von Neumann. Cf. [**60**]).]

5.6.9. Prove that the norm function of $C[0,1]$ cannot be derived from an inner product function.

5.6.10. Prove that the unit sphere $S(0,1)$ in V cannot contain a straight line segment. For this reason, the closed unit ball $\overline{B}(0,1)$ in V is called 'rotund', or *strictly convex*. [It is sufficient to show that for $\|u\| = \|v\| = 1$ and $u \neq v$, always $\|(u+v)/2\| < 1$.]

5.6.11. Show that the norms $\|\cdot\|_\infty$ and $\|\cdot\|_1$ on $\mathbb{R}^2$ cannot be derived from inner product functions. [Cf. Exercise 5.3.4.]

5.6.12. Describe the pairs $u, v \in V$ for which the triangle inequality becomes an equality.

5.6.13. In Exercises 5.4.11–5.4.13 on matrices one may obtain related results by using the 2-norm $\|A\| = \|A\|_2 = \left(\sum_{i,j} |\alpha_{ij}|^2\right)^{1/2}$ instead of the 1-norm. Verify the essential inequality

$$\|AB\| \leq \|A\| \, \|B\|$$

for the 2-norm.

CHAPTER 6

Orthogonal expansions and Fourier series

In inner product spaces V, there are orthogonal systems of elements, which under appropriate conditions form orthogonal bases. In terms of an orthogonal system $\{v_n\}$, every element $u \in V$ has an orthogonal expansion, which may or may not converge to u. For a nice theory of orthogonal expansions it is best to work with complete inner product spaces: Hilbert spaces. The general theory applies in particular to Fourier series in the space of square-integrable functions $L^2(-\pi, \pi)$. We will characterize orthogonal bases and use them to classify inner product spaces. Every Hilbert space turns out to be like a space $l^2(\Lambda)$ for an appropriate index set Λ; cf. Exercise 6.6.7.

In the present chapter V denotes an *inner product space*, unless there is an explicit statement to the contrary.

6.1. Orthogonal systems and expansions

Prototypes of such systems and expansions are the trigonometric orthogonal system

$$\frac{1}{2}, \cos x, \sin x, \cos 2x, \sin 2x, \cdots \quad \text{in} \;\; L^2(-\pi, \pi), \tag{6.1.1}$$

and the Fourier expansion of $f \in L^2(-\pi, \pi)$.

Definition 6.1.1. An *orthogonal system* in V is a subset $\{v_\lambda\}$, where λ runs over an index set Λ, of *nonzero* elements or vectors such that

$$v_\lambda \perp v_\mu \quad \text{for all} \;\; \lambda, \mu \in \Lambda \;\; \text{with} \;\; \lambda \neq \mu.$$

If the vectors v_λ are unit vectors, we speak of an *orthonormal* system.

In most applications the index set Λ is countably infinite. In theory involving that case, we usually take $\Lambda = \mathbb{N}$, the sequence of the positive integers.

Examples 6.1.2. The system $\{\sin nx\}$, $n \in \mathbb{N}$, is orthogonal in $L^2(0,\pi)$ and also in $L^2(-\pi,\pi)$. Likewise the system $\frac{1}{2}$, $\cos x$, $\cos 2x$, $\cdots$. The complex exponentials e^{inx}, $n \in \mathbb{Z}$, form an orthogonal system in $L^2(-\pi,\pi)$, or in L^2 on the unit circle Γ when we use arc length as underlying variable. The vectors $e_1 = (1,0,0,\cdots)$, $e_2 = (0,1,0,\cdots)$, $e_3 = (0,0,1,\cdots)$, $\cdots$ form an orthogonal system in $l^2 = l^2(\mathbb{N})$; cf. Example 5.5.3.

Let $\{v_n\}$, $n = 1,2,\cdots$ be a (finite or infinite) countable orthogonal system in V. A formal series $\sum c_n v_n$ is called an *orthogonal series*: the terms are pairwise orthogonal. We will denote the partial sum $\sum_{n=1}^{k} c_n v_n$ by s_k.

Proposition 6.1.3. *Suppose that* $\sum_{n=1}^{k} c_n v_n = u$, *or that* $\sum_{n=1}^{\infty} c_n v_n = u$, *that is,* $s_k \to u$ *in* V. *Then*

$$c_n = \frac{(u,v_n)}{(v_n,v_n)}, \quad n = 1,2,\cdots. \tag{6.1.2}$$

Proof. Take $k \geq n$. Then

$$(s_k, v_n) = \left(\sum_{j=1}^{k} c_j v_j, v_n\right) = \sum_{j=1}^{k} c_j (v_j, v_n) = c_n (v_n, v_n).$$

This equality completes the proof if $s_k = u$. In the case $\sum_{n=1}^{\infty} c_n v_n = u$, the continuity of inner products [Applications 5.6.6] shows that

$$(u, v_n) = \lim_{k\to\infty} (s_k, v_n) = c_n (v_n, v_n).$$

$\square$

Corollary 6.1.4. *Orthogonal systems are linearly independent. Orthogonal representations in terms of a given orthogonal system* $\{v_n\}$ *are unique: if* $u = \sum c_n v_n = \sum d_n v_n$, *then* $d_n = c_n$, $\forall\, n$.

Definition 6.1.5. The (orthogonal) *expansion* of u in V with respect to the (countable) orthogonal system $\{v_n\}$ is the formal orthogonal series

$$u \sim \sum c_n[u] v_n, \quad \text{where } c_n[u] = \frac{(u,v_n)}{(v_n,v_n)}, \quad n = 1,2,\cdots. \tag{6.1.3}$$

The notation $u \sim \cdots$ means that u has the expansion $\cdots$; there is no implication of convergence. One uses a corresponding definition in the case of an arbitrary orthogonal system $\{v_\lambda\}$, $\lambda \in \Lambda$.

Examples 6.1.6. We have seen in Section 1.6 that Fourier series of functions f in $L^2(-\pi,\pi)$ can be considered as orthogonal expansions, provided we do not combine the terms $a_n[f]\cos nx$ and $b_n[f]\sin nx$. The expansion of a function $f \in L^2(0,\pi)$ with respect to the orthogonal system $\sin x$, $\sin 2x$, $\sin 3x$, $\cdots$ is the Fourier sine series $\sum_{n=1}^{\infty} b_n[f]\sin nx$, where $b_n[f]$ is given by the expression $(2/\pi)\int_0^\pi f(x)\sin nx\,dx$.

Basic questions. Under what conditions will an orthogonal expansion $\sum c_n[u]v_n$ of u converge in V? Under what conditions will it converge to u?

It is easy to indicate a necessary condition, but for that we need some terminology. For any subset A of V, the (linear) span or hull $S(A)$ was defined as the linear subspace of V, consisting of all finite linear combinations of elements of A [Section 5.3]. The closure $W = \overline{S}(A)$ in V is also a linear subspace of V.

Definition 6.1.7. The subspace $W = \overline{S}(A)$ of V is called the *closed* (linear) *span* of A, or the closed subspace of V generated by A. If $\overline{S}(A) = V$, that is, if every element of V can be approximated arbitrarily well by finite linear combinations of elements of A, then A is called a *spanning set* for V.

A spanning orthogonal set A is also called a *complete orthogonal set*.

If an orthogonal series $\sum c_n v_n$ converges in V, the sum must belong to the closed span $\overline{S}(v_1, v_2, \cdots)$. Thus for the convergence of the expansion $\sum c_n[u]v_n$ to u it is *necessary* that u belong to $\overline{S}(v_1, v_2, \cdots)$. We will see below that this necessary condition is also *sufficient*.

Exercises. 6.1.1. Prove that the functions $\sin(n\pi x/a)$, $n \in \mathbb{N}$, form an orthogonal system in $L^2(0,a)$, and that the functions $e^{2\pi i n x}$, $n \in \mathbb{Z}$, form an orthogonal system in $L^2(0,1)$.

6.1.2. Suppose that

$$s_k(x) = \frac{1}{2}a_0 + \sum_{n=1}^{k}(a_n\cos nx + b_n \sin nx) \to f(x)$$

in $L^2(-\pi,\pi)$ as $k \to \infty$. Prove that

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx.$$

6.1.3. Let $\{v_n\}$ be an orthogonal system in V and suppose that $u = \sum c_n v_n$. Let $v \in V$ be such that $v - u \perp v_n$ for all n. Determine the expansion $\sum d_n v_n$ of v.

6.1.4. (Orthogonal expansions need not converge to the defining element) Determine the sum of the expansion of the constant function 1 with respect to the orthogonal system $\{\sin nx\}$, $n \in \mathbb{N}$, in $L^2(-\pi, \pi)$.

6.1.5. Show that the closed span $\overline{S}(A)$ of a subset $A \subset V$ is a linear subspace of V.

6.1.6. The notion "closed span $\overline{S}(A)$" of a subset $A \subset V$ makes sense in any normed vector space V. Determine $W = \overline{S}(A)$ if $V = C[a, b]$ and $A = 1, x, x^2, \cdots$.

6.1.7. Determine the closed span of the sequence $e_1, e_2, e_3, \cdots$ in l^2.

6.1.8. (A space containing an uncountable orthogonal system) Let V be the inner product space consisting of all finite sums of the form $f(x) = \sum c_\lambda e^{i\lambda x}$, $\lambda \in \mathbb{R}$, with

$$(f, g) = \lim_{a\to\infty} \frac{1}{2a} \int_{-a}^{a} f(x)\overline{g(x)}dx.$$

Show that (f, g) is indeed an inner product function, and that the functions $v_\lambda(x) = e^{i\lambda x}$, $\lambda \in \mathbb{R}$, form a (spanning) orthogonal system in V.

6.2. Best approximation property. Convergence theorem

We begin with a lemma on the computation of orthogonal projections.

LEMMA 6.2.1. *Let W be a finite dimensional subspace of V with orthogonal basis* [basis of pairwise orthogonal vectors] $v_1, \cdots, v_k$*, and let u be any element of V. Then the* orthogonal projection *of u onto W is given by*

$$w_0 = P_W u = \sum_{n=1}^{k} c_n v_n \quad \textit{with } c_n = \frac{(u, v_n)}{(v_n, v_n)}, \quad n = 1, \cdots, k. \tag{6.2.1}$$

PROOF. The existence of the orthogonal projection w_0 follows from Theorem 5.6.9 and Application 5.4.3, but it can also be proved independently. Indeed, the elements $w \in W$ have the form $\sum_1^k \gamma_n v_n$. The condition that $u - w$ be orthogonal to $W = S(v_1, \cdots, v_k)$ is equivalent to the condition that $u - w$ be orthogonal to all v_n, hence

$$\begin{aligned} 0 &= (u - w, v_n) = (u, v_n) - (w, v_n) \\ &= (u, v_n) - \sum_{j=1}^{k} \gamma_j (v_j, v_n) = (u, v_n) - \gamma_n (v_n, v_n), \quad n = 1, \cdots, k. \end{aligned}$$

These equations have the (unique) solution $\gamma_n = (u, v_n)/(v_n, v_n) = c_n$, $n = 1, \cdots, k$. □

THEOREM 6.2.2. *Let* $\sum c_n[u]v_n$ *be the expansion of* u *in* V *with respect to the (countable) orthogonal system* $\{v_n\}$ *in* V. *Then the partial sum*

$$s_k = s_k[u] = \sum_{n=1}^{k} c_n[u]v_n$$

is equal to the orthogonal projection $w_0 = P_W u$ *of* u *onto the span* $W = S(v_1, \cdots, v_k)$. *Thus the* partial sum $s_k = s_k[u]$ *is the element of* W *which* best approximates u *relative to the metric of* V*:*

$$\|u - w\| \geq \|u - s_k\| \quad \textit{for all } \; w \in W,$$

with equality only for $w = w_0 = s_k$.

PROOF. The first part follows from Definition 6.1.5 and Lemma 6.2.1, while the second part follows from Theorem 5.6.9; cf. Figure 5.11. □

APPLICATION 6.2.3. For functions $f \in L^2(-\pi, \pi)$ one has the following result. Among all trigonometric polynomials $w(x) = \alpha_0 + \sum_1^k (\alpha_n \cos nx + \beta_n \sin nx)$ of order k, the partial sum

$$s_k[f](x) = \frac{1}{2}a_0[f] + \sum_{n=1}^{k} \big(a_n[f] \cos nx + b_n[f] \sin nx\big)$$

of the Fourier series for f provides the best approximation to f in $L^2(-\pi, \pi)$:

$$\int_{-\pi}^{\pi} \big|f - w\big|^2 \geq \int_{-\pi}^{\pi} \big|f - s_k[f]\big|^2,$$

with equality only for $w = s_k[f]$.

EXAMPLE 6.2.4. Let $f(x) = x$ on $(-\pi, \pi)$. Then among all trigonometric polynomials of order k, the partial sum

$$s_k[f](x) = 2\sum_{n=1}^{k} \frac{(-1)^{n-1}}{n} \sin nx$$

of the Fourier series provides the best approximation to f relative to the metric of $L^2(-\pi, \pi)$. [Cf. also Exercise 2.1.3.]

THEOREM 6.2.5. *Let* $A = \{v_1, v_2, \cdots\}$ *be a countable orthogonal system in* V, *and let* u *be any element of* V. *Then*

(i) If (and only if) u *belongs to the closed span* $W = \overline{S}(A)$ *in* V, *the expansion* $\sum c_n[u]v_n$ *converges to* u *in* V*;*

(ii) If (and only if) A *spans the space* V, *that is,* $\overline{S}(A) = V$, *the expansion* $\sum c_n[u]v_n$ *converges to* u *for* every u *in* V.

PROOF. We give a proof for infinite systems A, but the proof is easily adjusted to the case of finite A.

(i) Take u in $\overline{S}(A)$. Then for every $\varepsilon > 0$, there is a finite linear combination $u_\varepsilon = \sum_{n=1}^{k(\varepsilon)} d_n(\varepsilon)v_n$ of elements of A such that $\|u - u_\varepsilon\| < \varepsilon$. Now consider any integer $k \geq k(\varepsilon)$, so that u_ε belongs to $W = W_k = S(v_1, \cdots, v_k)$. By the best-approximation property [Theorem 6.2.2], the partial sum $s_k = \sum_{n=1}^{k} c_n[u]v_n$ of the expansion of u is at least as close to u as u_ε. Thus

$$\|u - s_k\| \leq \|u - u_\varepsilon\| < \varepsilon.$$

Since this holds for every ε and for all $k \geq k(\varepsilon)$, we conclude that $s_k \to u$ in V as $k \to \infty$.

(ii) Supposing $\overline{S}(A) = V$, every u in V belongs to $\overline{S}(A)$, hence by part (i), every u in V is equal to the sum of its expansion $\sum c_n[u]v_n$. $\square$

In case (ii) every u in V has a unique representation $u = \sum c_n v_n$. We then call A an *orthogonal basis* for V. [Cf. Section 6.5 below.]

In order to apply the theorem to Fourier series, we need the following

PROPOSITION 6.2.6. *The trigonometric functions*

$$\frac{1}{2},\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots \tag{6.2.2}$$

form a complete or spanning orthogonal system in $L^2(-\pi, \pi)$. *The same is* (of course) *true for the exponential functions* e^{inx}, $n \in \mathbb{Z}$.

PROOF. We have to show that every function $f \in L^2(-\pi, \pi)$ can be approximated arbitrarily well by trigonometric polynomials. Let f and $\varepsilon > 0$ be given. By Integration Theory, the step functions [piecewise constant functions] s lie dense in $L^2(-\pi, \pi)$; cf. Example 5.5.4. Hence there is a step function s such that $d_2(f, s) = \|f - s\|_2 < \varepsilon$. To such a function s we can find a continuous function g on $[-\pi, \pi]$, with $g(-\pi) = g(\pi)$, such that $d_2(s, g) < \varepsilon$. [At points where s is discontinuous, one can cut off corners.] Next, by Weierstrass's Theorem 3.4.1, there is a trigonometric polynomial T such that $|g(x) - T(x)| < \varepsilon$ throughout $[-\pi, \pi]$, so that

$$d_2(g, T) = \left(\int_{-\pi}^{\pi} |g - T|^2\right)^{\frac{1}{2}} < \sqrt{2\pi}\varepsilon.$$

The triangle inequality finally shows that

$$d_2(f, T) \leq d_2(f, s) + d_2(s, g) + d_2(g, T) < \left(2 + \sqrt{2\pi}\right)\varepsilon.$$

$\square$

APPLICATION 6.2.7. For every function f in $L^2(-\pi,\pi)$, the *Fourier series converges to* f in the sense of L^2:

$$\int_{-\pi}^{\pi} \left|f - s_k[f]\right|^2 \to 0 \quad \text{as } k \to \infty.$$

The proof follows from Theorem 6.2.5 and Proposition 6.2.6. Indeed, the Fourier series for $f \in L^2(-\pi,\pi)$ is its orthogonal expansion with respect to the complete orthogonal system (6.2.2) [provided we do not combine the terms $a_n[f]\cos nx$ and $b_n[f]\sin nx$ in the series].

Exercises. 6.2.1. Compute the sine polynomial $\sum_{n=1}^k \beta_n \sin nx$ of order k which best approximates $f(x) = 1$ in $L^2(0,\pi)$.

6.2.2. What can one say about $f \in L^2(-\pi,\pi)$ if $\int_0^\pi f(x)\sin nx dx = 0$, $\forall\, n \in \mathbb{N}$?

6.2.3. Prove that the sine polynomials $\sum \beta_n \sin nx$ lie dense in $L^2(0,\pi)$, and likewise the cosine polynomials $\alpha_0 + \sum_{n\geq 1} \alpha_n \cos nx$.

6.2.4. For $f \in L^2(0,\pi)$ one can form both a Fourier sine series and a Fourier cosine series. Prove that both series converge to f in $L^2(0,\pi)$.

6.2.5. Describe the closed subspaces of $L^2(-\pi,\pi)$ generated by the orthogonal systems

$$A_1 = \{\sin x,\, \sin 2x,\, \sin 3x,\, \cdots\}, \quad \text{and} \;\; A_2 = \{1,\, \cos x,\, \cos 2x,\, \cdots\},$$

respectively.

6.3. Parseval formulas. Convergence of expansions

Combining Theorem 6.2.2 and the Pythagorean Theorem 5.6.1, we will obtain the following important results:

THEOREM 6.3.1. *Let $A = \{v_1, v_2, \cdots\}$ be a countable orthogonal system in V, and let $\sum c_n v_n = \sum c_n[u] v_n$ be the expansion of u in V with respect to A. Then*

(i) The numerical series $\sum |c_n|^2 \|v_n\|^2$ is convergent and has sum $\leq$ $\|u\|^2$ (Bessel's inequality)*;*

(ii) If (and only if) $\sum c_n v_n = u$ in V, one has $\sum |c_n|^2\|v_n\|^2 = \|u\|^2$ (Parseval formula)*;*

(iii) If $u = \sum c_n v_n$ in V and v has expansion $\sum d_n v_n$, then

$$(u,v) = \sum c_n \overline{d}_n \|v_n\|^2 \quad \text{(extended Parseval formula).}$$

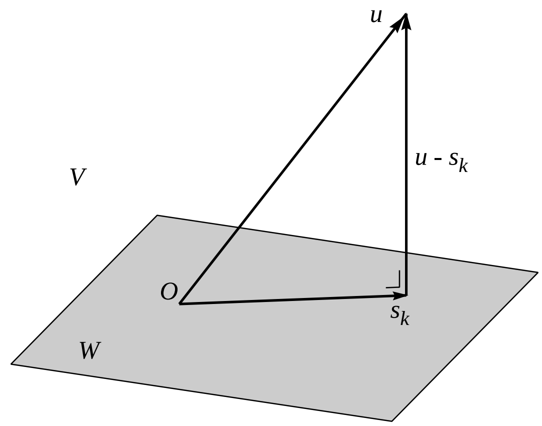

FIGURE 6.1

REMARK 6.3.2. The above results are named after the German astronomer-mathematician Friedrich W. Bessel (1784-1846; [**7**]) and the French mathematician Marc-Antoine Parseval (1755-1836; [**89**]); cf. [**90**].

PROOF OF THEOREM 6.3.1. Let k be any positive integer not exceeding the number of elements in A. Then the partial sum $s_k = \sum_{n=1}^{k} c_n v_n$ of the expansion of u is equal to the orthogonal projection of u onto the subspace $W = S(v_1, \cdots, v_k)$ of V [Theorem 6.2.2]. Hence in particular $u - s_k \perp s_k$. Thus Pythagoras gives the relations

$$\begin{aligned} \sum_{n=1}^{k} \|c_n v_n\|^2 &= \left\| \sum_{1}^{k} c_n v_n \right\|^2 = \|s_k\|^2 \\ &= \|u\|^2 - \|u - s_k\|^2 \le \|u\|^2; \end{aligned} \tag{6.3.1}$$

cf. Figure 6.1.

(i) In the case of an infinite system A, inequality (6.3.1) shows that the partial sums $\sigma_k = \sum_{n=1}^{k}$ of the numerical series $\sum_1^\infty \|c_n v_n\|^2$ are bounded by $\|u\|^2$. Hence that infinite series of nonnegative terms is convergent, and has sum $\le \|u\|^2$.

(ii) By (6.3.1), $\|u - s_k\|^2 = \|u\|^2 - \sum_1^k \|c_n v_n\|^2$. Hence for an infinite system A, the limit relation

$$\sum_{1}^{k} c_n v_n = s_k \to u \quad \text{in } V, \ \text{ or } \ \|u - s_k\|^2 \to 0 \ \text{ as } \ k \to \infty,$$

implies (and is implied by) the limit relation

$$\sigma_k = \sum_1^k \|c_n v_n\|^2 \to \|u\|^2.$$

In particular, if $\sum c_n v_n = u$, then the expansion coefficients c_n will satisfy Parseval's formula.

(iii) If v has expansion $\sum d_n v_n$, then $d_n = (v, v_n)/(v_n, v_n)$. Thus

$$(s_k, v) = \sum_1^k c_n(v_n, v) = \sum_1^k c_n\overline{(v, v_n)} = \sum_1^k c_n\overline{d}_n(v_n, v_n).$$

If A is infinite and $s_k \to u$ in V, the continuity of inner products now shows that

$$(u, v) = \lim_{k\to\infty}(s_k, v) = \sum_1^\infty c_n\overline{d}_n(v_n, v_n).$$

□

We will apply Theorem 6.3.1 to the special case of the trigonometric system (6.2.2) and the related system $\{e^{inx}\}$, $n \in \mathbb{Z}$, in $L^2(-\pi,\pi)$. The closed spans of these systems are equal to $L^2(-\pi,\pi)$. Thus we obtain the following

Corollaries 6.3.3. (i) *For any f in $L^2(-\pi,\pi)$, the Fourier coefficients $a_n = a_n[f]$, $b_n = b_n[f]$ and $c_n = c_n[f]$ satisfy the Parseval formulas*

$$\frac{|a_0|^2}{4}\,2\pi + \sum_1^\infty \left(|a_n|^2 + |b_n|^2\right)\pi = \sum_{-\infty}^\infty |c_n|^2\,2\pi = \int_{-\pi}^\pi |f(x)|^2 dx;$$

(ii) *For f, $g \in L^2(-\pi,\pi)$ one has the extended Parseval formula*

$$(f, g) = \int_{-\pi}^\pi f(x)\overline{g(x)}dx = 2\pi\sum_{-\infty}^\infty c_n[f]\overline{c}_n[g].$$

Examples 6.3.4. The Fourier series

$$2\sum_1^\infty \frac{(-1)^{n-1}}{n}\sin nx$$

for $f(x) = x$ on $(-\pi,\pi)$ [cf. Example 1.1.1] gives

$$\sum_1^\infty \frac{4}{n^2}\,\pi = \int_{-\pi}^\pi x^2 dx = \frac{2}{3}\,\pi^3, \quad \text{or} \quad \sum_1^\infty \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Similarly, the Fourier series

$$\frac{1}{3}\,\pi^2 + 4\sum_1^\infty \frac{(-1)^n}{n^2}\,\cos nx$$

for $f(x) = x^2$ on $(-\pi, \pi)$ [cf. Example 1.2.1] gives

$$\frac{\pi^4}{9}\,2\pi + \sum_1^\infty \frac{16}{n^4}\,\pi = \int_{-\pi}^{\pi} x^4 dx = \frac{2}{5}\,\pi^5, \quad \text{or} \quad \sum_1^\infty \frac{1}{n^4} = \frac{\pi^4}{90}.$$

As another nice application we mention the famous *Isoperimetric Theorem*. It says that among all simple closed curves Γ of given length L, a circle encloses the largest area A. In general,

$$4\pi A \le L^2 \qquad \text{(isoperimetric inequality)}; \tag{6.3.2}$$

see Exercise 6.3.11 and cf. [**57**].

Orthogonal expansions are orthogonal series to which we can apply the general Theorem 5.6.3. Thus Theorem 6.3.1 also has the following

COROLLARIES 6.3.5. *Let $A = \{v_1, v_2, \cdots\}$ be a countably infinite orthogonal system in V. Then:*

(i) For any u in V, the partial sums $s_k = \sum_1^k c_n[u]v_n$ of the expansion of u form a Cauchy sequence in V;

(ii) If V is a Hilbert space, all orthogonal expansions in V are convergent;

(iii) In a Hilbert space V, the expansion of u with respect to A converges to the orthogonal projection of u onto the closed subspace $W = \overline{S}(A)$ generated by A;

(iv) In a Hilbert space V, a formal series $\sum_1^\infty \gamma_n v_n$ is the expansion of an element of V if (and only if) the numerical series $\sum_1^\infty |\gamma_n|^2 \|v_n\|^2$ converges.

PROOF. Setting $c_n[u]v_n = u_n$, Bessel's inequality implies the convergence of the series $\sum_1^\infty \|u_n\|^2$. Assertions (i) and (ii) now follow from Theorem 5.6.3. From here on, let V be a Hilbert space.

(iii) The sum $w = \sum_1^\infty c_n[u]v_n = \lim s_k$ now exists in V and it must belong to $W = \overline{S}(A)$. We will show that $u - w \perp W$, so that $w = P_W u$.

Fix n and take $k \geq n$. Then by the continuity of inner products,

$$\begin{aligned}(w, v_n) &= \lim(s_k, v_n) = \lim_{k\to\infty} \sum_{j=1}^{k} c_j[u](v_j, v_n) \\ &= c_n[u](v_n, v_n) = (u, v_n).\end{aligned}$$

Hence $w - u \perp v_n$, $\forall\, n$. It follows that $w - u \perp A$, so that $w - u \perp S(A)$, and finally, $w - u \perp \overline{S}(A)$.

(iv) If $\sum_1^\infty |\gamma_n|^2 \|v_n\|^2$ converges, then the series $\sum_1^\infty \gamma_n v_n$ will converge to an element w in V by Theorem 5.6.3. The series $\sum_1^\infty \gamma_n v_n$ will be the expansion of w [Proposition 6.1.3]. □

Exercises. 6.3.1. Write down Parseval's formula for a function f in $L^2(0, 2\pi)$ and the complete orthogonal system $\{e^{inx}\}$, $n \in \mathbb{Z}$. Apply the formula to the function $f(x) = e^{\alpha x}$, $0 < x < 2\pi$, with real α.

6.3.2. Write down the Parseval formulas for the cosine series and the sine series of a function f in $L^2(0, \pi)$.

6.3.3. Let f be in $\mathcal{C}^1[0, \pi]$, $f(0) = f(2\pi) = 0$. Prove that

$$\int_0^\pi |f|^2 \leq \int_0^\pi |f'|^2.$$

For which functions f is there equality here?

6.3.4. Let f be in $L^2(-\pi, \pi)$. Prove that

$$d_2^2(f, s_k) = \int_{-\pi}^{\pi} \left|f - s_k[f]\right|^2 = \pi \sum_{n=k+1}^{\infty} \left(\left|a_n[f]\right|^2 + \left|b_n[f]\right|^2\right).$$

Also compute $d_2^2(f, \sigma_k)$, where $\sigma_k = \sigma_k[f] = (s_0 + s_1 + \cdots + s_{k-1})/k$. Does it surprise you that $d_2(f, \sigma_k) \geq d_2(f, s_k)$?

6.3.5. Compute the sum of the series $\sum_{n=1}^\infty \frac{1}{n^6}$.

6.3.6. Nobody has been able to express the sum of the series $\sum_1^\infty 1/n^3$ for $\zeta(3)$ in closed form. Express the sum $\sum_{p=1}^\infty 1/(2p-1)^3 = (7/8)\zeta(3)$ as an integral with the aid of the cosine series

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n} \quad \text{and} \quad -\frac{\pi^2}{8} + \sum_{p=1}^{\infty} \frac{\cos(2p-1)x}{(2p-1)^2};$$

cf. Exercises 1.1.4, 1.2.5.

6.3.7. Let $f(x) = 0$ on $(-\pi, 0)$, $f(x) = x$ on $(0, \pi)$. Determine the expansion of f with respect to the orthogonal system $\frac{1}{2}, \cos x, \cos 2x, \cdots$ in $L^2(-\pi, \pi)$. Calculate the sum of the expansion. [Cf. Corollaries 6.3.5.]

6.3.8. For which real values of α will the series $\sum_1^\infty (\sin nx)/n^\alpha$ converge in $L^2(0,\pi)$?

6.3.9. Let $\{\gamma_n\}$, $n \in \mathbb{Z}$, be an arbitrary sequence of complex numbers such that the series $\sum_{n=-\infty}^{\infty} |\gamma_n|^2$ converges. Prove that there is a function $f \in L^2(-\pi,\pi)$ with complex Fourier series $\sum_{-\infty}^{\infty} \gamma_n e^{inx}$.

6.3.10. Prove that the convergence and the sum of an orthogonal expansion in V are independent of the order of the terms.

6.3.11. Prove the isoperimetric inequality (6.3.2): $4\pi A \le L^2$ for piecewise smooth simple closed curves Γ. Determine all curves Γ for which there is equality.

Hint. Using arc length as parameter, Γ may be given by a formula $z = x + iy = g(s)$, $0 \le s \le L$, with $|g'(s)| \equiv 1$. One then has

$$A = \frac{1}{2}\int_{s=0}^{s=L} (xdy - ydx) = \frac{1}{2}\,\mathrm{Im}\,(g', g).$$

Without loss of generality one may take $L = 2\pi$ (so that g has period 2π). Also, the center of mass of Γ may be taken at the origin (so that $\int_0^{2\pi} g(s)ds = 0$).

6.4. Orthogonalization

Since orthogonal representations are so convenient, it is useful to know that for every sequence $\{u_1, u_2, \cdots\}$ of elements of V, there is an orthogonal sequence $\{v_1, v_2, \cdots\}$ of linear combinations of elements u_j with the *same span*.

Construction 6.4.1. (*Gram–Schmidt orthogonalization*) We start with an arbitrary (finite or infinite) sequence $\{u_1, u_2, \cdots\}$ of elements of V. Now define

$$\begin{aligned}
v_1 &= u_1;\\
v_2 &= \text{“part of } u_2 \text{ orthogonal to } v_1\text{”}\\
&= u_2 - \text{orthogonal projection of } u_2 \text{ onto } S(v_1)\\
&= u_2 - \lambda_{2,1} v_1,
\end{aligned}$$

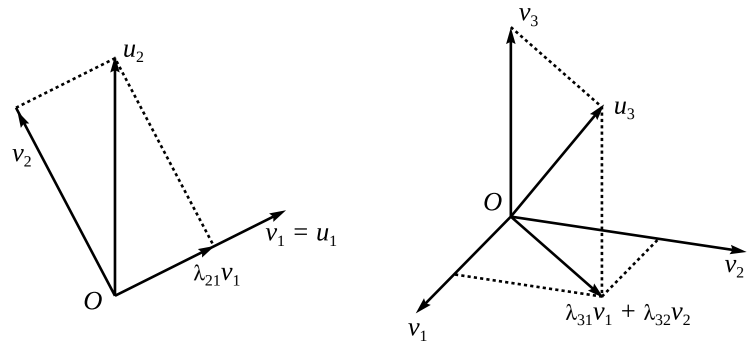

FIGURE 6.2

where the condition $v_2 \perp v_1$ gives $\lambda_{2,1} = (u_2, v_1)/(v_1, v_1)$ (but if $v_1 = 0$ we take $\lambda_{2,1} = 0$). In general, one defines

$$\begin{aligned} v_k &= \text{“part of } u_k \text{ orthogonal to } v_1, \cdots, v_{k-1}\text{”} \\ &= u_k - \text{orthogonal projection of } u_k \text{ onto } S(v_1, \cdots, v_{k-1}) \\ &= u_k - \lambda_{k,1}v_1 - \cdots - \lambda_{k,k-1}v_{k-1}, \end{aligned} \tag{6.4.1}$$

where the condition $v_k \perp v_j$ gives

$$\lambda_{k,j} = \frac{(u_k, v_j)}{(v_j, v_j)}, \quad j = 1, \cdots, k-1 \ \text{ (but if } v_j = 0 \text{ we take } \lambda_{k,j} = 0).$$

Cf. Figure 6.2. The construction is named after Jørgen P. Gram (Denmark, 1850–1916; [**40**]) and Erhard Schmidt (Germany, 1876–1959; [**108**]); cf. [**41**].

THEOREM 6.4.2. *Let $\{v_1, v_2, \cdots\}$ be the sequence of vectors in V obtained by orthogonalization of the sequence $\{u_1, u_2, \cdots\}$. Then:*

(i) Every vector v_k can be expressed as a linear combination of $u_1, \cdots, u_k$ in which u_k has coefficient 1*. Conversely, every vector u_k can be expressed as a linear combination of $v_1, \cdots, v_k$.*

(ii) For every n, $S(v_1, \cdots, v_n) = S(u_1, \cdots, u_n)$. Also, $S(v_1, v_2, \cdots) = S(u_1, u_2, \cdots)$. It follows that $\overline{S}(v_1, v_2, \cdots) = \overline{S}(u_1, u_2, \cdots)$; the sequence $\{v_1, v_2, \cdots\}$ generates the same closed subspace of V as the original sequence $\{u_1, u_2, \cdots\}$.

(iii) The vectors $v_1, v_2, \cdots$ are pairwise orthogonal. If (and only if) the vectors $u_1, u_2, \cdots$ are linearly independent, the vectors v_n are all $\neq 0$ (so that they form an orthogonal system according to Definition 6.1.1).

(iv) If the linear combinations of the vectors $u_1, u_2, \cdots$ lie dense in V and the vectors $u_1, u_2, \cdots$ are linearly independent, the vectors $v_1, v_2, \cdots$ form a spanning orthogonal set in V, hence an orthogonal basis.

PROOF. (i) Applying induction to (6.4.1), one readily shows that v_k can be expressed as a linear combination of $u_1, \cdots, u_k$ in which u_k has coefficient 1. For the other direction one may use (6.4.1) as it stands.

(ii) By (i), $v_k \in S(u_1, \cdots, u_k)$ and $u_k \in S(v_1, \cdots, v_k)$, hence every linear combination of $v_1, \cdots, v_n$ can be written as a linear combination of $u_1, \cdots, u_n$, and vice versa. The other assertions (ii) follow.

(iii) By (6.4.1), $v_k \perp$ all predecessors v_j, so that the vectors $v_1, v_2, \cdots$ are pairwise orthogonal. We also have the following equivalent assertions:

$$\begin{aligned}
&u_1, \cdots, u_n \text{ are linearly independent } \Leftrightarrow \dim S(u_1, \cdots, u_n) = n \\
&\Leftrightarrow \dim S(v_1, \cdots, v_n) = n \Leftrightarrow v_1, \cdots, v_n \text{ are linearly independent} \\
&\Leftrightarrow \text{none of the pairwise orthogonal vectors } v_1, \cdots, v_n \text{ is equal to } 0.
\end{aligned}$$

Thus if (and only if) the vectors $u_1, u_2, \cdots$ are linearly independent, *all* vectors v_n are $\neq 0$.

(iv) Assume $\overline{S}(u_1, u_2, \cdots) = V$. Then by (ii) also $\overline{S}(v_1, v_2, \cdots) = V$. The deletion of zero vectors in the sequence $\{v_n\}$ does not change the closed span, hence the nonzero vectors v_n form a spanning orthogonal set in V. □

EXAMPLE 6.4.3. (*Legendre polynomials*) [after Adrien-Marie Legendre (France, 1752–1833; [**78**]), who contributed to both pure and applied mathematics.] Orthogonalization of the sequence of powers $\{1, x, x^2, x^3, \cdots\}$ in $L^2(-1, 1)$ gives the following sequence of polynomials:

$$\begin{aligned}
&p_0(x) = 1, \;\; p_1(x) = x - \lambda_{1,0}1 = x, \\
&p_2(x) = x^2 - \lambda_{2,0}1 - \lambda_{2,1}x = x^2 - 1/3, \\
&p_3(x) = x^3 - \lambda_{3,0}1 - \lambda_{3,1}x - \lambda_{3,2}(x^2 - 1/3) = x^3 - (3/5)x, \text{ etc.}
\end{aligned}$$

The polynomials $p_n(x)$, $n \in \mathbb{N}_0$, form an orthogonal system; the degree of $p_n(x)$ is exactly n. It can be shown that $p_n(1) \neq 0$ for all n; see Section 7.1 below. Division of $p_n(x)$ by $p_n(1)$ gives the *Legendre polynomial* P_n:

$$P_n(x) \stackrel{\text{def}}{=} \frac{p_n(x)}{p_n(1)}, \quad \text{so that } \; P_n(1) = 1, \; \forall\, n. \tag{6.4.2}$$

One will find

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}\,(3x^2 - 1),$$

$$P_3(x) = \frac{1}{2}\,(5x^3 - 3x), \quad P_4(x) = \frac{1}{8}(35x^4 - 30x^3 + 3), \text{ etc.} \tag{6.4.3}$$

It is useful to know that

$$\|P_n\|^2 = \int_{-1}^{1} P_n^2(x)dx = \frac{1}{n+\frac{1}{2}}. \tag{6.4.4}$$

This and other results will be derived in Chapter 7.

Application 6.4.4. For any given $k \geq 0$, the Legendre polynomials $P_0, P_1, \cdots, P_k$ form an orthogonal basis for the subspace $W = S(1, x, \cdots, x^k)$ of $L^2(-1,1)$, which consists of the polynomials in x of degree $\leq k$ [restricted to the interval $(-1,1)$]. For functions $f \in L^2(-1,1)$ one can form the Legendre series: the expansion $\sum_{n=0}^{\infty} c_n[f]P_n$ with respect to the system $\{P_n\}$. The orthogonal projection of f onto W is equal to

$$s_k[f] = \sum_{n=0}^{k} c_n[f]P_n, \quad \text{where } c_n[f] = \frac{(f, P_n)}{(P_n, P_n)} = (n+1/2)\int_{-1}^{1} fP_n.$$

This is the polynomial of degree $\leq k$ which *best approximates* f on $(-1,1)$ in the sense of "least squares"; one has

$$\int_{-1}^{1} \left|f - s_k[f]\right|^2 \leq \int_{-1}^{1} \left|f - P\right|^2$$

for all polynomials P of degree $\leq k$, with equality only for $P = s_k[f]$; cf. Theorem 6.2.2.

Exercises. 6.4.1. Orthogonalize the sequence of vectors $u_1 = (1,1,1)$, $u_2 = (2,1,0)$, $u_3 = (1,0,0)$ in $\mathbb{E}^3$. Will orthogonalization of the sequence u_3, u_2, u_1 give the same result?

6.4.2. Orthogonalize the sequence $\{1, x, x^2\}$ in $L^2(0,1)$, and standardize to norm 1 through multiplication by suitable positive constants.

$\big[$Answer: $\big\{1,\ \sqrt{12}(x-\frac{1}{2}),\ \sqrt{180}(x^2 - x + \frac{1}{6})\big\}.\ \big]$

6.4.3. Compute $p_4(x)$ in Example 6.4.3 and verify the formula for $P_4(x)$ in (6.4.3).

6.4.4. Let $f(x) = |x|$, $-1 \leq x \leq 1$. Determine the polynomial P of degree ≤ 2 which best approximates f in $L^2(-1,1)$. Also compute $\|f - P\|^2$.

6.4.5. (Continuation) Next determine the linear combination T of 1, $\cos \pi x$, $\sin \pi x$ which best approximates f in $L^2(-1,1)$. Which of the two, P and T, provides a better approximation?

6.4.6. Show that the polynomials in x lie dense in $L^2(-1,1)$, and deduce that the Legendre polynomials P_n, $n \in \mathbb{N}_0$, form a spanning orthogonal set, or orthogonal basis, for $L^2(-1,1)$.

6.4.7. Let $\{u_1, u_2, \cdots\}$ be a sequence of vectors in V, $\{v_1, v_2, \cdots\}$ the sequence obtained by orthogonalization. Show that $\|v_n\| \le \|u_n\|$, $\forall\, n$.

6.4.8. The *Gram matrix* of vectors $u_1, \cdots, u_n$ in V is defined by

$$\begin{aligned} G(u_1, \cdots, u_n) &= \big[(u_i, u_j)\big]_{i,j=1,\cdots,n} \\ &= \begin{bmatrix} (u_1,u_1) & (u_1,u_2) & \dots & (u_1,u_n) \\ (u_2,u_1) & (u_2,u_2) & \dots & (u_2,u_n) \\ \dots & \dots & \dots & \dots \\ (u_n,u_1) & (u_n,u_2) & \dots & (u_n,u_n) \end{bmatrix}. \end{aligned}$$

Prove that the determinant, $\det G$, is invariant under orthogonalization:

$$\det G(u_1, \cdots, u_n) = \det G(v_1, \cdots, v_n) = \|v_1\|^2 \cdots \|v_n\|^2.$$

6.4.9. Show that $u_1, \cdots, u_n$ are linearly independent in V if and only if $\det G(u_1, \cdots, u_n) \neq 0$.

6.4.10. Let $A = [\alpha_{ij}]$ be an $n \times n$ complex matrix. Verify that

$$A\overline{A}^T = \big[(u_i, u_j)\big],$$

where $u_1, \cdots, u_n$ denote the row vectors of A, considered as elements of $\mathbb{U}^n$. Deduce *Hadamard's inequality*

$$|\det A| \le \|u_1\| \cdots \|u_n\|.$$

Interpret the inequality geometrically when A is a real matrix.

[Jacques Salomon Hadamard (France, 1865–1963; [**42**]) is perhaps best known for the proof of the *prime number theorem* in 1896. The theorem was proved independently - in the same year - by the Belgian mathematician Charles-Jean de la Vallée Poussin (1866–1962, [**121**]). The prime number theorem says that $\pi(x)$, the number of primes $\le x$, behaves like $x/\log x$ for large x. More precisely, cf. [**96**],

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} \to 1 \quad \text{as } x \to \infty.]$$

6.5. Orthogonal bases

A spanning orthogonal sequence $\{v_1, v_2, \cdots\}$ in V is also called an orthogonal basis, because every element u in V will have a unique representation $u = c_1v_1 + c_2v_2 + \cdots$; cf. Theorem 6.2.5. Here the coefficients c_n are the expansion coefficients of u: $c_n = c_n[u] = (u, v_n)/(v_n, v_n)$. The order of the terms in the expansion is not important.

For general orthogonal systems we formulate

DEFINITION 6.5.1. An orthogonal system $\{v_\lambda\}$, $\lambda \in \Lambda$, in V is called an *orthogonal basis* for V if every element u in V has a unique representation as the sum of a finite or convergent infinite series of terms $c_\lambda v_\lambda$.

In every representation of u as a finite or infinite sum $\sum c_\lambda v_\lambda$, the coefficient c_λ must be equal to $(u, v_\lambda)/(v_\lambda, v_\lambda)$, hence the representation is unique (up to the order of the terms). It is the expansion of u with respect to the system $\{v_\lambda\}$, and it never has more than a countable number of nonzero terms; see Lemma 6.5.3 below.

It is natural to ask which inner product spaces have a *countable* orthogonal basis. Suppose that the space V has such an orthogonal basis $\{v_n\}$. Then V must be *separable*, that is, V must contain a countable set of elements $u_1, u_2, \cdots$ which lies dense in V. Indeed, the finite linear combinations $\sum c_n v_n$ must be dense in V, and these can be approximated by finite combinations $\sum \gamma_n v_n$ with "rational" coefficients γ_n, that is, $\mathrm{Re}\, c_n$ and $\mathrm{Im}\, c_n$ rational. The latter combinations form a countable set. There is also a converse result:

THEOREM 6.5.2. *Every separable inner product space V has a countable orthogonal basis.*

Indeed, a dense sequence of elements $u_1, u_2, \cdots$ is *a fortiori* a spanning sequence. Orthogonalization will produce a spanning sequence of pairwise orthogonal vectors $v_1, v_2, \cdots$ [Theorem 6.4.2]. Omitting all zero vectors v_n from that sequence, one obtains a countable orthogonal basis.

LEMMA 6.5.3. *Let $\{v_\lambda\}$, $\lambda \in \Lambda$, be an orthogonal system in V. Then for every element u in V, at most countably many of the expansion coefficients $c_\lambda[u] = (u, v_\lambda)/(v_\lambda, v_\lambda)$ are different from zero.*

PROOF. For any finite subset Λ_0 of Λ, Bessel's inequality shows that

$$\sum_{\lambda \in \Lambda_0} \left| c_\lambda[u] \right|^2 \|v_\lambda\|^2 \le \|u\|^2; \tag{6.5.1}$$

cf. Theorem 6.3.1. It follows that the number of λ's for which $\left|c_\lambda[u]\right|\|v_\lambda\|$ is ≥ 1 is at most equal to $\|u\|^2$. Similarly, the number of λ's for which $\left|c_\lambda[u]\right|\|v_\lambda\|$ is less than 1 but greater than or equal to $\frac{1}{2}$ is bounded by $4\|u\|^2$. The number of λ's for which $\frac{1}{2} > \left|c_\lambda[u]\right|\|v_\lambda\| \geq \frac{1}{3}$ is bounded by $9\|u\|^2$, etc. Thus the nonzero terms $c_\lambda[u]v_\lambda$ in the expansion of u can be arranged in a *sequence* according to decreasing norm – they form a countable set. $\square$

The proof of Theorem 6.2.5 may be adapted to the case of general orthogonal systems to give

THEOREM 6.5.4. *An orthogonal system $\{v_\lambda\}$ in V is an orthogonal basis if and only if the finite linear combinations of the vectors v_λ lie dense in V.*

For Hilbert spaces, there is the following useful characterization of orthogonal bases:

THEOREM 6.5.5. *In a complete inner product space V, an orthogonal system $\{v_\lambda\}$ is an orthogonal basis if and only if it is a* maximal *orthogonal system.*

Such maximality means that there is no vector y in V that is orthogonal to all vectors v_λ, except the vector $y = 0$.

PROOF OF THE THEOREM. (i) Let $\{v_\lambda\}$ be an orthogonal basis of V and suppose that $y \in V$ is orthogonal to every v_λ. Then $(y, v_\lambda) = 0$ for all $\lambda \in \Lambda$, hence y has the expansion 0. Since y must be equal to the sum of its expansion, $y = 0$.

(ii) Let $\{v_\lambda\}$ be a maximal orthogonal system in V and let u be an arbitrary element of V. Considering only the nonzero terms, we form the expansion of u, $\sum c_{\lambda_n} v_{\lambda_n}$. By Bessel's inequality, the series $\sum \left|c_{\lambda_n}[u]\right|^2 \|v_{\lambda_n}\|^2$ must converge, and hence, V being complete, the series $\sum c_{\lambda_n} v_{\lambda_n}$ converges to an element $w \in V$ [Corollaries 6.3.5].

The difference $y = u - w$ will be orthogonal to every vector v_λ. This is clear if λ is of the form λ_n, but it is also true if λ is different from all λ_n. The maximality of the orthogonal system $\{v_\lambda\}$ now shows that $y = 0$, hence $u = w = \sum c_{\lambda_n} v_{\lambda_n}$. Since u was arbitrary, it follows that $\{v_\lambda\}$ is an orthogonal basis of V. $\square$

*The characterization in Theorem 6.5.5 may be used to show that every Hilbert space V, no matter how large, has an orthogonal basis. The proof requires a form of the axiom of choice, such as Zorn's Lemma (after Max

Zorn, Germany–USA, 1906–1993; [**127**]); cf. [**128**], or see the book [**119**]. By that Lemma, V will contain a maximal orthogonal system. One can also show that all orthogonal bases of a given Hilbert space have the same cardinal number. This cardinal number is sometimes called the *orthogonal dimension* of the space. In the subsequent theory, we will restrict outselves to separable spaces V.

Exercises. 6.5.1. Prove that the vectors $e_1, e_2, e_3, \cdots$ of Example 6.1.2 form an orthonormal basis for $l^2 = l^2(\mathbb{N})$.

6.5.2. Which of the following systems are orthogonal bases of the given spaces? Explain your answers.

$$\begin{aligned}
&\{\sin x, \sin 2x, \sin 3x, \cdots\} \quad \text{in} \;\; L^2(0,\pi);\\
&\{\cos x, \cos 2x, \cos 3x, \cdots\} \quad \text{in} \;\; L^2(0,\pi);\\
&\{e^{inx},\; n \in \mathbb{Z}\} \quad \text{in} \;\; L^2(-\pi,\pi);\\
&\{P_0, P_1, P_2, P_3, \cdots\} \quad \text{in} \;\; L^2(-1,1).
\end{aligned}$$

6.5.3. Let $\{\phi_0, \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \cdots\}$ be an orthogonal basis for $L^2(-a,a)$ such that $\phi_0, \phi_2, \phi_4, \cdots$ are even functions, while $\phi_1, \phi_3, \phi_5, \cdots$ are odd functions. Prove that the system $\{\phi_0, \phi_2, \phi_4, \cdots\}$ is an orthogonal basis for $L^2(0,a)$, and likewise the system $\{\phi_1, \phi_3, \phi_5, \cdots\}$.

6.5.4. Same questions as in Exercise 6.5.2 for the systems

$$\begin{aligned}
&\{\sin(\pi x/a), \sin(2\pi x/a), \sin(3\pi x/a), \cdots\} \quad \text{in} \;\; L^2(0,a);\\
&\{P_0, P_2, P_4, \cdots\} \quad \text{in} \;\; L^2(0,1);\\
&\{\cos x, \sin 2x, \cos 3x, \sin 4x, \cdots\} \quad \text{in} \;\; L^2(-\pi/2,\pi/2);\\
&\{\sin x, \cos 2x, \sin 3x, \cos 4x, \cdots\} \quad \text{in} \;\; L^2(-\pi/2,\pi/2);\\
&\{\sin x, \sin 3x, \sin 5x, \cdots\} \quad \text{in} \;\; L^2(0,\pi/2);\\
&\{\cos x, \cos 3x, \cos 5x, \cdots\} \quad \text{in} \;\; L^2(0,\pi/2).
\end{aligned}$$

6.5.5. Let $L^2(a,b;w)$, where $w(x)$ is an almost everywhere positive (measurable) function on (a,b), denote the Hilbert space of the functions $f(x)$ such that $f(x)\sqrt{w(x)}$ belongs to $L^2(a,b)$, with the inner product

$$(f,g) = \int_a^b f(x)\overline{g(x)}w(x)dx.$$

Show that the following systems are orthogonal bases of the given spaces:

$$\{T_0(x), T_1(x), T_2(x), \cdots\} \quad \text{in } L^2\Big(-1, 1; \frac{1}{\sqrt{1-x^2}}\Big) \quad \text{[Exercise 3.4.4]};$$
$$\{P_0(\cos\theta), P_1(\cos\theta), P_2(\cos\theta), \cdots\} \quad \text{in } L^2(0, \pi; \sin\theta).$$

6.5.6. Let V be a Hilbert space, W a (separable) closed subspace. Prove that every element $u \in V$ has an orthogonal projection $P_W u$. How can one compute $P_W u$?

6.5.7. Let V be an inner product space, A an orthogonal basis for V. Prove that A is also an orthogonal basis for the completion $\hat{V}$ of V.

6.5.8. Let W be the subspace of $L^2(0, \pi)$ consisting of all finite linear combinations of the functions $x, \sin 2x, \sin 3x, \cdots$. Prove that $A = \{\sin 2x, \sin 3x, \cdots\}$ is a maximal orthogonal system in W, but not an orthogonal basis.

6.5.9. Prove Theorem 6.5.4, and give an example of an inner product space V with an uncountable orthogonal basis.

6.5.10. Let A be an orthogonal system in V, and let E be a subset of V whose linear span $S(E)$ lies dense in V. Suppose that the expansion of every element $u \in E$ with respect to A converges to u in V. Prove that A is an orthogonal basis for V.

6.5.11. Let $\{\phi_1, \phi_2, \phi_3, \cdots\}$ be an orthonormal system in $L^2(0, 1)$.

(i) Let $0 \le a \le 1$. Determine the expansion of the step function s_a such that $s_a(x) = 1$ on $[0, a]$, $s_a(x) = 0$ on $(a, 1]$.

(ii) Deduce that

$$\sum_{n=1}^{\infty} \left| \int_0^a \phi_n(x) dx \right|^2 \le a. \tag{6.5.2}$$

(iii) Prove that one has equality in (6.5.2) for every $a \in [0, 1]$ if and only if $\{\phi_1, \phi_2, \phi_3, \cdots\}$ is an orthonormal basis for $L^2(0, 1)$.

6.5.12. Let the functions $f_n(x)$, $n \in \mathbb{N}$, form an orthogonal basis for $L^2(a, b)$, and let the functions $g_k(y)$, $k \in \mathbb{N}$, form an orthogonal basis for $L^2(c, d)$. Prove that the products $f_n(x)g_k(y)$, $n, k = 1, 2, \cdots$, form an orthogonal basis for L^2 on the domain $(a < x < b,\ c < y < d)$.

Deduce that the functions $e^{i(nx+ky)}$, $n, k = 0, \pm 1, \pm 2, \cdots$, form an orthogonal basis for L^2 on the rectangle $(-\pi < x < \pi,\ -\pi < y < \pi)$.

6.6. Structure of inner product spaces

We will first show that all orthogonal bases of a separable inner product space have the same number of elements, or more accurately, the same cardinal number. In the following it is convenient, and no loss of generality, to restrict ourselves to orthonormal bases and systems.

LEMMA 6.6.1. *Let* $\{v_\lambda\}$*,* $\lambda \in \Lambda$*, be an orthonormal system in the separable inner product space* V*. Then the system* $\{v_\lambda\}$ *is countable.*

PROOF. Observe that $\|v_\lambda - v_\mu\|^2 = 2$ whenever $\lambda \neq \mu$, so that the open balls $B(v_\lambda, \sqrt{2})$, $\lambda \in \Lambda$, are pairwise disjoint. Now let $\{u_1, u_2, \cdots\}$ be a sequence which lies dense in the separable space V. For given $\lambda \in \Lambda$, the ball $B(v_\lambda, \sqrt{2})$ will contain certain elements u_n; the one with the lowest index will be called u_{n_λ}. Doing this for each λ, we obtain a one-to-one correspondence between the elements of the system $\{v_\lambda\}$ and a subset of the positive integers. Conclusion: the system $\{v_\lambda\}$ is either finite, or countably infinite. □

THEOREM 6.6.2. *Let* V *be a separable inner product space different from just a zero vector. Then all orthonormal bases of* V *have the same cardinal number, which is either a positive integer or countably infinite.*

PROOF. By Lemma 6.6.1 every orthonormal basis of V is countable. Let $A = \{v_1, v_2, \cdots\}$ be such a basis. Then there are two possibilities:

(i) A is finite, $A = \{v_1, \cdots, v_n\}$, say. In this case A is an ordinary algebraic basis for V, since every element of V has a unique representation $\sum_{j=1}^n c_j v_j$. Thus by Linear Algebra, every linearly independent set in V has at most n elements. In particular all orthonormal bases of V are finite, and hence ordinary algebraic bases. Since the latter all have the same number of elements, so do all orthonormal bases.

(ii) A is infinite, $A = \{v_n\}$, $n \in \mathbb{N}$. In this case the (algebraic) dimension of V must be infinite; cf. part (i). Hence all orthonormal bases of V must be infinite, and by Lemma 6.6.1 they must be countably infinite. □

We will now classify the separable real and complex inner product spaces. To that end we need a suitable concept of isomorphism.

DEFINITION 6.6.3. Two inner product spaces V and V' are called *isomorphic*, notation $V \cong V'$, if there is a one to one map T of V onto V' which commutes with addition and multiplication by scalars:

$$T(u_1 + u_2) = Tu_1 + Tu_2, \quad T\lambda u = \lambda Tu, \ \forall\, \lambda,$$

and which preserves inner products:

$$(Tu_1, Tu_2)_{V'} = (u_1, u_2)_V.$$

The first condition says that T is linear, and the final condition may be expressed by saying that T must be an *isometry*. Indeed, if T preserves inner products, it will automatically preserve norms and distances, and vice versa.

THEOREM 6.6.4. *Let V be a separable inner product space $\neq \{0\}$. Then one of the following three cases must pertain:*

(i) $\dim V$ *is equal to a positive integer n. In this case V is isomorphic to* $\mathbb{E}^n$ (it it is a real space)*, or to* $\mathbb{U}^n$ (if it is complex)*;*

(ii) $\dim V = \infty$ *and V is complete. In this case V is isomorphic to* $l^2 = l^2(\mathbb{N})$ (if it is a complex space)*, or to* l^2_{re} (if it is real)*;*

(iii) $\dim V = \infty$ *and V is incomplete. In that case V is isomorphic to a dense subspace of l^2 or l^2_{re}.*

PROOF. We will discuss the case where V is a (separable) infinite dimensional complex Hilbert space. In this case every orthonormal basis of V has the form $\{v_n\}$, $n \in \mathbb{N}$. Fixing such a basis, the elements $u \in V$ are precisely the sums $\sum_{n=1}^{\infty} c_n v_n$, with $c_n \in \mathbb{C}$, $\sum_{n=1}^{\infty} |c_n|^2 < \infty$; cf. Corollaries 6.3.5. We now define a map T of V to l^2 by setting

$$T \sum_{n=1}^{\infty} c_n v_n = \{c_1, c_2, c_3, \cdots\} = \sum_{n=1}^{\infty} c_n e_n. \tag{6.6.1}$$

Here $\{e_n\}$, $n \in \mathbb{N}$, is the standard orthonormal basis of $l^2 = l^2(\mathbb{N})$; cf. Exercise 6.5.1. By the definition of l^2 [Example 5.5.3], the map T is one to one and onto, it commutes with addition and multiplication by scalars, and it preserves inner products:

$$\begin{aligned}\left(\sum c_n v_n, \sum d_k v_k\right)_V &= \sum c_n \overline{d}_n \|v_n\|^2 \\ = \sum c_n \overline{d}_n &= \left(\sum c_n e_n, \sum d_k e_k\right)_{V'}.\end{aligned}$$

Cf. the extended Parseval formula in Theorem 6.3.1.

If V is an incomplete (complex) separable inner product space, $\dim V$ must be infinite; cf. Theorem 6.4.2. Thus V has an orthonormal basis $\{v_n\}$, $n \in \mathbb{N}$. This time the linear map T with the rule (6.6.1) will establish an isomorphism of V with a dense but incomplete subspace of l^2. $\square$

Exercises. 6.6.1. Prove part (i) of Theorem 6.6.4.

6.6.2. Explicitly describe an isomorphism between $l^2(\mathbb{Z})$ and $l^2(\mathbb{N})$. [Cf. Example 5.5.3.]

6.6.3. Let V and V' be isomorphic inner product spaces. Prove that any completions $\hat{V}$ and $\hat{V}'$ are also isomorphic. In particular any two completions of a given inner product space V are isomorphic.

6.6.4. Let V and V' be isomorphic and let V be complete. Prove that V' is also complete.

6.6.5. Let $\{v_\lambda\}$ be an orthonormal basis of V and let T be an isomorphism of V onto V'. Prove that $\{Tv_\lambda\}$ is an orthonormal basis of V'.

6.6.6. Describe the completion $\hat{V}$ of the "large" inner product space V in Exercise 6.1.8.

6.6.7. How would you define the Hilbert space $l^2(\Lambda)$, where Λ is an arbitrary index set? Prove that every complex Hilbert space V is isomorphic to some space $l^2(\Lambda)$.

CHAPTER 7

Classical orthogonal systems and series

Both in pure and applied mathematics, one meets a large variety of orthogonal systems besides the trigonometric functions. Boundary value problems for differential equations are an important source. In fact, there are large classes of eigenvalue problems for differential equations, whose standardized eigenfunctions form orthogonal bases. Thus the "Legendre eigenvalue problem" of mathematical physics leads to the Legendre polynomials; cf. [**79**] and Section 8.3. In this and the next chapter we will study these and other orthogonal polynomials from different points of view.

7.1. Legendre polynomials: Properties related to orthogonality

In Example 6.4.3, the Legendre polynomials $P_n(x)$ were obtained by orthogonalization of the sequence of powers $\{1, x, x^2 \cdots\}$ in $L^2(-1, 1)$, and subsequent standardization of the resulting pairwise orthogonal polynomials $p_0(x)$, $p_1(x)$, $p_2(x)$, $\cdots$ through multiplication by suitable constants:

$$P_n(x) = \frac{1}{p_n(1)}\, p_n(x), \quad n = 0, 1, 2, \cdots . \tag{7.1.1}$$

[Here it was assumed that $p_n(1) \neq 0$; a proof will be given below.] Thus $P_n(x)$ is a polynomial in x of precise degree n, so that every polynomial in x of degree $\leq n$ can be expressed as a linear combination of $P_0(x), P_1(x), \cdots,$ $P_n(x)$. It is convenient to formulate the following

Definition 7.1.1. The *Legendre polynomial* $P_n(x)$ is the unique polynomial of degree n in x, which is orthogonal to $1, x, \cdots, x^{n-1}$ in $L^2(-1, 1)$, and for which $P_n(1) = 1$.

The existence and uniqueness of P_n can be proved by linear algebra, cf. Exercises 7.1.1, 7.1.2, but the following *construction* does more: it gives an explicit representation. Let $P_n(x)$ be *any* polynomial of precise degree n which is orthogonal to $1, x, \cdots, x^{n-1}$ (there surely is such a polynomial; cf.

Exercise 7.1.1):

$$(P_n(x), x^s) = \int_{-1}^{1} P_n(x)x^s dx = 0, \quad s = 0, 1, \cdots, n-1.$$

We now introduce auxiliary polynomials

$$P_{n,1}(x) = \int_{-1}^{x} P_n(t)dt, \;\; P_{n,2}(x) = \int_{-1}^{x} P_{n,1}(t)dt, \;\cdots,$$
$$P_{n,k}(x) = \int_{-1}^{x} P_{n,k-1}(t)dt, \;\cdots,\; P_{n,n}(x) = \int_{-1}^{x} P_{n,n-1}(t)dt,$$

so that $P_{n,k}(x)$ has precise degree $n+k$. For $n \geq 1$,

$$P_{n,1}(1) = \int_{-1}^{1} P_n \cdot 1 = \big(P_n(x), x^0\big) = 0, \;\; P_{n,1}(-1) = 0.$$

Hence, integrating by parts,

$$(s+1)\big(P_{n,1}(x), x^s\big) = \int_{-1}^{1} P_{n,1}(x)dx^{s+1}$$
$$= \big[P_{n,1}(x)x^{s+1}\big]_{-1}^{1} - \int_{-1}^{1} P_n(x)x^{s+1}dx = 0, \;\; s = 0, 1, \cdots, n-2.$$

Next, for $n \geq 2$,

$$P_{n,2}(1) = \int_{-1}^{1} P_{n,1} \cdot 1 = 0, \;\; P'_{n,2}(1) = P_{n,1}(1) = 0,$$
$$P'_{n,2}(-1) = P_{n,2}(-1) = 0,$$
$$(s+1)(P_{n,2}(x), x^s) = -(P_{n,1}(x), x^{s+1}) = 0, \;\; s = 0, 1, \cdots, n-3.$$

Thus, inductively, for $n \geq k$,

$$P_{n,k}(1) = P'_{n,k}(1) = \cdots = P_{n,k}^{(k-1)}(1) = 0,$$
$$P_{n,k}(-1) = P'_{n,k}(-1) = \cdots = P_{n,k}^{(k-1)}(-1) = 0,$$
$$(P_{n,k}(x), x^s) = 0, \;\; s = 0, 1, \cdots, n-k-1.$$

For $k = n$ we run out of orthogonality relations, but find

$$P_{n,n}(1) = P'_{n,n}(1) = \cdots = P_{n,n}^{(n-1)}(1) = 0,$$
$$P_{n,n}(-1) = P'_{n,n}(-1) = \cdots = P_{n,n}^{(n-1)}(-1) = 0.$$

It now follows from Taylor's formula for $P_{n,n}(x)$ around the point $x = 1$ that

$$\begin{aligned}
P_{n,n}(x) &= P_{n,n}(1) + P'_{n,n}(1)(x-1) + \cdots + P^{(n-1)}_{n,n}(1)\,\frac{(x-1)^{n-1}}{(n-1)!} \\
&\qquad + \cdots + P^{(2n)}_{n,n}(1)\frac{(x-1)^{2n}}{(2n)!}, \\
&= P^{(n)}_{n,n}(1)\frac{(x-1)^n}{n!} + \cdots + P^{(2n)}_{n,n}(1)\frac{(x-1)^{2n}}{(2n)!}.
\end{aligned}$$

Thus the polynomial $P_{n,n}(x)$ of degree $2n$ is divisible by $(x-1)^n$. It is likewise divisible by $(x+1)^n$, hence by $(x^2-1)^n$. Conclusion:

$$P_{n,n}(x) = \alpha_n(x^2-1)^n, \quad P_n(x) = \alpha_n D^n(x^2-1)^n, \;\; D = \frac{d}{dx}\,,$$

where α_n is a constant. We finally compute $P_n(1)$ by Leibniz's formula for the n-th derivative of a product:

$$\begin{aligned}
P_n(1) &= \alpha_n\Big[D^n\{(x-1)^n(x+1)^n\}\Big]_{x=1} \\
&= \alpha_n \sum_{k=0}^{n} \binom{n}{k} \Big[D^{n-k}(x-1)^n D^k(x+1)^n\Big]_{x=1}.
\end{aligned}$$

The terms on the right are all equal to zero except for the term with $k = 0$:

$$P_n(1) = \alpha_n \binom{n}{0}\, n!\, 2^n.$$

Thus we can impose the condition $P_n(1) = 1$ and it gives $\alpha_n = 1/(2^n n!)$:

Theorem 7.1.2. (*Rodrigues' formula* for the Legendre polynomials). *One has*

$$\begin{aligned}
P_n(x) &= \frac{1}{2^n n!}\, D^n(x^2-1)^n \\
&= \frac{1\cdot 3\cdots(2n-1)}{n!}\, x^n - \frac{1\cdot 3\cdots(2n-3)}{2\cdot(n-2)!}\, x^{n-2} + \cdots\,. \qquad (7.1.2)
\end{aligned}$$

Notice that P_n is an even function when n is even, and an odd function when n is odd.

Formulas for orthogonal polynomials such as the one above are named after the French banker and mathematician Olinde Rodrigues (1795–1851; [**103**]); cf. [**104**].

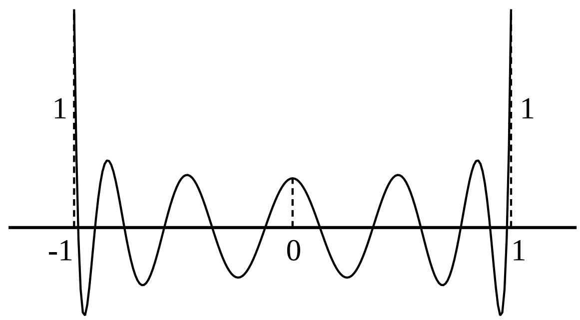

FIGURE 7.1

Formula (7.1.2) will provide information on the general appearance of the graph of $P_n(x)$. When $n \geq 1$, the even polynomial $(x^2-1)^n$ has zeros of multiplicity n at $x=1$ and $x=-1$. Thus the derivative $D(x^2-1)^n$ (which is an odd polynomial) has zeros of multiplicity $n-1$ at $x=\pm 1$, and by Rolle's theorem, it has at least one zero between -1 and $+1$. In view of the degree $2n-1$ of the derivative, there can be only one such zero, and it must be simple; it lies at the origin, of course. When $n \geq 2$ one finds that the (even) polynomial $D^2(x^2-1)^n$ of degree $2n-2$ has zeros of multiplicity $n-2$ at ± 1, and exactly two simple zeros between -1 and $+1$. For $n \geq k$, the polynomial $D^k(x^2-1)^n$ of degree $2n-k$ has zeros of multiplicity $n-k$ at ± 1, and exactly k simple zeros between -1 and $+1$. Taking $k=n$, we obtain

PROPOSITION 7.1.3. *All n zeros of the Legendre polynomial $P_n(x)$ are real and simple, and they lie between -1 and $+1$.*

The derivative $P_n'(x)$ will have exactly $n-1$ simple zeros; they separate the zeros of $P_n(x)$. Thus the graph of $P_n(x)$ on $\mathbb{R}$ has precisely $n-1$ relative extrema. They occur at points between -1 and $+1$ which alternate with the zeros. Beyond the last minimum point the graph is rising and free of inflection points. On the closed interval $[-1,1]$ there will be $n+1$ relative extrema, including the end points; cf. Figure 7.1.

It will follow from Exercise 7.1.11 that the successive relative maximum values of $|P_n(x)|$ on $[0,1]$ form an increasing sequence.

Since the polynomials in x lie dense in $C[-1,1]$ while the continuous functions lie dense in $L^2(-1,1)$ [cf. Example 5.5.4], every function $f(x)$ in $L^2(-1,1)$ can be approximated arbitrarily well in L^2 sense by linear combinations of Legendre polynomials. Thus the Legendre polynomials form an orthogonal basis for $L^2(-1,1)$; cf. Theorem 6.5.4. With a little work their norms may be obtained from Rodrigues' formula:

$$\begin{aligned} \|P_n\|^2 &= \int_{-1}^1 P_n P_n = -\int_{-1}^1 P_{n,1}P_n' = \cdots = (-1)^n \int_{-1}^1 P_{n,n}P_n^{(n)} \\ &= (-1)^n \int_{-1}^1 \alpha_n (x+1)^n (x-1)^n \cdot \alpha_n (2n)!\, dx \\ &= (-1)^{n-1}\alpha_n^2 (2n)! \int_{-1}^1 \frac{(x+1)^{n+1}}{n+1} \cdot n(x-1)^{n-1} dx = \cdots \qquad (7.1.3) \\ &= \alpha_n^2 (2n)! \int_{-1}^1 \frac{(x+1)^{2n}}{(n+1)\cdots(2n)}\, n(n-1)\cdots 1\, dx = \cdots = \frac{1}{n+\frac{1}{2}}. \end{aligned}$$

THEOREM 7.1.4. (Basis property) *Every function f in $L^2(-1,1)$ is equal to the sum of its Legendre expansion or* Legendre series*,*

$$f = \sum_{n=1}^{\infty} c_n[f]P_n, \quad c_n[f] = (n+1/2)\int_{-1}^1 f P_n.$$

Here the convergence is L^2 convergence: $\int_{-1}^1 \left|f - s_k[f]\right|^2 \to 0$ as $k \to \infty$.

Orthogonal systems such as $\{P_n\}$ satisfy a three-term recurrence relation by which P_{n+1} may be expressed in terms of P_n and P_{n-1}. Indeed, observing that the leading coefficient in P_{n+1} is equal to $(2n+1)/(n+1)$ times the leading coefficient in P_n [see Theorem 7.1.2], one finds that

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) = Q_{n-1}(x),$$

a polynomial of degree $\leq n-1$. Here the left-hand side is orthogonal to $1, x, \cdots, x^{n-2}$ in $L^2(-1,1)$, hence $Q_{n-1}(x) = \beta_{n-1}P_{n-1}(x)$, a constant multiple of $P_{n-1}(x)$. The constant β_{n-1} may be evaluated by setting $x=1$: $\beta_{n-1} = -n$.

PROPOSITION 7.1.5. (Recurrence relation) *One has $P_0(x) = 1$, $P_1(x) = x$, and*

$$(n+1)P_{n+1}(x) - (2n+1)xP_n(x) + nP_{n-1}(x) = 0 \qquad (n \geq 1).$$

A differential equation for $P_n(x)$ may be obtained in a similar manner. Let $R_n(x)$ be the polynomial $\{(1-x^2)P_n'(x)\}'$ of degree n. Integration by parts shows that $R_n(x) \perp 1, x, \cdots, x^{n-1}$ in $L^2(-1,1)$:

$$\int_{-1}^{1} R_n(x)x^s dx = -s\int_{-1}^{1}(1-x^2)P_n'(x)x^{s-1}dx$$
$$= s\int_{-1}^{1} P_n(x)\{(s-1)x^{s-2}-(s+1)x^s\}dx = 0, \quad 0 \le s \le n-1.$$

It follows that $R_n(x) = \gamma_n P_n(x)$. Comparison of the leading coefficients shows that $\gamma_n = -n(n+1)$. Conclusion:

PROPOSITION 7.1.6. (Differential equation) *The Legendre polynomial* $y = P_n(x)$ *satisfies the differential equation*

$$\{(1-x^2)y'\}' + n(n+1)y = 0.$$

There are other ways to obtain the Legendre differential equation; cf. Examples 8.3.1 below.

Another important consequence of the orthogonality is Gauss's "quadrature formula" (after Carl Friedrich Gauss, Germany, 1777–1855; [**35**]). This is a formula for *numerical integration*; see Exercise 7.1.6 and cf. [**36**].

Exercises. 7.1.1. Let W_k denote the subspace $S(1, x, \cdots, x^k)$ of dimension $k+1$ in $L^2(-1,1)$. Prove that the vectors in W_n that are orthogonal to W_{n-1} form a 1-dimensional subspace. Deduce that the orthogonality condition $P_n \perp W_{n-1}$ determines P_n up to a constant multiple.

7.1.2. Let P_n be any real polynomial of precise degree n such that $\int_{-1}^{1} P_n Q = 0$ for all real polynomials Q of degree $\le n-1$. Deduce directly that P_n must change sign at least n times (hence precisely n times) between -1 and $+1$. In particular such a polynomial cannot vanish at an end point ± 1.

7.1.3. Prove that the polynomials $D^n\{(x-a)^n(x-b)^n\}$, $n \in \mathbb{N}_0$, form an orthogonal system, and in fact, an orhogonal basis, in $L^2(a,b)$.

7.1.4. Use the recurrence relation, Proposition 7.1.5, to compute P_2, P_3 and P_4 from $P_0(x) = 1$, $P_1(x) = x$.

7.1.5. Use the differential equations for P_n and P_k to verify that $(P_n, P_k) = 0$ whenever $k \ne n$.

7.1.6. (*Gauss quadrature*) Let $x_1 < x_2 < \cdots < x_k = x_{n,k} < \cdots < x_n$ denote the consecutive zeros of P_n. Prove that there are constants $\lambda_j = \lambda_{n,j}$ such that

$$\int_{-1}^{1} f(x)dx = \sum_{j=1}^{n} \lambda_j f(x_j) \tag{7.1.4}$$

for all polynomials $f(x)$ of degree $\le 2n-1$.

[The coefficients λ_j are determined by the condition that (7.1.4) must be correct for $f(x) = 1, x, \cdots, x^{n-1}$, hence for all polynomials f of degree $\le n-1$. Furthermore, any f of degree $\le 2n-1$ can be written as $P_n Q + R$, with $\deg Q < n$, $\deg R < n$. Conclusion?]

7.1.7. (Continuation) (a) compute the numbers x_j and λ_j for $n = 2$ and $n = 3$.

(b) Prove that the constants $\lambda_j = \lambda_{n,j}$ are all positive. [Choose f cleverly!]

7.1.8. Suppose $g \in C[-1, 1]$ has all its power moments equal to zero: $\int_{-1}^1 g(x)x^n dx = 0$, $\forall\, n \in \mathbb{N}_0$. Prove that $g(x) \equiv 0$. Can you prove a corresponding result for $f \in L^2(-1, 1)$? For $f \in L^1(-1, 1)$?

7.1.9. Obtain the Legendre series for $f(x) = |x|$ on $[-1, 1]$. Prove that the series converges uniformly on $[-1, 1]$. Does the series converge pointwise to $|x|$?

[Observe that for even n, $c_n[f] = (2n + 1)P_{n,2}(0) = \cdots$.]

7.1.10. Use Rodrigues' formula to show that

$$
\begin{aligned}
&P_n(x) = \sum_{0 \le k \le \frac{1}{2}n} (-1)^k \, \frac{1 \cdot 3 \cdots (2n - 2k - 1)}{2^k \, k!(n - 2k)!} \, x^{n-2k}, \\
&P_{2m}(0) = (-1)^m 2^{-2m} \binom{2m}{m} \sim (-1)^m/\sqrt{\pi m}, \\
&P'_{2m+1}(0) = (2m + 1)P_{2m}(0).
\end{aligned}
$$

7.1.11. Use the differential equation in Proposition 7.1.6 to show that the function

$$
v(x) = v_n(x) = P_n^2(x) + \frac{1}{n(n + 1)}\,(1 - x^2)P_n'(x)^2
$$

is strictly increasing on $[0, 1]$. Deduce that the relative maxima of $|P_n(x)|$ on $[0, 1]$ form an increasing sequence, and that

$$
|P_n(x)| \le P_n(1) = 1 \quad \text{on} \;\; [-1, 1].
$$

7.1.12. (Continuation) Show that $\int_0^1 v_n = 2\int_0^1 P_n^2 = 1/(n + 1/2)$,

$$
v_n(0) \sim \frac{2}{\pi n}, \quad v_n(x) \le \frac{1}{(n + \frac{1}{2})(1 - x)} \;\; \text{on} \;\; [0, 1).
$$

7.1.13. Show that $P_n' - P_{n-2}'$ is orthogonal to $1, x, \cdots, x^{n-2}$ on $(-1,1)$, and deduce that $P_n' - P_{n-2}' = (2n-1)P_{n-1}$. Next prove that

$$P_n' = (2n-1)P_{n-1} + (2n-5)P_{n-3} + (2n-9)P_{n-5} + \cdots,$$

and deduce that $|P_n'(x)| \le P_n'(1) = n(n+1)/2$ on $[-1,1]$.

7.1.14. Compute $\int_{-1}^{1} x^n P_s(x)dx$ for $0 \le s \le n$, and deduce that x^n is equal to

$$\sum_{0\le k\le \frac{1}{2}n} \frac{n(n-1)\cdots(2k+2)}{(2n-2k+1)(2n-2k-1)\cdots(2k+3)}\,(2n-4k+1)P_{n-2k}(x).$$

7.2. Other orthogonal systems of polynomials

All classical systems of orthogonal polynomials can be obtained by orthogonalization of the sequence of powers

$$\{1, x, x^2, \cdots\}, \tag{7.2.1}$$

and standardization of the resulting polynomials through multiplication by suitable constants. On the interval $(-1,1)$ different weight functions lead to different orthogonal systems. Thus orthogonalization of the sequence (7.2.1) relative to the weight function $w(x) = (1-x^2)^{-\frac{1}{2}}$ leads to the *Chebyshev polynomials*

$$T_n(x) \stackrel{\text{def}}{=} \cos n\theta\Big|_{\cos\theta = x}, \quad n = 0, 1, 2, \cdots; \tag{7.2.2}$$

cf. Section 3.4 and Exercise 7.2.1 below. Similarly, orthogonalization of the sequence (7.2.1) relative to the weight function $w(x) = (1-x^2)^{+\frac{1}{2}}$ on $(-1,1)$ leads to the so-called Chebyshev polynomials of the second kind:

$$U_n(x) \stackrel{\text{def}}{=} \frac{\sin(n+1)\theta}{\sin\theta}\Bigg|_{\cos\theta = x}, \quad n = 0, 1, 2, \cdots. \tag{7.2.3}$$

More generally, the weight functions

$$(1-x^2)^\alpha,\ \alpha > -1, \quad \text{and} \quad (1-x)^\alpha(1+x)^\beta,\ \alpha > -1,\ \beta > -1$$

on $(-1,1)$ lead to the *ultraspherical*, and the *Jacobi* polynomials, respectively. [Carl G.T. Jacobi, German mathematician, 1809–1851; [**58**].]

Spherical polynomials and associated Legendre functions. We will consider the important weight function $(1-x^2)^k$, $k \in \mathbb{N}_0$, on $(-1,1)$, which leads to the so-called spherical polynomials. For $k = 0$ these are simply the Legendre polynomials P_n, $n \ge 0$. For $k = 1$ one will obtain (scalar multiples of) their derivatives P_n', $n \geq 1$. Indeed, by the differential equation for P_n [Proposition 7.1.6],

$$\begin{aligned}\int_{-1}^{1} P_n'(x)P_s'(x)(1-x^2)dx &= \Big[(1-x^2)P_n'(x)P_s(x)\Big]_{-1}^{1} \\ &\quad - \int_{-1}^{1} \big\{(1-x^2)P_n'(x)\big\}' P_s(x)dx \\ &= n(n+1)\int_{-1}^{1} P_n(x)P_s(x)dx = 0, \quad \forall\, s \neq n.\end{aligned}$$

Repeated differentiation of the differential equation for P_n shows that the k^{th} derivatives $z = P_n^{(k)}(x)$ satisfy the differential equation

$$-\big\{(1-x^2)^{k+1}z'\big\}' = (n-k)(n+k+1)(1-x^2)^k z. \tag{7.2.4}$$

It now follows inductively that the polynomials $P_n^{(k)}$ form an orthogonal system in $L^2(-1,1;(1-x^2)^k)$. Indeed, we know this for $k=0$ (and $k=1$). Assuming the result for order k, we will see that (7.2.4) gives it for order $k+1$. Abbreviating $P_n^{(m)}(x)$ to $P_n^{(m)}$ in the following integrals and omitting dx, one has

$$\begin{aligned}\int_{-1}^{1} P_n^{(k+1)}P_s^{(k+1)}(1-x^2)^{k+1} &= \Big[(1-x^2)^{k+1}P_n^{(k+1)}P_s^{(k)}\Big]_{-1}^{1} \\ &\quad - \int_{-1}^{1} \big\{(1-x^2)^{k+1}P_n^{(k+1)}\big\}' P_s^{(k)} \\ &= (n-k)(n+k+1)\int_{-1}^{1} P_n^{(k)}P_s^{(k)}(1-x^2)^k = 0, \quad \forall\, s \neq n.\end{aligned}$$

Multiplying the "spherical polynomials" $P_n^{(k)}(x)$ by the square root of the weight function, hence by $(1-x^2)^{\frac{1}{2}k}$, we obtain an orthogonal system in $L^2(-1,1)$. In fact, one can say more:

THEOREM 7.2.1. *For every $k \in \mathbb{N}_0$, the functions*

$$P_n^k(x) \overset{\text{def}}{=} (1-x^2)^{\frac{1}{2}k}P_n^{(k)}(x), \quad n = k, k+1, \cdots,$$

called associated Legendre functions *of order k, form an* orthogonal basis *of* $L^2(-1,1)$. *One has*

$$\|P_n^k\|^2 = \int_{-1}^{1} \big|P_n^k(x)\big|^2 dx = \frac{(n+k)!}{(n-k)!}\,\frac{1}{n+\frac{1}{2}}.$$

Proof. The functions P_n^k, $n \geq k$, will form a *maximal* orthogonal system in $L^2(-1,1)$. Indeed, suppose that for $g \in L^2(-1,1)$,

$$\int_{-1}^{1} (1-x^2)^{\frac{1}{2}k} P_n^{(k)}(x) g(x) dx = 0, \ \ \forall\, n \geq k.$$

Then all power moments of the L^2 function $(1-x^2)^{\frac{1}{2}k} g(x)$ on $(-1,1)$ will be equal to zero, since $P_n^{(k)}$ has precise degree $n-k$. Thus $(1-x^2)^{\frac{1}{2}k} g(x)$ has Legendre series 0, hence $g(x) = 0$ almost everywhere, so that $g = 0$ in $L^2(-1,1)$; cf. Theorem 7.1.4.

Furthermore, by (7.2.4) with $k-1, k-2, \cdots, 0$ instead of k,

$$\begin{aligned}
&\int_{-1}^{1} \left|P_n^k\right|^2 = \int_{-1}^{1} (1-x^2)^k P_n^{(k)} \cdot P_n^{(k)} = -\int_{-1}^{1} \left\{(1-x^2)^k P_n^{(k)}\right\}' P_n^{(k-1)} \\
&= \{n-(k-1)\}(n+k) \int_{-1}^{1} \left\{P_n^{k-1}\right\}^2 = \cdots \\
&= [\{n-(k-1)\} \cdots n] \cdot [(n+k) \cdots (n+1)] \int_{-1}^{1} P_n^2.
\end{aligned}$$

□

First substituting $z = (1-x^2)^{-\frac{1}{2}k} y$ in (7.2.4), and in a second step, setting $x = \cos\theta$, one obtains the associated Legendre equation of order k, and the polar associated Legendre equation of order k, respectively:

Proposition 7.2.2. *The associated Legendre function*

$$y = P_n^k(x) = (1-x^2)^{\frac{1}{2}k} P_n^{(k)}(x), \quad k \in \mathbb{N}_0,$$

satisfies the differential equation

$$-\left\{(1-x^2)y'\right\}' + \frac{k^2}{1-x^2}\, y = n(n+1)y, \quad -1 < x < 1,$$

and the related function

$$w = P_n^k(\cos\theta) = (\sin\theta)^k P_n^{(k)}(\cos\theta)$$

satisfies the differential equation

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta \frac{dw}{d\theta}\right) + \frac{k^2}{\sin^2\theta}\, w = n(n+1)w, \quad 0 < \theta < \pi.$$

Laguerre polynomials (after Edmond Laguerre, France, 1834–1886; [**71**]); cf. [**72**]). On unbounded intervals, weight functions are indispensable, since over such intervals, the powers x^n fail to be integrable. The simplest weight function for the interval $(0, \infty)$ is e^{-x}. Orthogonalization of the sequence (7.2.1) in $L^2(0, \infty; e^{-x})$, and subsequent standardization through multiplication by suitable constants, lead to the Laguerre polynomials.

Definition 7.2.3. The Laguerre polynomial $L_n(x)$ is the unique polynomial of degree n in x which is orthogonal to $1, x, \cdots, x^{n-1}$ in $L^2(0, \infty; e^{-x})$, and for which $L_n(0) = 1$.

The existence and uniqueness of L_n can be proved by linear algebra or by the following explicit *construction*. Let L_n be any polynomial of precise degree n such that

$$(L_n, x^s) = \int_0^\infty L_n(x) x^s e^{-x} dx = 0, \quad s = 0, 1, \cdots, n-1.$$

One may now introduce auxiliary functions $L_{n,k}(x)$ by setting

$$L_{n,k}(x) e^{-x} = \int_0^x L_{n,k-1}(t) e^{-t} dt, \quad k = 1, \cdots, n; \quad L_{n,0} = L_n. \tag{7.2.5}$$

Induction on k will show that $L_{n,k}(x)$ is a polynomial of precise degree n such that

$$\begin{aligned} &(L_{n,k}, x^s) = 0 \quad \text{for } \ s = 0, 1, \cdots, n-k-1, \ \text{ and} \\ &L_{n,k}(x) = \mathcal{O}(x^k) \quad \text{as } \ x \searrow 0. \end{aligned} \tag{7.2.6}$$

Indeed, assume that (7.2.5) holds for some $k < n$. Then for $n \geq 1$, (repeated) integration by parts will give

$$L_{n,k+1}(x) e^{-x} = p(x) e^{-x} + c \int_0^x e^{-t} dt,$$

where $\deg p = n$ and c is a constant. However, since $(L_{n,k}, 1) = 0$, (7.2.5) with $k+1$ instead of k shows that $L_{n,k+1}(x) e^{-x} \to 0$ as $x \to \infty$, hence $c = 0$, so that $L_{n,k+1}(x) = p(x)$. Furthermore, for $s \leq n-k-2$, application of (7.2.5) with $k+1$ instead of k and (7.2.6) show that

$$\begin{aligned} (s+1)(L_{n,k+1}, x^s) &= \int_0^\infty \left\{ L_{n,k+1}(x) e^{-x} \right\} dx^{s+1} \\ &= -\int_0^\infty L_{n,k}(x) e^{-x} x^{s+1} dx = 0. \end{aligned}$$

That $L_{n,k+1}(x) = \mathcal{O}(x^{k+1})$ at 0 follows immediately from (7.2.6) and (7.2.5).

Conclusion for $k = n$: $L_{n,n}(x)$ is a polynomial of precise degree n that is divisible by x^n, hence $L_{n,n}(x) = \alpha_n x^n$. It follows that

$$L_n(x)e^{-x} = D_n\big\{L_{n,n}(x)e^{-x}\big\} = \alpha_n D^n\big(x^n e^{-x}\big).$$

Setting $x = 0$ one finds that $L_n(0) = \alpha_n\, n!$. Thus we can impose the condition $L_n(0) = 1$ and it gives $\alpha_n = 1/n!$.

THEOREM 7.2.4. (*Rodrigues type formula* for the Laguerre polynomials)*:*

$$L_n(x) = \frac{1}{n!}\, e^x D^n\big(x^n e^{-x}\big) = \sum_{k=0}^{n} \binom{n}{k} \frac{(-x)^k}{k!}, \quad n = 0, 1, 2, \cdots . \tag{7.2.7}$$

PROPERTIES 7.2.5. All n *zeros* of the Laguerre polynomial $L_n(x)$ are positive real and simple. The *norm* of L_n is equal to 1:

$$\|L_n\|^2 = \int_0^\infty L_n^2(x)e^{-x}dx = 1.$$

One has $L_0(x) = 1$, $L_1(x) = -x + 1$ and the *recurrence relation*

$$(n+1)L_{n+1}(x) + (x - 2n - 1)L_n(x) + nL_{n-1}(x) = 0.$$

Furthermore $y = L_n(x)$ satisfies the *differential equation*

$$xy'' + (1 - x)y' + ny = 0.$$

The Laguerre polynomials $L_n(x)$, $n \in \mathbb{N}_0$, form an orthonormal basis for $L^2(0, \infty; e^{-x})$. Equivalently, the *Laguerre functions*

$$L_n(x)e^{-\frac{1}{2}x}, \quad n \in \mathbb{N}_0,$$

form an *orthonormal basis* for $L^2(0, \infty)$. A standard proof is based on the theory of Laplace transforms [Chapter 11]; cf. Exercise 7.2.11.

Exercises. 7.2.1. Orthogonalize the sequence of powers $\{1, x, x^2 \cdots\}$ in $L^2\left(-1, 1; (1 - x^2)^{-\frac{1}{2}}\right)$ to obtain polynomials $t_n(x)$, $n \in \mathbb{N}_0$. Show that $t_n(1) \neq 0$, and that division of $t_n(x)$ by $t_n(1)$ gives the Chebyshev polynomial $T_n(x)$.

Hint. The trigonometric polynomials $t_n(\cos\theta)$ of order n, $n \in \mathbb{N}_0$, form an orthogonal system in $L^2(0, \pi)$. Hence $t_n(\cos\theta)$ is a linear combination of $1/2, \cos\theta, \cdots, \cos n\theta$. Etc.

7.2.2. Orthogonalize the sequence $\{1, x, x^2 \cdots\}$ in $L^2\left(-1, 1; (1-x^2)^{+\frac{1}{2}}\right)$ to obtain polynomials $u_n(x)$, $n \in \mathbb{N}_0$. Show that $u_n(1) \neq 0$, and that multiplication of $u_n(x)$ by a suitable constant gives the Chebyshev polynomial of the second kind $U_n(x)$.

7.2.3. Derive the differential equation (7.2.4) for $z = P_n^{(k)}(x)$.

7.2.4. Derive the differential equations for the associated Legendre functions in Proposition 7.2.2.

7.2.5. Prove that the functions $P_n^k(\cos\theta)$, $n \geq k$, form an orthogonal basis of $L^2(0, \pi; \sin\theta)$ for every $k \in \mathbb{N}_0$.

7.2.6. The polynomial $P_{n,k}(x)$ of degree $n+k$ [Section 7.1] is divisible by $(1-x)^k$ and orthogonal to $1, x, \cdots, x^{n-k-1}$ in $L^2(-1,1)$. Prove that $P_{n,k}(x)$ must be a scalar multiple of $(1-x^2)^k P_n^{(k)}(x)$.

7.2.7. Prove that the n zeros of $L_n(x)$ are positive real and distinct.

7.2.8. Prove that

$$\begin{aligned}(L_n, L_n) = \int_0^\infty L_n(x) L_n(x) e^{-x} dx &= -(L_{n,1}, L_n') = \cdots \\ &= (-1)^n (L_{n,n}, L_n^{(n)}) = 1.\end{aligned}$$

7.2.9. Derive the recurrence relation for the Laguerre polynomials after showing that $(n+1)L_{n+1} + (x - 2n - 1)L_n$ is a polynomial of degree $\leq n-1$ that is orthogonal to $1, x, \cdots, x^{n-2}$ in $L^2(0, \infty; e^{-x})$.

7.2.10. Derive the differential equation for $y = L_n(x)$ after showing that $(D-1)\{xL_n'(x)\}$ is a polynomial of degree n that is orthogonal to $1, x, \cdots, x^{n-1}$ in $L^2(0, \infty; e^{-x})$.

7.2.11. Use the following facts about Laplace transforms [which will be proved later] to show that the Laguerre functions $L_n(x)e^{-\frac{1}{2}x}$, $n \in \mathbb{N}_0$, form an orthogonal basis of $L^2(0, \infty)$:

(i) For g in $L^1(0, \infty)$ or $L^2(0, \infty)$, the Laplace transform

$$(\mathcal{L}g)(s) = \int_0^\infty g(x) e^{-sx} dx$$

is analytic in the right half-plane $\{\operatorname{Re} s > 0\}$;

(ii) The nth derivative of $(\mathcal{L}g)(s)$ is given by

$$(\mathcal{L}g)^{(n)}(s) = \int_0^\infty (-x)^n g(x) e^{-sx} dx, \quad \operatorname{Re} s > 0;$$

(iii) If $\mathcal{L}g = 0$, then $g = 0$.

Hint. If $g(x) \in L^2(0,\infty)$ is orthogonal to $L_n(x)e^{-\frac{1}{2}x}$, $\forall\, n$, then $g(x) \perp x^n e^{-\frac{1}{2}x}$ in $L^2(0,\infty)$, $\forall\, n$. What can you conclude about $(\mathcal{L}g)(s)$ then?

7.2.12. The so-called generalized Laguerre polynomials $L_n^{(\alpha)}(x)$, $n \in \mathbb{N}_0$, are obtained by orthogonalization of the sequence of powers $\{1, x, x^2, \cdots\}$ in $L^2(0,\infty; x^\alpha e^{-x})$, $\alpha > -1$, and standardization so as to make $L_n^{(\alpha)}(0)$ equal to $\binom{n+\alpha}{n}$. Prove that

$$L_n^{(\alpha)}(x)x^\alpha e^{-x} = \frac{1}{n!}\, D^n\left(x^{n+\alpha}e^{-x}\right) = \sum_{k=0}^{n} \binom{n+\alpha}{n-k} \frac{(-x)^k}{k!}.$$

7.2.13. Show that the generalized Laguerre polynomials $y = L_n^{(\alpha)}(x)$ satisfy the differential equation $xy'' + (\alpha + 1 - x)y' + ny = 0$.

7.3. Hermite polynomials and Hermite functions

We finally consider the doubly infinite interval $(-\infty,\infty)$. In this case the simplest weight function is e^{-x^2}. Orthogonalization of the sequence $\{1, x, x^2, \cdots\}$ in $L^2\left(-\infty,\infty; e^{-x^2}\right)$, and subsequent standardization through multiplication by suitable constants, lead to the Hermite polynomials (named after Charles Hermite, France, 1822–1901; [**46**]); cf. [**47**]).

Definition 7.3.1. The Hermite polynomial $H_n(x)$ is the unique polynomial of degree n in x which is orthogonal to the powers $1, x, \cdots, x^{n-1}$ in $L^2\left(-\infty,\infty; e^{-x^2}\right)$ and has leading coefficient 2^n.

The definition leads to the following simple formula:

Theorem 7.3.2. (*Rodrigues type formula* for the Hermite polynomials)*:*

$$H_n(x) = (-1)^n e^{x^2} D^n e^{-x^2} = 2^n\left(x^n - \frac{n(n-1)}{2^2}\, x^{n-2} + \cdots\right). \tag{7.3.1}$$

In particular

$$H_0(x) = 1,\ \ H_1(x) = 2x,\ \ H_2(x) = 4x^2 - 2,\ \ H_3(x) = 8x^3 - 12x.$$

Remark 7.3.3. In Mathematical Statistics it is customary to use a slightly different weight function, namely, $e^{-\frac{1}{2}x^2}$. The corresponding (modified) Hermite polynomials $\tilde{H}_n(x)$ have properties very similar to those of the polynomials $H_n(x)$; see Exercise 7.3.11.

For the proof of Theorem 7.3.2 one introduces auxiliary functions $H_{n,k}(x)$ by setting

$$H_{n,k}(x)e^{-x^2} = \int_{-\infty}^{x} H_{n,k-1}(t)e^{-t^2}dt, \quad k = 1, \cdots, n, \quad H_{n,0}(x) = H_n(x).$$

One then proves by induction on k that $H_{n,k}(x)$ is a polynomial of precise degree $n-k$ such that

$$(H_{n,k}, x^s) = \int_{-\infty}^{\infty} H_{n,k}(x)x^s e^{-x^2}dx = 0, \quad s = 0, 1, \cdots, n-k-1. \tag{7.3.2}$$

For $k = n$ the conclusion will be that $H_{n,n}(x)$ is a constant α_n, so that

$$H_n(x)e^{-x^2} = \alpha_n D^n e^{-x^2};$$

the value $(-1)^n$ for α_n corresponds to leading coefficient 2^n in $H_n(x)$.

PROPERTIES 7.3.4. All n *zeros* of the Hermite polynomial $H_n(x)$ are real and simple. The *norm* of H_n is equal to $2^{\frac{1}{2}n}(n!)^{\frac{1}{2}}\pi^{\frac{1}{4}}$:

$$\|H_n\|^2 = \int_{-\infty}^{\infty} H_n^2(x)e^{-x^2}dx = 2^n\, n!\, \pi^{\frac{1}{2}}.$$

One has the (differential) *recurrence relations*

$$DH_n = H_n' = 2nH_{n-1}, \quad H_{n+1} = (2x - D)H_n,$$

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0. \tag{7.3.3}$$

The polynomial $H_n(x)$ satisfies the *differential equation*

$$y'' - 2xy' + 2ny = 0. \tag{7.3.4}$$

More important than the Hermite polynomials are the Hermite functions:

DEFINITION 7.3.5. The *normalized Hermite functions* are given by

$$h_n(x) \stackrel{\text{def}}{=} \rho_n H_n(x)e^{-\frac{1}{2}x^2}, \quad n \in \mathbb{N}_0, \quad \rho_n = 2^{-\frac{1}{2}n}\,(n!)^{-\frac{1}{2}}\,\pi^{-\frac{1}{4}}. \tag{7.3.5}$$

The Hermite functions form a very important orthonormal basis of $L^2(\mathbb{R})$. A standard proof is based on a moment theorem; see Chapter 9 and cf. Exercise 7.3.10. We observe here that the substitution $H_n(x) = (1/\rho_n)e^{\frac{1}{2}x^2}h_n(x)$ in (7.3.3) gives the following *fundamental relations*:

$$(x + D)h_n = \frac{\rho_n}{\rho_{n-1}}\, 2nh_{n-1} = \sqrt{2n}\, h_{n-1}, \tag{7.3.6}$$

$$(x - D)h_n = \frac{\rho_n}{\rho_{n+1}}\, h_{n+1} = \sqrt{2n+2}\, h_{n+1}. \tag{7.3.7}$$

Applying the second relation with $n+1$ replaced by $n, n-1, \cdots, 1$, we obtain

$$\begin{aligned}\frac{h_n}{\rho_n} &= (x-D)\,\frac{h_{n-1}}{\rho_{n-1}} = (x-D)^2\,\frac{h_{n-2}}{\rho_{n-2}} = \cdots \\ &= (x-D)^n\,\frac{h_0}{\rho_0} = (x-D)^n e^{-\frac{1}{2}x^2}.\end{aligned}$$

Since $D(xh_n) = h_n + xDh_n$, it also follows from (7.3.6), (7.3.7) that

$$\begin{aligned}(x^2-D^2)h_n &= \{(x-D)(x+D)+1\}h_n \\ &= (x-D)\sqrt{2n}\,h_{n-1} + h_n = (2n+1)h_n.\end{aligned}$$

We have thus proved

PROPOSITION 7.3.6. *The normalized Hermite function h_n can be represented by the formula*

$$h_n(x) = \rho_n(x-D)^n e^{-\frac{1}{2}x^2}, \quad n \in \mathbb{N}_0, \tag{7.3.8}$$

where $\rho_n = 2^{-\frac{1}{2}n}\,(n!)^{-\frac{1}{2}}\,\pi^{-\frac{1}{4}}$. The function $y = h_n(x)$ satisfies the differential equation

$$(x^2-D^2)y = (2n+1)y. \tag{7.3.9}$$

Exercises. 7.3.1. Derive the Rodrigues type formula for $H_n(x)$ in Theorem 7.3.2.

Hint. For $k < n$, integration by parts shows that

$$H_{n,k+1}(x)e^{-x^2} = \int_{-\infty}^{x} H_{n,k}(t)e^{-t^2}dt = p(x)e^{-x^2} + c\int_{-\infty}^{x} e^{-t^2}dt,$$

where $p(x)$ is a polynomial of precise degree $n-k-1$ and c is a constant. Now use (7.3.2) for the given k to show that $c = 0$, and then prove that $\big(H_{n,k+1}(x), x^s\big) = 0$ for $s = 0, 1, \cdots, n-k-2$.

7.3.2. Show that the n zeros of $H_n(x)$ are real and distinct.

7.3.3. Prove that

$$\begin{aligned}(H_n, H_n) &= (-1)^n\big(H_{n,n}, H_n^{(n)}\big) \\ &= (-1)^n\alpha_n 2^n n!\int_{-\infty}^{\infty} e^{-x^2}dx = 2^n n!\,\pi^{\frac{1}{2}}.\end{aligned}$$

7.3.4. Prove the relations (7.3.3).

Hint. Show that H_n' is a polynomial of degree $n-1$ which is orthogonal to $1, x, \cdots, x^{n-2}$ in $L^2(\mathbb{R}; e^{-x^2})$, and that $(2x - D)H_n$ is a polynomial of degree $n+1$ which is orthogonal to $1, x, \cdots, x^n$.

7.3.5. Derive the differential equation (7.3.4) for $y = H_n(x)$.

7.3.6. Deduce the relations (7.3.6), (7.3.7) from (7.3.3).

7.3.7. How many relative maxima and minima does $H_n(x)$ have on $\mathbb{R}$? How many does $h_n(x)$ have? Determine the largest value of x for which $h_n(x)$ has an inflection point. Deduce that the last extremum of $h_n(x)$ occurs at a point $x < \sqrt{2n+1}$.

7.3.8. Use the differential equation for h_n to prove that the function

$$v(x) = v_n(x) = h_n(x)^2 + \frac{1}{2n+1-x^2}\, h_n'(x)^2$$

is strictly increasing on the interval $[0, \sqrt{2n+1}]$. Deduce that the successive relative maxima of $|h_n(x)|$ on $[0, \infty)$ form an increasing sequence. Make a rough sketch of the graph for $y = h_n(x)$ on $[0, \infty)$.

7.3.9. The even polynomials $H_0, H_2, H_4, \cdots$ form an orthogonal system in $L^2(0, \infty; e^{-x^2})$. Deduce that $H_{2n}(\sqrt{x})$ is a scalar multiple of the (generalized) Laguerre polynomial $L_n^{(-\frac{1}{2})}(x)$. Likewise, show that $H_{2n+1}(\sqrt{x})/\sqrt{x}$ is a scalar multiple of $L_n^{(\frac{1}{2})}(x)$. [Cf. Exercise 7.2.12.]

7.3.10. Prove that the even Hermite functions $h_{2n}(x)$, $n \in \mathbb{N}_0$, form an orthogonal basis for $L^2(0, \infty)$. [Cf. Exercise 7.2.11.]

7.3.11. (*Modified Hermite polynomials $\tilde{H}_n(x)$ of Mathematical Statistics*) One may obtain the polynomials $\tilde{H}_n(x)$ by orthogonalization of the sequence of powers $\{1, x, x^2, \cdots\}$ in $L^2\left(\mathbb{R}; e^{-\frac{1}{2}x^2}\right)$ – no standardization is necessary. Show that

$$\begin{aligned}
&\tilde{H}_n(x) = (-1)^n e^{\frac{1}{2}x^2} D^n e^{-\frac{1}{2}x^2};\\
&\|\tilde{H}_n\|^2 = \int_{-\infty}^{\infty} \tilde{H}_n(x)^2 e^{-\frac{1}{2}x^2} dx = n!\,(2\pi)^{\frac{1}{2}};\\
&\tilde{H}_{n+1}(x) - x\tilde{H}_n(x) + n\tilde{H}_{n-1}(x) = 0, \quad \tilde{H}_0(x) = 1, \quad \tilde{H}_1(x) = x,\\
&\tilde{H}_2(x) = x^2 - 1, \quad \tilde{H}_3(x) = x^3 - 3x, \quad \tilde{H}_4(x) = x^4 - 6x^2 + 3;\\
&D\tilde{H}_n(x) = n\tilde{H}_{n-1}(x), \quad (x - D)\tilde{H}_n(x) = \tilde{H}_{n+1}(x);\\
&\tilde{H}_n''(x) - x\tilde{H}_n'(x) + n\tilde{H}_n(x) = 0.
\end{aligned}$$

7.4. Integral representations and generating functions

The Legendre polynomials $P_n(x)$. Cauchy's formula for the n^{th} derivative of an analytic function $f(s)$ [see Complex Analysis] has the form

$$D^n f(x) = \frac{n!}{2\pi i}\int_C \frac{f(s)}{(s-x)^{n+1}}\,ds. \tag{7.4.1}$$

Here C may be any positively oriented contour (piecewise smooth Jordan curve) around the point x. Of course, C must be such that f is analytic in the domain Ω enclosed by C and continuous on the closure $\overline{\Omega}$.

Thinking of Rodrigues' formula for $P_n(x)$ [Theorem 7.1.2], we substitute $f(s) = (s^2-1)^n/(2^n n!)$ into (7.4.1) to obtain

Proposition 7.4.1. (*Schläfli's integral*) (after the Swiss mathematician Ludwig Schläfli, 1814-1895; [**107**].) *One has*

$$P_n(x) = \frac{1}{2\pi i}\int_C \frac{(s^2-1)^n}{2^n(s-x)^{n+1}}\,ds.$$

It is natural to take for C a circle with center at the point x. Setting

$$s = x + \rho e^{i\phi}, \quad -\pi \le \phi \le \pi,$$

where ρ is a nonzero constant (which need not be real!), one obtains

$$\begin{aligned} P_n(x) &= \frac{1}{2\pi i}\int_{-\pi}^{\pi} \frac{\left(x^2-1+\rho^2 e^{2i\phi}+2x\rho e^{i\phi}\right)^n}{2^n \rho^{n+1} e^{(n+1)i\phi}}\,\rho e^{i\phi} i\,d\phi \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{x + \frac{(x^2-1)e^{-i\phi}+\rho^2 e^{i\phi}}{2\rho}\right\}^n d\phi. \end{aligned}$$

The formula becomes simpler when one takes $\rho^2 = x^2-1$, a choice which is legitimate whenever $x \ne \pm 1$. The resulting integrand is even in ϕ. One thus obtains, for either choice of the square root $\rho = (x^2-1)^{\frac{1}{2}}$,

Proposition 7.4.2. (*Laplace's integral*) *One has*

$$P_n(x) = \frac{1}{\pi}\int_0^{\pi} \left\{x + (x^2-1)^{\frac{1}{2}}\cos\phi\right\}^n d\phi.$$

By inspection the formula is valid also for $x = \pm 1$, hence it holds for all real and complex x. Taking in particular $-1 \le x \le 1$, the absolute value of the integrand $\left\{x \pm i(1-x^2)^{\frac{1}{2}}\cos\phi\right\}^n$ is equal to

$$\left\{x^2 + (1-x^2)\cos^2\phi\right\}^{\frac{1}{2}n} = \left\{1-(1-x^2)\sin^2\phi\right\}^{\frac{1}{2}n} \le 1.$$

COROLLARY 7.4.3. *For* $-1 \le x \le 1$ *one has* $|P_n(x)| \le 1$.

Careful analysis shows that $|P_n(x)|$ is also bounded by

$$\sqrt{\frac{2}{\pi n(1-x^2)}}$$

on $(-1,1)$; see Szegő [**117**] Theorem 7.3.3, and cf. (8.2.5) below.

We will now derive a generating function for the Legendre polynomials $P_n(x)$.

DEFINITION 7.4.4. Let $\{u_n\}$, $n \in \mathbb{N}_0$, be a sequence of numbers or functions. A *generating function* for the sequence $\{u_n\}$ is an analytic function $g(w)$ with a development of the type $\sum_{n=0}^{\infty} u_n w^n$ or $\sum_{n=0}^{\infty} u_n w^n/n!$.

Using Laplace's integral for $P_n(\cos\theta)$, $\theta \in \mathbb{R}$, and taking $|w| < 1$, we find

$$\begin{aligned} g(w) &\stackrel{\text{def}}{=} \sum_{n=0}^{\infty} P_n(\cos\theta) w^n \\ &= \sum_0^{\infty} \frac{1}{\pi} \int_0^{\pi} (\cos\theta + i\sin\theta\cos\phi)^n d\phi \cdot w^n = \frac{1}{\pi}\int_0^{\pi} \sum_0^{\infty} \cdots\, d\phi \\ &= \frac{1}{\pi}\int_0^{\pi} \frac{1}{1-(\cos\theta + i\sin\theta\cos\phi)w}\, d\phi. \end{aligned} \tag{7.4.2}$$

Here the inversion of the order of summation and integration is justified by uniform convergence relative to ϕ. The final integral is readily evaluated:

LEMMA 7.4.5. *For* $|a| > |b|$ *one has*

$$I = \frac{1}{\pi}\int_0^{\pi} \frac{d\phi}{a + b\cos\phi} = \frac{1}{\sqrt{a^2-b^2}} = \frac{1}{a}\,\text{p.v}\left(1 - \frac{b^2}{a^2}\right)^{-\frac{1}{2}}.$$

PROOF. This is a simple exercise in Complex Function Theory. Setting $e^{i\phi} = z$, one obtains an integral along the positively oriented unit circle $C(0,1)$ [center 0, radius 1]:

$$\begin{aligned} I &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{d\phi}{a+b\cos\phi} = \frac{1}{2\pi}\int_{C(0,1)} \frac{1}{a + \frac{1}{2}b(z+1/z)}\,\frac{dz}{iz} \\ &= \frac{1}{2\pi i}\,\frac{2}{b}\int_{C(0,1)} \frac{1}{z^2 + 2(a/b)z + 1}\,dz \qquad (b \ne 0). \end{aligned}$$

We write the denominator as $(z-z_1)(z-z_2)$, where z_1 and z_2 are the roots inside and outside the unit circle, respectively:

$$z_1 = -(a/b) + (a/b)(1-b^2/a^2)^{\frac{1}{2}}, \quad z_2 = -(a/b) - (a/b)(1-b^2/a^2)^{\frac{1}{2}}.$$

Then the residue theorem gives

$$I = \frac{2}{b} \times (\text{residue of integrand at } z_1) = \frac{2}{b}\,\frac{1}{z_1 - z_2} = \frac{1}{a(1-b^2/a^2)^{\frac{1}{2}}}.$$

□

Returning to (7.4.2), we obtain

$$\begin{aligned} g(w) &= \frac{1}{\left\{(1-w\cos\theta)^2 - (iw\sin\theta)^2\right\}^{\frac{1}{2}}} \\ &= \frac{1}{(1-2w\cos\theta + w^2)^{\frac{1}{2}}} = (1-we^{i\theta})^{-\frac{1}{2}}(1-we^{-i\theta})^{-\frac{1}{2}}. \end{aligned} \tag{7.4.3}$$

The result will certainly be valid for small $|w|$, say $|w| < \frac{1}{2}$. By analytic continuation, it is valid throughout the unit disc $B(0,1)$, provided one takes the analytic branch of the square root $(1-2w\cos\theta+w^2)^{\frac{1}{2}}$ which is equal to 1 for $w=0$. In applications, w is usually real: $w = r \in (-1,1)$ or $w = r \in [0,1)$. In those cases our square root is real and positive.

We thus obtain

Proposition 7.4.6. (*Generating function*) for the Legendre polynomials) *One has*

$$\frac{1}{(1-2r\cos\theta + r^2)^{\frac{1}{2}}} = \sum_{n=1}^{\infty} P_n(\cos\theta) r^n, \quad \theta \in \mathbb{R}, \quad 0 \le r < 1.$$

Many properties of the Legendre polynomials may be derived directly from the generating function; cf. Exercises 7.4.2–7.4.7 and Examples 8.3.1.

The Laguerre polynomials. The Rodrigues formula [Theorem 7.2.4] and the Cauchy formula (7.4.1) immediately give the *integral representation*

$$L_n(x) = \frac{1}{2\pi i} \int_C \frac{s^n e^{x-s}}{(s-x)^{n+1}}\, ds, \tag{7.4.4}$$

where C may be any positively oriented contour around the point x. For a crude bound on $|L_n(x)|$ when $x \neq 0$ we take for C a circle of radius $2|x|$

about the point x:

$$|L_n(x)| \le \frac{1}{2\pi}\,\frac{|2x|^n e^{|x|}}{|x|^{n+1}}\,2\pi|x| = 2^n e^{|x|}.$$

Since $|s/(s-x)| \le 2$ on the circle C, (7.4.4) now shows that for $|w| < 1/2$ [appealing to uniform convergence],

$$\begin{aligned}\sum_{n=0}^{\infty} L_n(x)w^n &= \frac{1}{2\pi i}\int_C \sum_{n=0}^{\infty}\left(\frac{ws}{s-x}\right)^n \frac{e^{x-s}}{s-x}\,ds \\ &= \frac{1}{2\pi i}\int_C \frac{1}{1-\frac{ws}{s-x}}\,\frac{e^{x-s}}{s-x}\,ds = \frac{1}{1-w}\,\frac{1}{2\pi i}\int_C \frac{e^{x-s}}{s-\frac{x}{1-w}}\,ds \\ &= \frac{1}{1-w}\,\exp\left(\frac{-xw}{1-w}\right). \qquad (7.4.5)\end{aligned}$$

In the final step we have used the residue theorem:7. the point $s = x/(1-w)$ lies inside C. The *generating function* in the last member is analytic for $w \ne 1$, hence its power series expansion in the first member must be valid for all $|w| < 1$ and every x.

The Hermite polynomials. The Rodrigues type formula [Theorem 7.3.2] and Cauchy's formula (7.4.1) give the *integral*

$$\frac{H_n(x)}{n!} = \frac{(-1)^n}{2\pi i}\int_{C_x} \frac{e^{x^2-s^2}}{(s-x)^{n+1}}\,ds = \frac{1}{2\pi i}\int_{C_0} \frac{e^{2xw-w^2}}{w^{n+1}}\,dw, \qquad (7.4.6)$$

where C_a stands for a positively oriented contour about the point a. The final member represents the coefficient of w^n in the power series for e^{2xw-w^2}. Thus we immediately obtain the following *generating function:*

$$\sum_{n=0}^{\infty} H_n(x)\,\frac{w^n}{n!} = e^{2xw-w^2}. \qquad (7.4.7)$$

This formula is valid for all w and x.

Exercises. 7.4.1. Use Laplace's integral with $x = \cos\theta$ to show that

$$\left|\frac{d}{d\theta}\,P_n(\cos\theta)\right| \le n, \qquad \left|\frac{d}{d\theta}\left\{\sin\theta\,\frac{d}{d\theta}\,P_n(\cos\theta)\right\}\right| \le n^2.$$

7.4.2. Show that

$$P_n(\cos\theta) = \sum_{k=0}^{n} \gamma_k\gamma_{n-k}\cos(n-2k)\theta, \quad \gamma_k = \frac{1\cdot 3\cdots(2k-1)}{2\cdot 4\cdots(2k)}.$$

Hint. Expand the factors of the generating function, $(1-re^{i\theta})^{-\frac{1}{2}}$ and $(1-re^{-i\theta})^{-\frac{1}{2}}$.

7.4.3. Deduce from Exercise 7.4.2 that

$$|P_n(\cos\theta))| \le P_n(1) = 1, \quad |P_n'(\cos\theta)| \le P_n'(1) = n(n+1)/2.$$

7.4.4. Show that for real or complex x and all sufficiently small $|r|$,

$$\frac{1}{(1-2xr+r^2)^{\frac{1}{2}}} = \sum_{n=0}^{\infty} P_n(x)r^n. \tag{7.4.8}$$

7.4.5. Using (7.4.8) as a definition for $P_n(x)$, show that $P_n(x)$ is a polynomial in x of precise degree n such that $P_n(1) = 1$. Determine $P_0(x)$, $P_1(x)$ and $P_2(x)$ directly from (7.4.8).

7.4.6. Use (7.4.8) and differentiation with respect to r to show that

$$(1-2xr+r^2)\sum_{1}^{\infty} nP_n(x)r^{n-1} - (x-r)\sum_{0}^{\infty} P_n(x)r^n = 0.$$

Then use this result to derive the recurrence relation for the Legendre polynomials [Proposition 7.1.5].

7.4.7. Use a direct computation(!) to show that for $0 < r,\ s < 1$,

$$\int_{-1}^{1} \frac{dx}{(1-2xr+r^2)^{\frac{1}{2}}(1-2xs+s^2)^{\frac{1}{2}}} = \frac{1}{\sqrt{rs}} \log\frac{1+\sqrt{rs}}{1-\sqrt{rs}}.$$

Deduce from this that the coefficients $P_n(x)$ in the expansion (7.4.8) satisfy the orthogonality relations

$$\int_{-1}^{1} P_n(x)P_k(x)dx = \begin{cases} 0 & \text{for } k \ne n, \\ \frac{1}{n+\frac{1}{2}} & \text{for } k = n. \end{cases}$$

7.4.8. (Another generating function for the Legendre polynomials) Use Laplace's integral to show that

$$\begin{aligned}\sum_{n=0}^{\infty} P_n(\cos\theta)\frac{w^n}{n!} &= e^{w\cos\theta}\frac{1}{\pi}\int_0^{\pi} e^{iw\sin\theta\cos\phi}d\phi \\ &= e^{w\cos\theta}J_0(w\sin\theta),\end{aligned}$$

where $J_0(z)$ denotes the Bessel function of order zero. [For the final step, cf. Chapter 12.]

7.4.9. Compute a generating function for the Chebyshev polynomials $T_n(x)$.

7.4.10. Show that for $0 \leq r < 1$,

$$\sum_{n \geq k} P_n^{(k)}(x) r^{n-k} = 1 \cdot 3 \cdots (2k-1)(1 - 2xr + r^2)^{-k-\frac{1}{2}}.$$

7.4.11. Verify the step from the first to the second integral in formula (7.4.6).

7.4.12. Use Exercise 7.2.11 to show that

$$\sum_{n=0}^{\infty} L_n^{(\alpha)}(x) w^n = (1-w)^{-\alpha-1} \exp\left(\frac{-xw}{1-w}\right).$$

CHAPTER 8

Eigenvalue problems related to differential equations

In this chapter we will encounter some of the same orthogonal systems as in Chapter 7, but now differential equations of mathematical physics will play the leading role. The first two sections review the theory of second-order linear differential equations. Following that, we study Sturm–Liouville problems: eigenvalue problems for differential operators. As one application we obtain solutions of Laplace's equation in $\mathbb{R}^3$ – so-called harmonic functions. This study leads to spherical harmonics and Laplace series; cf. also Kellogg [**61**].

8.1. Second order equations. Homogeneous case

We will review some facts that are proved in the Theory of Ordinary Differential Equations; cf. [**21**], [**54**]. In order to apply the results below one has to put the (linear) differential equation into the *standard form*

$$y'' + f_1(x)y' + f_2(x)y = g(x), \quad a < x < b. \tag{8.1.1}$$

It is assumed throughout that f_1, f_2 and g are continuous on (a, b), or at least locally integrable, that is, integrable over every bounded closed subinterval. By a *solution* of equation (8.1.1) is meant an indefinite integral $y = \phi(x)$ of order two [that is, an indefinite integral of an indefinite integral] which satisfies the differential equation almost everywhere on (a, b). In the case of continuous f_1, f_2 and g, the solutions will then be of class $\mathcal{C}^2$, and they will satisfy the equation everywhere on (a, b).

Proposition 8.1.1. *For arbitrary x_0 in (a, b) and arbitrary constants c_0 and c_1, equation* (8.1.1) *has a unique solution on* (a, b) *that satisfies the "initial conditions"*

$$y(x_0) = c_0, \quad y'(x_0) = c_1. \tag{8.1.2}$$

If f_1, f_2 and g are continuous or integrable from the point a on we may also take $x_0 = a$, that is, there is then a unique solution $y = \phi(x)$ on (a, b) such that $\phi(a) = c_0$ [or $\phi(a+) = c_0$] and $\phi'(a) = c_1$ [or $\phi'(a+) = c_1$].

The solutions of the homogeneous equation: (8.1.1) with $g = 0$, will form a two-dimensional linear space. Linearly independent solutions ϕ_1 and ϕ_2 of the homogeneous equation cannot vanish at the same point. [Why not?]

*In the usual proof of Proposition 8.1.1, the initial value problem is converted to a system of Volterra integral equations (after Vito Volterra, Italy, 1860–1940; [**122**]) that may be solved by iteration. Setting $y = y_1$, $y' = y_2$, equation (8.1.1) goes over into the system of differential equations

$$y_1' = y_2, \quad y_2' = -f_2 y_1 - f_1 y_2 + g.$$

Integrating from x_0 to x and using (8.1.2), the initial value problem takes the equivalent form

$$\begin{aligned} y_1(x) &= c_1 + \int_{x_0}^{x} y_2(t)dt, \\ y_2(x) &= c_2 + \int_{x_0}^{x} \{-f_2(t)y_1(t) - f_1(t)y_2(t) + g(t)\}dt. \end{aligned}$$

Beginning with a first approximation for y_1 and y_2 under the integral signs, for example $y_{1,1}(t) \equiv c_1$, $y_{2,1}(t) \equiv c_2$, one determines a second approximation $y_{1,2}(x)$, $y_{2,2}(x)$ from the equations above, etc. One can show that the successive approximations will converge to a solution of the system.

For the standard second-order differential equations of mathematical physics there is a practical alternative, the *power series method*. We will first discuss the *homogeneous* equation

$$y'' + f_1(x)y' + f_2(x)y = 0, \tag{8.1.3}$$

now with *analytic* coefficients f_1 and f_2.

PROPOSITION 8.1.2. *If f_1 and f_2 are analytic on* (a, b) [hence on a complex neighborhood of every point x_0 in (a, b)], *all solutions of* (8.1.3) *are analytic on (a, b), and conversely. If f_1 and f_2 are analytic on the disc $B(x_0, r)$, the solution of the initial value problem* (8.1.3), (8.1.2) *on the interval $x_0 - r < x < x_0 + r$ or on $x_0 \le x < x_0 + r$ is given by a convergent power series of the form*

$$\phi(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n.$$

Here the coefficients $c_2, c_3, c_4, \cdots$ can be determined recursively from c_0 and c_1.

Example 8.1.3. The initial value problem

$$y'' + y = 0, \quad -b < x < b; \quad y(0) = c_0, \;\; y'(0) = c_1,$$

has the solution $\phi(x) = \sum_0^\infty c_n x^n$, where by termwise differentiation,

$$0 = \sum_2^\infty n(n-1)c_n x^{n-2} + \sum_0^\infty c_n x^n = \sum_2^\infty \big\{n(n-1)c_n + c_{n-2}\big\}x^{n-2}.$$

By the uniqueness of power-series representations, one must have

$$n(n-1)c_n + c_{n-2} = 0, \quad \forall\, n \geq 2.$$

It will follow that

$$(2k)!\, c_{2k} = (-1)^k c_0, \quad (2k+1)!\, c_{2k+1} = (-1)^k c_1.$$

It often happens that one or both end points of (a, b) are *singular* points for the differential equation (8.1.3). In the case of a this means that at least one of the functions f_1 and f_2 fails to be integrable from a on. Simple examples are

$$y'' + \frac{1}{x}\, y' + y = 0 \quad \text{on} \;\; (0, \infty) \tag{8.1.4}$$

(*Bessel's equation* of order zero);

$$y'' - \frac{2x}{1-x^2}\, y' + \frac{\lambda}{1-x^2}\, y = 0 \quad \text{on} \;\; (-1, 1) \tag{8.1.5}$$

(*general Legendre equation*).

In practice, a singular end point is frequently a so-called regular singular point:

Definition 8.1.4. The (singular) point a is called a *regular singular point* for the equation (8.1.3) on (a, b) if

$$f_1(x) = \frac{A(x)}{x-a} \quad \text{and} \quad f_2(x) = \frac{B(x)}{(x-a)^2},$$

where $A(x) = a_0 + a_1(x-a) + \cdots$ and $B(x) = b_0 + b_1(x-a) + \cdots$ are analytic in a neighborhood of a. A similar definition holds for the end point b.

Proposition 8.1.5. *Let a be a regular singular point for equation* (8.1.3) *on* (a, b), *with $A(x)$ and $B(x)$ in Definition 8.1.4 analytic in the disc $B(a, r)$.*

Then the equation has one or two solutions on $a < x < a + r$ [or for $0 < |x - a| < r$] *of the form*

$$\phi(x) = \sum_{n=0}^{\infty} c_n (x-a)^{\rho+n} \quad \text{with } c_0 = 1.$$

Here the number ρ *must satisfy the so-called* indicial equation

$$\rho(\rho - 1) + a_0\rho + b_0 = 0. \tag{8.1.6}$$

For at least one root ρ *(of maximal real part), there is a solution* $\phi(x)$ *as indicated; the coefficients* $c_1, c_2, \cdots$ *may be determined recursively.* [There is of course a corresponding result for the end point b.]

Examples 8.1.6. The simplest example is given by the so-called *equidimensional equation*

$$y'' + \frac{a_0}{x}\, y' + \frac{b_0}{x^2}\, y = 0.$$

Here x^ρ is a solution if and only if ρ satisfies the indicial equation. In the case of Bessel's equation (8.1.4), the indicial equation is $\rho^2 = 0$, and its only root is $\rho = 0$. Setting $y = \sum_0^\infty c_n x^n$ with $c_0 = 1$, one obtains the condition

$$\sum_2^\infty n(n-1)c_n x^{n-2} + \sum_1^\infty n c_n x^{n-2} + \sum_2^\infty c_{n-2} x^{n-2} = 0.$$

Hence $c_1 = 0$ and $n^2 c_n + c_{n-2} = 0$ for $n \geq 2$. The solution is the *Bessel function* of order zero,

$$J_0(x) \stackrel{\text{def}}{=} 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2} - \frac{x^6}{2^2 4^2 6^2} + \cdots .$$

In the case of the Legendre equation (8.1.5), both -1 and $+1$ are regular singular points with indicial equation $\rho^2 = 0$.

How do we find a *"second solution"* of equation (8.1.3) if the method of Proposition 8.1.5 gives only one? We could of course expand about a different point. However, if we are interested in the behavior of the second solution near the singular point a, it is preferable to use

Proposition 8.1.7. *Let* $\phi_1(x)$ *be any solution of the homogeneous equation* (8.1.3) *on* (a, b) *different from the zero solution. Then the general solution has the form* $\phi = C_1\phi_1 + C_2\phi_2$*, where*

$$\phi_2(x) = \phi_1(x) \int_{x_2}^{x} \frac{1}{\phi_1^2(t)} \exp\left\{- \int_{x_1}^{t} f_1(s)ds\right\} dt. \tag{8.1.7}$$

Here x_1 and x_2 may be chosen arbitrarily in (a, b).

This result is obtained by substituting $y = z\phi_1$ in equation (8.1.3).

EXAMPLE 8.1.8. In the case of Legendre's equation (8.1.5) with $\lambda = n(n+1)$ one may take $\phi_1(x) = P_n(x)$; cf. Proposition 7.1.6. Setting $x_1 = 0$, one obtains

$$\int_{x_1}^{t} f_1(s)ds = \int_0^t \frac{-2s}{1-s^2}\,ds = \log(1-t^2),$$
$$\phi_2(x) = P_n(x)\int_{x_2}^{x} \frac{1}{P_n^2(t)}\,\frac{1}{1-t^2}\,dt.$$

Taking x_2 very close to 1 and x even closer, the approximation $P_n(t) \approx 1$ on $[x_2, x]$ shows that $\phi_2(x)$ becomes infinite like $-(1/2)\log(1-x)$ as $x \nearrow 1$. Thus the only solutions of (8.1.5) with $\lambda = n(n+1)$ that remain bounded as $x \nearrow 1$ are the scalar multiples of $\phi_1(x) = P_n(x)$.

Exercises. 8.1.1. Compute the even and odd power series solutions $\phi(x) = \phi(x, \lambda) = \sum_0^\infty c_n x^n$ of the general Legendre equation (8.1.5). Show that the radii of convergence are equal to one, unless the series break off. For which values of λ will this happen?

8.1.2. Let $\phi_1 \not\equiv 0$ be a special solution of equation (8.1.3). Set $y = z\phi_1$ to obtain the general solution (8.1.7).

8.1.3. Compute the coefficients in the power series solution $\phi_1(x) = \phi_1(x, \lambda) = \sum_0^\infty c_n(1-x)^n$ of the general Legendre equation (8.1.5) for which $c_0 = 1$. Show that the power series has radius of convergence two unless it breaks off.

8.1.4. (Continuation) How does the "second solution" of (8.1.5) behave near $x = 1$?

8.1.5. Consider the differential equation

$$y'' + \frac{2}{x}\,y' + \left(1 - \frac{2}{x^2}\right) y = 0.$$

Show that the equation has a power series solution $\phi_1(x) = \sum_0^\infty c_n x^n \not\equiv 0$ which converges for all x. Express the general solution ϕ in terms of ϕ_1. Determine the behavior near $x = 0$ of a solution ϕ_2 which is not a scalar multiple of ϕ_1.

8.1.6. Carefully discuss the behavior of the solutions of the differential equation

$$y'' - \frac{1}{x}\,y' + \left(1 + \frac{1}{x^2}\right) y = 0$$

near the point $x = 0$.

8.1.7. Discuss the solutions of *Bessel's equation of order* $\nu\,(\geq 0)$:

$$y'' + \frac{1}{x}\,y' + \left(1 - \frac{\nu^2}{x^2}\right) y = 0, \quad 0 < x < \infty.$$

The solution which behaves like $x^\nu/\{2^\nu\Gamma(\nu+1)\}$ near $x = 0$ is called $J_\nu(x)$. For the complete power series, see Definition 11.7.3 below.

8.1.8. Let f_1, f_2 and g be analytic for $|x| < r$. Prove that the initial value problem (8.1.1), (8.1.2) with $x_0 = 0$ has a unique formal power series solution $\phi(x) = \sum_0^\infty c_n x^n$.

[A power series is called a *formal* solution if termwise differentiation and substitution into the equation make the coefficients of all powers x^n on the left equal to those on the right.]

8.2. Non-homogeneous equation. Asymptotics

We now return to equation (8.1.1), assuming that we know two linearly independent solutions of the corresponding homogeneous equation.

PROPOSITION 8.2.1. *Let ϕ_1 and ϕ_2 be linearly independent solutions of the homogeneous equation* (8.1.3) *on* (a, b). *Then the general solution of the non-homogeneous equation* (8.1.1) *has the form*

$$\phi(x) = C_1\phi_1(x) + C_2\phi_2(x) + \int_{x_0}^{x} \frac{\phi_1(t)\phi_2(x) - \phi_2(t)\phi_1(x)}{\omega(t)}\, g(t)dt,$$

where

$$\omega(x) = \phi_1(x)\phi_2'(x) - \phi_2(x)\phi_1'(x) = \omega(x_1)\exp\left\{-\int_{x_1}^{x} f_1(s)ds\right\}.$$

Here x_0 and x_1 may be chosen arbitrarily in (a, b). The integral represents the special solution $\phi_0(x)$ that satisfies the initial conditions $y(x_0) = y'(x_0) = 0$.

The standard derivation of this result uses the method of "variation of constants". That is, one tries to solve the nonhomogeneous equation by setting $y = z_1\phi_1(x) + z_2\phi_2(x)$, where z_1 and z_2 are unknown functions. The problem is simplified by imposing the additional condition $z_1'\phi_1 + z_2'\phi_2 = 0$.

APPLICATION 8.2.2. In the case of the "model equation"

$$y'' + \nu^2 y = g(x) \quad \text{on} \ \ (a,b), \quad \nu \ \text{ a positive constant,} \tag{8.2.1}$$

one may take $\phi_1(x) = \cos \nu(x - x_0)$, $\phi_2(x) = (1/\nu)\sin \nu(x - x_0)$, for which one has $\omega(x) \equiv 1$. The general solution may then be written as

$$\phi(x) = \phi(x_0)\cos \nu(x - x_0) + \phi'(x_0)\,\frac{\sin \nu(x - x_0)}{\nu} + \int_{x_0}^{x} \frac{\sin \nu(x - t)}{\nu}\, g(t)dt.$$

This result is very important for the *asymptotic study* of oscillatory functions that satisfy certain differential equations, such as Legendre polynomials, Hermite functions, Bessel functions, etc. The aim is to obtain a good approximation for large values of a parameter or variable. It is then necessary to put the appropriate differential equation (8.1.3) into the form (8.2.1). This requires a transformation which *removes* the first-derivative term. Assuming that f_1 can be written as an indefinite integral, the removal will be achieved by setting $y = f \cdot z$ with an appropriate function f, cf. Exercise 8.2.1:

LEMMA 8.2.3. *The substitution*

$$y = \exp\left\{-\frac{1}{2}\int_{x_1}^{x} f_1(s)ds\right\} \cdot z$$

transforms the differential equation (8.1.3) *into the equation*

$$z'' + F_2(x)z = 0, \quad \textit{where} \ \ F_2 = f_2 - (1/4)f_1^2 - (1/2)f_1'. \tag{8.2.2}$$

One would now hope that $F_2(x)$ can be put into the form $\nu^2 - g_1(x)$ with a relatively small function $g_1(x)$, so that the product $g_1(x)z(x)$ may be treated as a perturbation term $g(x)$ on the right-hand side of the equation. As an illustration of the procedure we will obtain an *asymptotic formula* for the Legendre polynomials P_n as $n \to \infty$; see (8.2.5) below.

APPLICATION 8.2.4. Setting $x = \cos\theta$ in the differential equation for $P_n(x)$, one obtains the so-called *polar Legendre equation* for $w = P_n(\cos\theta)$:

$$w'' + (\cot\theta)w' + n(n+1)w = 0, \quad 0 < \theta < \pi; \tag{8.2.3}$$

cf. Proposition 7.2.2. For the removal of the first-derivative term one may set $w = f \cdot y$ with $f(\theta) = \exp\big\{ - (1/2)\int_{\pi/2}^{\theta} \cot s\, ds\big\} = (\sin\theta)^{-\frac{1}{2}}$. It is thus found that $y = (\sin\theta)^{\frac{1}{2}} P_n(\cos\theta)$ satisfies the differential equation

$$\begin{aligned} & y'' + \left\{ n(n+1) - \frac{1}{4}\cot^2\theta - \frac{1}{2}(\cot\theta)' \right\} y \\ &= y'' + \left\{ n(n+1) + \frac{1}{4} + \frac{1}{4\sin^2\theta} \right\} y = 0. \end{aligned}$$

Treating $-(1/4)(\sin\theta)^{-2}y = -(1/4)(\sin\theta)^{-3/2}P_n(\cos\theta)$ as a perturbation $g(\theta)$, we find that $y = (\sin\theta)^{\frac{1}{2}}P_n(\cos\theta)$ satsfies the following equation of the desired type:

$$y'' + (n+1/2)^2 y = -(1/4)(\sin\theta)^{-3/2}P_n(\cos\theta), \quad 0 < \theta < \pi. \tag{8.2.4}$$

Thus, taking n even so that $P_n(0) \neq 0$, $P_n'(0) = 0$, and using $\theta_0 = \pi/2$ as base point, Application 8.2.2 will give the representation

$$\begin{aligned} (\sin\theta)^{\frac{1}{2}} P_n(\cos\theta) &= P_n(0)\cos(n+1/2)(\theta - \pi/2) \\ &+ \frac{1}{n+\frac{1}{2}} \int_{\pi/2}^{\theta} \frac{1}{4}(\sin t)^{-3/2} \sin\{(n+1/2)(x-t)\} P_n(\cos t)\,dt. \end{aligned}$$

We now limit ourselves to a fixed interval $\delta \le \theta \le \pi - \delta$ with $\delta > 0$. Observing that $P_n(0) = \mathcal{O}(n^{-\frac{1}{2}})$ while $|P_n(\cos\theta)| \le 1$, one first obtains a uniform estimate $(\sin\theta)^{\frac{1}{2}}P_n(\cos\theta) = \mathcal{O}(n^{-\frac{1}{2}})$. Introducing this estimate into the integral and observing that $P_n(0) = (-1)^{n/2}(\pi n/2)^{-\frac{1}{2}} + \mathcal{O}(n^{-3/2})$, one may conclude that

$$P_n(\cos\theta) = \sqrt{2}\,(\pi n \sin\theta)^{-\frac{1}{2}} \cos\big\{(n+1/2)\theta - \pi/4\big\} + \mathcal{O}(n^{-3/2}), \tag{8.2.5}$$

uniformly for $\delta \le \theta \le \pi - \delta$. The final result is also true for odd n.

Exercises. 8.2.1. Let ϕ_1 and ϕ_2 be linearly independent solutions of the homogeneous equation (8.1.3). Determine functions z_1 and z_2 such that the combination $y = z_1\phi_1(x) + z_2\phi_2(x)$ satisfies the non-homogeneous equation (8.1.1). Cf. Proposition 8.2.1.

8.2.2. Starting with equation (8.1.3), determine $f(x)$ such that the substitution $y = f(x)z$ leads to a differential equation for z of the form $z'' + F_2(x)z = 0$.

8.2.3. (Continuation) Now complete the proof of Lemma 8.2.3.

8.2.4. Let $Z_\nu(x)$ be any solution on $(0,\infty)$ of Bessel's equation of order $\nu\,(\ge 0)$. Show that $\phi(x) = x^{\frac{1}{2}}Z_\nu(x)$ satisfies a differential equation of the form

$$z'' + z = \frac{\mu}{x^2}\,z, \quad \mu \in \mathbb{R}.$$

Compute μ and determine the general solution of Bessel's equation of order $\nu = 1/2$.

8.2.5. Let $\phi(x)$ be any real solution on $(0, \infty)$ of the differential equation in Exercise 8.2.4.

(i) Prove that $\phi(x)$ satisfies an integral equation of the form

$$\phi(x) = A\sin(x - \alpha) + \int_{x_0}^{x} \frac{\mu}{t^2}\,\phi(t)\sin(x - t)\,dt.$$

(ii) Deduce that $\phi(x)$ remains bounded as $x \to +\infty$.

Hint. Taking $x_0 \geq 2|\mu|$ and $A > 0$, one will have $|\phi(x)| \leq 2A$ on (x_0, ∞). Indeed, the supposition that there would be a [smallest] value $x_1 > x_0$ with $|\phi(x_1)| = 2A$ would lead to a contradiction.

(iii) Show that for $x \to \infty$,

$$\begin{aligned} \phi(x) &= B\sin(x - \beta) - \int_{x}^{\infty} \frac{\mu}{t^2}\,\phi(t)\sin(x - t)\,dt \\ &= B\sin(x - \beta) + \mathcal{O}(1/x). \end{aligned} \tag{8.2.6}$$

(iv) Deduce that the real Bessel functions $Z_\nu(x)$ for $\nu \geq 0$ behave like

$$B\,\frac{\sin(x - \beta)}{x^{\frac{1}{2}}} + \mathcal{O}\left(\frac{1}{x^{\frac{3}{2}}}\right) \quad \text{as } x \to +\infty.$$

*8.2.6. Prove that the formal power series obtained in Exercise 8.1.8 converges for $|x| < r$, so that it represents an actual solution there.

Hint. Because of Propositions 8.2.1, 8.1.7 and Lemma 8.2.3, it will be sufficient to consider the case $g = f_1 = 0$. Setting $f_2(x) = \sum_0^\infty b_n x^n$ one will have $|b_n| \leq AK^n$ for any constant $K > 1/r$. Prove inductively that $|c_n| \leq CK^n$.

*8.2.7. For a proof of Proposition 8.1.5 it is sufficient to discuss the case $a = 0$, $A(x) = 0$; cf. Lemma 8.2.3. Accordingly, consider the equation $y'' + \{B(x)/x^2\}y = 0$, where $B(x)$ is analytic for $|x| < r$. Show that for a formal solution of the form

$$\phi(x) = \sum_0^\infty c_n x^{\rho + n} \quad \text{with } c_0 = 1, \tag{8.2.7}$$

the number ρ must satisfy the indicial equation $\rho(\rho - 1) + b_0 = 0$. One of the roots ρ_1, ρ_2, say ρ_1, must have real part $\geq 1/2$. Prove that the differential equation does have a formal solution (8.2.7) with $\rho = \rho_1$. Show moreover that for $\rho_2 \neq 0, -1/2, -1, \cdots$, there is also a formal solution (8.2.7) with

$\rho = \rho_2$. It may be proved as in Exercise 8.2.6 that the formal series solutions converge for $0 < |x| < r$, hence they represent actual solutions there.

8.3. Sturm–Liouville problems

This is the name given to two-point boundary value problems for second order differential equations of a certain form. Prototype is the eigenvalue problem which occurs in the case of the vibrating string [cf. Exercise 1.3.1]:

EXAMPLE 8.3.1. Determine the values λ ("eigenvalues") for which the two-point boundary value problem

$$-y'' = \lambda y, \quad 0 < x < \pi; \quad y(0) = 0, \;\; y(\pi) = 0, \tag{8.3.1}$$

has nonzero solutions $y = y(x)$ ("eigenfunctions").

The reason for choosing the present form, with *minus* y'' on the left, is that in this way the eigenvalues will be positive. Indeed, for complex $\lambda \neq 0$, the solutions of the differential equation are

$$y = C_1 e^{\sqrt{-\lambda}\,x} + C_2 e^{-\sqrt{-\lambda}\,x} = A\cos\sqrt{\lambda}\,x + B\sin\sqrt{\lambda}\,x.$$

Only the multiples of $\sin\sqrt{\lambda}\,x$ satisfy the first boundary condition. The second boundary condition now requires that $\sin\sqrt{\lambda}\,\pi$ be equal to zero. Since $\sin z = 0$ if and only if $z = n\pi$, the eigenvalues λ of our problem must satisfy the condition $\sqrt{\lambda} = n$, or $\lambda = n^2$, with $\pm n \in \mathbb{N}$. Here $\pm n$ give the same eigenvalue n^2. The value $\lambda = 0$ does not work in eigenvalue problem (8.3.1): a solution $y = A + Bx$ of the differential equation $-y'' = 0$ cannot be an eigenfunction. Thus the characteristic pairs are

$$\lambda = n^2, \quad y = B\sin nx, \quad n = 1, 2, \cdots \quad (B \neq 0).$$

When one speaks of eigenvalues, there must be a *linear operator* L around. In the present example it will be the operator with rule $Ly = -y''$, whose domain D consists of the $\mathcal{C}^2$ functions $y(x)$ on $[0, \pi]$ for which $y(0) = y(\pi) = 0$. This is a so-called *positive operator* in $L^2(0, \pi)$, cf. Exercise 8.3.1:

$$(Ly, y) = -\int_0^\pi y''\,\overline{y} = \int_0^\pi y'\,\overline{y}' \geq 0, \quad \forall\, y \in D.$$

In general we will consider differential equations which, through multiplication by a suitable function, can be [and have been] put into the *standard form*

$$-\{p(x)y'\}' + q(x)y = \lambda w(x)y, \quad a < x < b. \tag{8.3.2}$$

Here p, q and w are to be real-valued, with w positive almost everywhere, and usually p as well. The functions $1/p$, q and w must be locally integrable on (a, b). A *solution* of the differential equation is an indefinite integral ϕ on (a, b), for which $p\phi'$ is also an indefinite integral, and which is such that the differential equation is satisfied almost everywhere.

Regular Sturm–Liouville problems. For the time being we assume that $1/p$, q and w are integrable over the whole interval (a, b). Then the solutions ϕ of the differential equation will be indefinite integrals on the closed interval $[a, b]$ and the same will be true for $p\phi'$. The differential equation (8.3.2) will now have a unique solution for every pair of initial conditions $y(a) = c_0$, $(py')(a) = c_1$, and likewise for conditions $y(b) = d_0$, $(py')(b) = d_1$. Imposing boundary conditions of the form

$$c_0(py')(a) - c_1 y(a) = 0, \quad d_0(py')(b) - d_1 y(b) = 0, \tag{8.3.3}$$

(with c_j, d_j real, and at least one $c_j \neq 0$, at least one $d_j \neq 0$) we speak of a *regular Sturm–Liouville problem.* [References to Sturm and Liouville are given at the end of this section.]

Theorem 8.3.2. *The eigenvalues of a regular Sturm–Liouville problem* (8.3.2), (8.3.3) *are* real*, and it is sufficient to consider real eigenfunctions.* [The other eigenfunctions are just scalar multiples of the real ones.] *Eigenfunctions ϕ_1 and ϕ_2 belonging to different eigenvalues λ_1 and λ_2 are* orthogonal *to each other on (a, b) with respect to the weight function w:*

$$\int_a^b \phi_1 \overline{\phi}_2 w = 0.$$

Proof. Let (λ_1, ϕ_1) and (λ_2, ϕ_2) be arbitrary characteristic pairs (eigenpairs) of our problem. Then

$$-(p\phi_j')' + q\phi_j = \lambda_j w \phi_j \ \text{ a.e. on } \ (a, b), \quad \phi_j \not\equiv 0.$$

Multiplying the relation for ϕ_1 by ϕ_2, the relation for ϕ_2 by ϕ_1 and subtracting, we find

$$(\lambda_1 - \lambda_2) w \phi_1 \phi_2 = -(p\phi_1')'\phi_2 + (p\phi_2')'\phi_1 = \{-p\phi_1'\phi_2 + p\phi_2'\phi_1\}'.$$

Integration over (a, b) [or over $[\alpha, \beta] \subset (a, b)$ and passage to the limit as $\alpha \searrow a$, $\beta \nearrow b$] thus gives

$$(\lambda_1 - \lambda_2) \int_a^b \phi_1 \phi_2 w = \begin{vmatrix} \phi_1(x) & \phi_2(x) \\ p\phi_1'(x) & p\phi_2'(x) \end{vmatrix}_{x=a}^{x=b}. \tag{8.3.4}$$

Now ϕ_1 and ϕ_2 both satisfy the boundary conditions (8.3.3). Hence for $x = a$ the rows of the determinant in (8.3.4) are proportional, and likewise for $x = b$. It follows that the right-hand side of (8.3.4) is equal to zero, so that

$$\int_a^b \phi_1\phi_2 w = 0 \quad \text{whenever} \;\; \lambda_1 \neq \lambda_2. \tag{8.3.5}$$

Conclusions. Suppose for a moment that (λ_1, ϕ_1) is a characteristic pair (eigenpair) with nonreal λ_1. Then $(\lambda_2, \phi_2) = (\overline{\lambda}_1, \overline{\phi}_1)$ is a characteristic pair with $\lambda_2 \neq \lambda_1$. Hence by (8.3.5), $\int_a^b \phi_1\overline{\phi}_1 w = 0$, so that $\phi_1 \equiv 0$ [since $w > 0$ a.e.]. This contradiction proves that all eigenvalues must be real.

Taking λ real, all coefficients in the differential equation (8.3.2) are real, hence the special solution ϕ_0 for which $\phi_0(a) = c_0$, $(p\phi_0')(a) = c_1$ will be real. All solutions ϕ of (8.3.2) that satisfy the first condition (8.3.3) are scalar multiples of ϕ_0. In particular every eigenfunction of our problem is a scalar multiple of a real eigenfunction ϕ_0 [and the eigenvalues have multiplicity one]. Restricting ourselves to real eigenfunctions, formula (8.3.5) expresses the orthogonality of eigenfunctions belonging to different eigenvalues. □

DEFINITION 8.3.3. The *Sturm–Liouville operator* L corresponding to our regular Sturm–Liouville problem has the rule

$$Ly = \frac{1}{w}\big\{ -(py')' + qy\big\}.$$

The domain D consists of all indefinite integrals $y = y(x)$ on $[a, b]$ for which py' is also an indefinite integral, while y and py' satisfy the boundary conditions (8.3.3). Restricting the domain to $D_0 = \big\{u \in D : Lu \in L^2(a, b; w)\big\}$, we obtain a so-called *symmetric operator* L in $L^2(a, b; w)$:

$$\begin{aligned}
(Lu, v) &= \int_a^b \big\{ -(pu')' + qu\big\}\overline{v} \\
&= -\big[pu'\overline{v}\big]_a^b + \int_a^b pu'\overline{v}' + \int_a^b qu\overline{v} \\
&= \begin{vmatrix} u(x) & \overline{v}(x) \\ pu'(x) & p\overline{v}'(x)\end{vmatrix}_{x=a}^{x=b} + \int_a^b u\big\{ -(p\overline{v}')' + q\overline{v}\big\} \\
&= (u, Lv), \quad \forall\, u,\, v \in D_0.
\end{aligned} \tag{8.3.6}$$

SINGULAR STURM–LIOUVILLE PROBLEMS. If not all three functions $1/p$, q, w are integrable from a on, the end point a is said to be singular for the differential equation (8.3.2). In general, there will then be no solution of (8.3.2) that satisfies the initial conditions $y(a) = c_0$, $(py')(a) = c_1$. In practice a is usually a regular singular point and then a standard boundary condition will be

$$y(x) \text{ must have a finite limit as } x \searrow a. \tag{8.3.7}$$

Together with the differential equation this condition will often imply that $(py')(x) \to 0$ as $x \searrow a$. Corresponding remarks apply to b if it is a regular singular point.

Theorem 8.3.2 has an extension to many singular Sturm–Liouville problems. In fact, the basic formula (8.3.4) often remains valid, provided we interpret $\int_a^b$ as the limit of $\int_\alpha^\beta$ as $\alpha \searrow a$ and $\beta \nearrow b$, and similarly for the right-hand side of (8.3.4).

Example 8.3.4. We consider the general Legendre equation (8.1.5) on $(-1, 1)$. Comparing

$$y'' - \frac{2x}{1-x^2}\,y' \quad \text{with} \quad -(py')' = -py'' - p'y',$$

one finds that the standard form (8.3.2) will be obtained through multiplication by $-p$, where $p'/p = -2x/(1-x^2)$. Taking $\log p = \log(1-x^2)$ so that $p = 1 - x^2$, we obtain the standard form

$$-\{(1-x^2)y'\}' = \lambda y \quad \text{on} \;\; (-1, 1). \tag{8.3.8}$$

Since both end points are singular, one imposes the boundary conditions

$$y(x) \text{ must approach finite limits as } x \to \pm 1. \tag{8.3.9}$$

These conditions arise naturally in problems of Potential Theory; see Section 8.4.

The theory of Section 8.1 involving the indicial equation shows that equation (8.3.8) or (8.1.5) has a solution of the form

$$\phi(x) = \phi(x, \lambda) = \sum_{n=0}^{\infty} c_n (1-x)^n, \;\; c_n = c_n(\lambda), \;\; \text{with} \;\; c_0 = 1.$$

This solution will be analytic at least for $|x - 1| < 2$. The "second solution" around the point $x = 1$ will have a logarithmic singularity there, hence the boundary condition at $x = 1$ is satisfied only by the scalar multiples of ϕ. Around the point $x = -1$ the situation is similar; in fact, the "good solution" there is just $\phi(-x, \lambda)$. For λ to be an eigenvalue, $\phi(x, \lambda)$ must be a scalar multiple of $\phi(-x, \lambda)$; if it is, $\phi(x, \lambda)$ will be an eigenfunction.

Formula (8.3.4) will be applicable to eigenfunctions ϕ and then the right-hand side will be zero, since $(1-x^2)\phi'(x) \to 0$ as $x \to \pm 1$. Thus the eigenvalues must be real, and eigenfunctions belonging to different eigenvalues are pairwise orthogonal in $L^2(-1,1)$.

What else can we say about the characteristic pairs? Replacing $1-x$ by t, the differential equation becomes

$$t(2-t)\frac{d^2y}{dt^2} + 2(1-t)\frac{dy}{dt} + \lambda y = 0.$$

Setting $y = \phi(x,\lambda) = \sum_0^\infty c_n t^n$ one readly obtains the recurrence relation

$$c_{n+1} = \frac{n(n+1)-\lambda}{2(n+1)^2}\, c_n, \quad n = 0,1,2,\cdots; \;\; c_0 = 1.$$

Hence the power series for $\phi(x,\lambda)$ has radius of convergence $R = \lim c_n/c_{n+1} = 2$, *unless it breaks off*, in which case it reduces to a polynomial. If λ is an eigenvalue, $\phi(x,\lambda) = c\phi(-x,\lambda)$ must have an analytic extension across the point -1 and, in fact, across all points of the circle $|x-1| = 2$. [Indeed, -1 is the only singular point of the differential equation on that circle.] In this case the series for $\phi(x,\lambda)$ must have radius of convergence $R > 2$, and thus the power series *must* break off. This happens precisely if for some n, $c_{n+1} = 0$ or $\lambda = n(n+1)$. Then $\phi(x,\lambda)$ reduces to a polynomial $p_n(x)$ of exact degree n. We thus obtain the characteristic pairs

$$\lambda = n(n+1), \quad \phi(x) = \phi(x,\lambda) = p_n(x), \quad n = 0,1,2,\cdots.$$

Since the polynomials p_n form an orthogonal system in $L^2(-1,1)$, while $p_n(1) = 1$, they are precisely the Legendre polynomials P_n!

Basis property of eigenfunctions. Taking for granted that the eigenvalues of our Sturm–Liouville problems all have multiplicity one, we may take one eigenfunction ϕ_n to every eigenvalue λ_n to obtain a so-called *complete system of eigenfunctions* $\{\phi_n\}$. For a regular Sturm–Liouville problem such a system will be an orthogonal basis of the space $L^2(a,b;w)$, and the same is true for the most common singular Sturm–Liouville problems. One proves this by considering the inverse $T = L^{-1}$ of the Sturm–Liouville operator L (assuming for simplicity that 0 is not an eigenvalue, so that L is one to one). The operator T has the same eigenfunctions as L and it is an integral operator in $L^2(a,b;w)$ with square-integrable symmetric kernel. By the theory of Hilbert and Schmidt for such integral operators, the eigenfunctions of T (for different eigenvalues) form an orthogonal basis of $L^2(a,b;w)$. [See Functional Analysis.]

History. Sturm–Liouville problems were named after the French mathematicians Charles-François Sturm (1803–1855; [**115**]) and Joseph Liouville (1809–1882; [**82**]); cf. [**116**]. Hilbert–Schmidt integral operators were named after Hilbert and Schmidt, whose names we have met before.

Exercises. 8.3.1. Let L be a positive linear operator in an inner product space V, that is, the domain D and the range R of L belong to V, and $(Lv, v) \geq 0$ for all $v \in D$. Prove that any eigenvalue λ of L must be real and ≥ 0.

8.3.2. Prove that any eigenvalue of a symmetric linear operator L in an inner product space V must be real, and that eigenvectors belonging to different eigenvalues must be orthogonal to each other.

8.3.3. Consider a regular Sturm–Liouville problem (8.3.2), (8.3.3) with $p > 0$, $q \geq \beta$ and $w > 0$ a.e. on (a, b), and with $c_0 c_1 \geq 0$ and $d_0 d_1 \leq 0$. Prove that the eigenvalues λ must be $\geq \beta$.

8.3.4. Consider the (zero order) Bessel eigenvalue problem

$$y'' + \frac{1}{x}\, y + \lambda y = 0, \quad 0 < x < 1; \;\; y(x) \text{ finite at } x = 0, \; y(1) = 0.$$

(i) Show that the solutions of the differential equation that satisfy the boundary condition at 0 have the form $y = CJ_0\left(\sqrt{\lambda}\, x\right)$.

(ii) Prove that all eigenvalues λ are real and > 0, and that eigenfuntions belonging to different eigenvalues are orthogonal to each other on $(0, 1)$ relative to the weight function $w(x) = \cdots$.

(iii) Characterize the eigenvalues and show that they form an infinite sequence $\lambda_n \to \infty$. [Cf. Exercise 8.2.5, part (iv).]

8.3.5. Consider the associated Legendre eigenvalue problem of integral order $k \geq 0$:

$$\begin{aligned} &-\left\{(1 - x^2) y'\right\}' + \frac{k^2}{1 - x^2}\, y = \lambda y, \quad -1 < x < 1; \\ &y(x) \;\text{ finite at }\; x = \pm 1. \end{aligned}$$

(i) Show that the eigenfunctions must have the form $(1 - x)^{\frac{1}{2}k} g(x)$ at $x = 1$ and $(1 + x)^{\frac{1}{2}k} h(x)$ at $x = -1$, where g and h are analytic at $x = 1$, and $x = -1$, respectively.

(ii) Substitute $y = (1 - x^2)^{\frac{1}{2}k} z$ and determine the differential equation for z.

(iii) Show that the z-equation has a solution of the form

$$z = \phi(x) = \phi(x, \lambda) = \sum_{0}^{\infty} c_n (1-x)^n \quad \text{with } \ c_0 = 1.$$

Obtain a recurrence relation for the coefficients $c_n = c_n(\lambda)$, and show that the series for ϕ has radius of convergence 2 unless it breaks off.

(iv) Show that the characteristic pairs of the associated Legendre problem have the form

$$\lambda = (n+k)(n+k+1), \ \ y = c(1-x^2)^{\frac{1}{2}k} p_n(x), \ \ n = 0, 1, 2, \cdots,$$

where $p_n(x)$ is a polynomial of precise degree n.

(v) What orthogonality property do the eigenfunctions have? Relate the polynomials p_{n-k}, $n \geq k$, to certain known polynomials.

8.3.6. Consider the Hermite eigenvalue problem

$$y'' - 2xy' + \lambda y = 0, \ \ -\infty < x < \infty; \ \ |y(x)| \ll e^{x^2} \ \text{ at } \pm\infty.$$

Show that, in general, the differential equation has even and odd power series solutions that grow roughly like e^{x^2} as $x \to \pm\infty$. Prove that substantially smaller solutions exist only if $\lambda = 2n$, $n \in \mathbb{N}_0$. What sort of functions are the eigenfunctions? What orthogonality property do they have? Relate the eigenfunctions to known functions.

8.3.7. The linear *harmonic oscillator* of quantum mechanics leads to the following eigenvalue problem (cf. [**43**]):

$$-y'' + x^2 y = Ey, \ \ -\infty < x < \infty; \ \ \int_{-\infty}^{\infty} |y(x)|^2 dx \ \text{ finite.}$$

[Roughly speaking, $|y(x)|^2 dx$ represents the probability to find the "oscillating particle" in the interval $(x - dx/2, x + dx/2)$. The eigenvalues E correspond to the possible energy levels.]

One expects solutions of the differential equation that behave roughly like $e^{\pm\frac{1}{2}x^2}$ at $\pm\infty$, so that it is reasonable to substitute $y = e^{-\frac{1}{2}x^2} z$. Next use the preceding Exercise to deduce that the eigenvalues are $E = 2n + 1$, $n \in \mathbb{N}_0$. What are the corresponding eigenfunctions?

A more natural treatment of the linear harmonic oscillator will be given in Section 9.7.

8.4. Laplace equation in $\mathbb{R}^3$; polar coordinates

We will explore some connections between Potential Theory in $\mathbb{R}^3$, Legendre polynomials, and associated Legendre functions. A typical problem

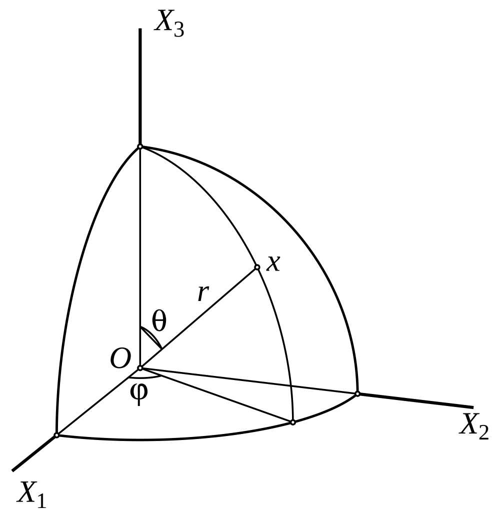

FIGURE 8.1

would be the *Dirichlet problem* for Laplace's equation in the open unit ball $B = B(0,1)$. Here one looks for a solution of Laplace's equation in B, in other words, a *harmonic function*, with prescribed boundary function f on the unit sphere $S = S(0,1)$. An important role is played by *spherical harmonics*: a spherical harmonic of order n is the restriction to S of a homogeneous harmonic polynomial of degree n in x_1, x_2, x_3; cf. Section 8.5.

The Laplace operator Δ_3 occurs in the differential equations for many physical phenomena; cf. [**75**]. Here we need its form in polar coordinates r, θ, ϕ. The latter are given by the relations

$$x_1 = r\sin\theta\cos\phi, \quad x_2 = r\sin\theta\sin\phi, \quad x_3 = r\cos\theta,$$

with $r = \|x\| = |x| \geq 0$, $0 \leq \theta \leq \pi$ and $-\pi < \phi \leq \pi$; cf. Figure 8.1. *Laplace's equation* now becomes

$$\begin{aligned} \Delta_3 u &\equiv \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} \\ &\equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2} = 0; \end{aligned} \tag{8.4.1}$$

cf. Exercises 8.4.1, 8.4.2. In the simpler applications we will have functions u with *axial symmetry*, around the X_3-axis, say. Such functions u are

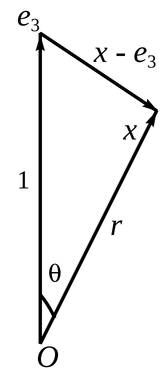

FIGURE 8.2

independent of the angle ϕ. In this case Laplace's equation becomes

$$r^2\Delta_3 u \equiv \frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) = 0. \tag{8.4.2}$$

EXAMPLES 8.4.1. The important solution with *spherical symmetry* about the origin, that is, a solution $u = u(r)$ depending only on $r = |x|$, is $u = 1/r$; cf. Exercise 8.4.3. This is a solution of Laplace's equation in $\mathbb{R}^3 \setminus \{0\}$. Since the Laplacian Δ_3 is translation invariant, $u(x) = 1/|x - a|$ is harmonic in $\mathbb{R}^3 \setminus \{a\}$. In particular $u = u(r,\theta) = 1/|x - e_3|$ [where $e_3 = (0,0,1)$] is harmonic in the unit ball $B(0,1)$. Now

$$\frac{1}{|x - e_3|} = \frac{1}{(1 - 2r\cos\theta + r^2)^{\frac{1}{2}}}$$

(Figure 8.2) is the generating function of the Legendre polynomials $P_n(\cos\theta)$ [Proposition 7.4.6]. It follows that the sum of the series

$$u(r,\theta) = \sum_{n=0}^{\infty} P_n(\cos\theta) r^n$$

satisfies Laplace's equation in B. Here we may apply the operator $r^2\Delta_3$ term by term:

$$\begin{aligned} &r^2\Delta_3 u(r,\theta) \\ &= \sum_{n=0}^{\infty}\left[n(n+1)P_n(\cos\theta) + \frac{1}{\sin\theta}\frac{d}{d\theta}\left\{\sin\theta\frac{d}{d\theta}P_n(\cos\theta)\right\}\right] r^n = 0. \end{aligned}$$

[The differentiated series will be uniformly convergent for $0 \le r \le r_0 < 1$; cf. the estimates in Exercise 7.4.1.] By the uniqueness theorem for power series representations, it follows that the coefficient of r^n must be equal to zero for every $n \in \mathbb{N}_0$:

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left\{\sin\theta\frac{d}{d\theta}P_n(\cos\theta)\right\} = n(n+1)P_n(\cos\theta), \quad 0 < \theta < \pi. \tag{8.4.3}$$

Here we have obtained another derivation of the differential equation for the Legendre polynomials in polar form! [Cf. Proposition 7.2.2.] The corresponding Sturm–Liouville problem or *Legendre eigenvalue problem* will be

$$\begin{gathered} -\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dw}{d\theta}\right) = \lambda w, \quad 0 < \theta < \pi, \quad \text{where} \\ w(\theta) \ \text{ must approach finite limits as } \theta \searrow 0 \text{ and } \theta \nearrow \pi. \end{gathered} \tag{8.4.4}$$

Observe that the boundary conditions are imposed by the geometry: for $\theta = 0$ and for $\theta = \pi$, the point (r, θ) lies on the axis, and there $w(\theta)r^n$ must be continuous. Taking $x = \cos\theta \in [-1, 1]$ as independent variable and setting $w(\theta) = y(x)$, the present Sturm–Liouville problem goes over into the one discussed in Example 8.3.4.

PROPOSITION 8.4.2. *Every harmonic function in B with* axial symmetry *(around the X_3-axis) may be represented by an absolutely convergent series*

$$u(r, \theta) = \sum_{n=0}^{\infty} c_n P_n(\cos\theta) r^n. \tag{8.4.5}$$

PROOF. Since the functions $P_n(\cos\theta)$, $n \in \mathbb{N}_0$, form an orthogonal basis of $L_2(0, \pi; \sin\theta)$, cf. Exercise 7.2.5, every function in that space can be represented by a series $\sum_0^\infty d_n P_n(\cos\theta)$. In particular, for fixed $r \in (0, 1)$, a harmonic function $u(r, \theta)$ in B has the L^2 convergent representation

$$\begin{gathered} u(r, \theta) = \sum_{n=0}^{\infty} v_n(r) P_n(\cos\theta), \quad \text{with} \\ v_n(r) = (n + 1/2)\int_0^\pi u(r, t) P_n(\cos t)\sin t\, dt. \end{gathered} \tag{8.4.6}$$

Since u is a C^∞ function of r, so is v_n, and we may compute the r-derivative of $r^2 dv_n/dr$ by differentiation under the integral sign. Using equation

(8.4.2), it now follows from repeated integration by parts and equation (8.4.3) that $v_n(r)$ must satisfy the equidimensional equation

$$r^2\,\frac{d^2v}{dr^2}+2r\,\frac{dv}{dr}-n(n+1)v=0.$$

The basic solutions are r^n and r^{-n-1}. Since our function $v_n(r)$ must have a finite limit as $r\searrow 0$, we conclude that $v_n(r)=c_nr^n$, where c_n is a constant.

The convergence of the series (8.4.5) for $r=r_1\in(0,1)$ in $L^2(0,\pi;\sin\theta)$ implies absolute and uniform convergence for $0\le r\le r_0<r_1$. Indeed, by Bessel's inequality or the Parseval relation, $|c_n|^2\|P_n\|^2r_1^{2n}\le\|u(r_1,\theta)\|^2$, so that $|c_n|\le A(r_1)\sqrt{(n+1/2)}\,r_1^{-n}$. $\square$

REMARKS 8.4.3. If $u(r,\theta)$ has a continuous boundary function $f(\theta)$ on S, it is plausible that the coefficients c_n must be the Legendre coefficients of f in $L^2(0,\pi;\sin\theta)$. For a proof that the series in Proposition 8.4.2 with these coefficients c_n actually solves the Dirichlet problem, it is best to use the Poisson integral for the unit ball; cf. Section 8.5.

The function $P_n(\cos\theta)r^n$ is harmonic in $\mathbb{R}^n$. It may be expressed as a homogeneous polynomial in x_1,x_2,x_3 of degree n. Indeed, for $0\le k\le n/2$,

$$r^n\cos^{n-2k}\theta=r^{2k}(r\cos\theta)^{n-2k}=\left(x_1^2+x_2^2+x_3^2\right)^kx_3^{n-2k}.$$

Thus $P_n(\cos\theta)$ is a special spherical harmonic of order n.

We now turn to a description of *arbitrary harmonic functions* $u(r,\theta,\phi)$ in B. Since the geometry requires periodicity in the variable ϕ with period 2π, we can expand

$$u(r,\theta,\phi)=\sum_{k\in\mathbb{Z}}u_k(r,\theta)e^{ik\phi},\quad\text{where}$$

$$u_k(r,\theta)=\frac{1}{2\pi}\int_{-\pi}^{\pi}u(r,\theta,\phi)e^{-ik\phi}d\phi. \tag{8.4.7}$$

We observe that for harmonic u, *every term* in the series must be harmonic. One may base a proof on termwise application of the operator $r^2\Delta_3$: the resulting Fourier series must have sum zero. An equivalent procedure is to apply the part of $r^2\Delta_3$ that involves derivatives with respect to r and θ to the integral for u_k. After using equation (8.4.1) under the integral sign, one would apply two integrations by parts. Both methods lead to the equation

$$\frac{\partial}{\partial r}\left(r^2\,\frac{\partial u_k}{\partial r}\right)+\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial u_k}{\partial\theta}\right)=\frac{k^2}{\sin^2\theta}\,u_k, \tag{8.4.8}$$

$0 < r < 1$, $0 < \theta < \pi$.

Just like equation (8.4.2), this equation will have *product solutions* $u_k(r,\theta) = v(r)w(\theta)$. Indeed, substituting such a product and separating variables, (8.4.8) leads to the requirement that

$$\frac{1}{v}\frac{d}{dr}\left(r^2\frac{dv}{dr}\right) = \frac{1}{w}\left\{-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dw}{d\theta}\right) + \frac{k^2}{\sin^2\theta}w\right\}$$

for all (r,θ). Here the left-hand side would be independent of θ, while the right-hand side would be independent of r. Thus for equality the two members must be independent of both r and θ, hence they must be equal to the same constant, which we call λ. For $w(\theta)$ we thus obtain the following Sturm–Liouville problem, the *associated Legendre problem* of order $|k|$ in polar form:

$$\begin{gathered} -\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{dw}{d\theta}\right) + \frac{k^2}{\sin^2\theta}w = \lambda w, \quad 0 < \theta < \pi, \\ w(\theta) \text{ must have finite limits as } \theta \searrow 0 \text{ and } \theta \nearrow \pi. \end{gathered} \tag{8.4.9}$$

In fact, for continuity of $u_k(r,\theta)e^{ik\phi} = v(r)w(\theta)e^{ik\phi}$ on the axis, we must have $w(\theta) \to 0$ as $\theta \searrow 0$ or $\theta \nearrow \pi$ when $k \neq 0$.

Taking $x = \cos\theta \in [-1,1]$ as independent variable and setting $w(\theta) = y(x)$, (8.4.9) becomes the Sturm–Liouville problem of Exercise 8.3.5. By the method described there and by Theorem 7.2.1, the characteristic pairs of (8.4.9) are found to be

$$\lambda = n(n+1), \quad w = cP_n^{|k|}(\cos\theta) = c(\sin\theta)^{|k|}P_n^{(|k|)}(\cos\theta),$$

$n = |k|, |k|+1, \cdots$. A matching function $v(r)$ will be r^n [the differential equation for v will be the same as in the proof of Proposition 8.4.2]. We have thus found infinitely many product solutions

$$u_k(r,\theta) = v(r)w(\theta) = cr^nP_n^{|k|}(\cos\theta), \quad n = |k|, |k|+1, \cdots$$

of equation (8.4.8) that may be used in formula (8.4.7). More generally, one could use superpositions of such product solutions,

$$u_k(r,\theta) = \sum_{n\geq|k|} c_{nk}P_n^{|k|}(\cos\theta)r^n. \tag{8.4.10}$$

Proposition 8.4.4. *If u is a harmonic function in B of the form $u_k(r,\theta)e^{ik\phi}$, the factor $u_k(r,\theta)$ may be represented by a series* (8.4.10) *that converges in $L^2(0,\pi;\sin\theta)$, and converges absolutely.*

The proof is similar to the proof of Proposition 8.4.2 [which is the special case $k=0$]. One has to observe that $u_k(r,\theta)$ must be in $L^2(0,\pi;\sin\theta)$ and that the functions $P_n^{|k|}(\cos\theta)$, $n \geq k$, form an orthogonal basis of that space; cf. Theorem 7.2.1. For the absolute convergence one may use the inequalities

$$\begin{aligned} &|c_{nk}|\,\|P_n^{|k|}\|\,r_1^n \leq \|u_k(r_1,\theta)\| \qquad (r_1<1) \quad \text{and} \\ &\sup|P_n^{|k|}| \leq \sqrt{(n+1/2)}\,\|P_n^{|k|}\|; \end{aligned} \tag{8.4.11}$$

cf. Exercise 8.5.6 below.

THEOREM 8.4.5. *Every harmonic function u in the unit ball B may be represented by absolutely convergent series*

$$\begin{aligned} u(r,\theta,\phi) &= \sum_{k\in\mathbb{Z}} u_k(r,\theta)e^{ik\phi} = \sum_{k\in\mathbb{Z}}\left\{\sum_{n\geq|k|} c_{nk}P_n^{|k|}(\cos\theta)r^n\right\}e^{ik\phi} \\ &= \sum_{n\in\mathbb{N}_0,\,|k|\leq n} c_{nk}P_n^{|k|}(\cos\theta)e^{ik\phi}r^n \\ &= \sum_{n=0}^{\infty}\left\{\sum_{-n\leq k\leq n} c_{nk}P_n^{|k|}(\cos\theta)e^{ik\phi}\right\}r^n. \end{aligned} \tag{8.4.12}$$

PROOF. Ignoring questions of convergence, the expansions are obtained by combining (8.4.7) and (8.4.10). Let us now look at the double series in (8.4.12), the series on the middle line. For $r_1 \in (0,1)$, repeated application of Bessel's inequality shows that

$$\begin{aligned} |c_{nk}|^2\left\|P_n^{|k|}(\cos\theta)\right\|^2\,r_1^{2n} &\leq \int_0^\pi |u_k(r_1,\theta)|^2\sin\theta\,d\theta \\ &\leq \frac{1}{2\pi}\int_0^\pi\left\{\int_{-\pi}^{\pi}|u(r_1,\theta,\phi)|^2 d\phi\right\}\sin\theta\,d\theta. \end{aligned}$$

Thus, using the second part of (8.4.11),

$$\left|c_{nk}P_n^{|k|}(\cos\theta)e^{ik\phi}r^n\right| \leq \sqrt{(n+1/2)}\,C(r_1)\left(\frac{r}{r_1}\right)^n, \quad \forall\,k,\ \forall\,n\geq|k|.$$

It follows that the double series in (8.4.12) is absolutely and uniformly convergent for $0 \leq r \leq r_0 < r_1$. The absolute convergence justifies the various rearrangements in (8.4.12). □

In order to solve the Dirichlet problem for Laplace's equation in the ball B, one would try to make $u(1,\theta,\phi)$ equal to a prescribed function $f(\theta,\phi)$

on $S = \partial B$. Thus one would like to represent $f(\theta, \phi)$ by a double series of the form $\sum c_{nk} P_n^{|k|}(\cos\theta) e^{ik\phi}$; see Section 8.5.

Exercises. 8.4.1. Setting $x = r\cos\phi$, $y = r\sin\phi$, show that

$$\frac{\partial}{\partial x} = \cos\phi \frac{\partial}{\partial r} - \frac{\sin\phi}{r}\frac{\partial}{\partial\phi}, \quad \frac{\partial}{\partial y} = \sin\phi\frac{\partial}{\partial r} + \frac{\cos\phi}{r}\frac{\partial}{\partial\phi}.$$

Deduce that

$$\Delta_2 u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial\phi^2}.$$

8.4.2. Setting $x_1 = s\cos\phi$, $x_2 = s\sin\phi$ while keeping $x_3 = x_3$, and subsequently setting $x_3 = r\cos\theta$, $s = r\sin\theta$, show that

$$\begin{aligned}\Delta_3 u \equiv \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} &= \frac{\partial^2 u}{\partial s^2} + \frac{1}{s}\frac{\partial u}{\partial s} + \frac{1}{s^2}\frac{\partial^2 u}{\partial\phi^2} + \frac{\partial^2 u}{\partial x_3^2} \\ &= \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial\theta^2} + \frac{\cot\theta}{r^2}\frac{\partial u}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2}.\end{aligned}$$

8.4.3. Obtain the general solution of Laplace's equation $\Delta_3 u = 0$ which is spherically symmetric about the origin [so that u depends only on $r = |x|$].

8.4.4. Show that the solutions of Laplace's equation in the unit ball with axial symmetry relative to the X_3-axis are uniquely determined by their values on the interval $0 < x_3 < 1$ of that axis.

8.4.5. Ignoring Proposition 7.4.6, use Exercise 8.4.4 to obtain a series representation for the axially symmetric harmonic function $1/|x - e_3|$ in the unit ball.

8.4.6. Determine all product solutions $u(r, \theta) = v(r)w(\theta)$ of Laplace's equation on $\mathbb{R}^3 \setminus \{0\}$. Single out the solutions that vanish at infinity.

8.4.7. Obtain a formula for the general axially symmetric solution of Laplace's equation in the exterior of the unit sphere that vanishes at infinity.

8.4.8. Given that $u(r, \theta, \phi)$ is a solution of Laplace's equation $\Delta_3 u = 0$ in some domain $\Omega \subset \mathbb{R}^3$, prove that

$$v(r, \theta, \phi) \stackrel{\text{def}}{=} \frac{1}{r}\, u\left(\frac{1}{r}, \theta, \phi\right)$$

is a solution of Laplace's equation $\Delta_3 v = 0$ in the domain Ω', obtained by *inversion* of Ω with respect to the unit sphere.

[Inversion of the point (r, θ, ϕ) gives the point $(1/r, \theta, \phi)$; v is called the Kelvin transform of u. [Named after the British mathematical physicist Lord Kelvin (William Thomson), 1824–1907; [**62**]; cf. [**63**].]

8.5. Spherical harmonics and Laplace series

We introduce the notation

$$W_{nk}(\theta,\phi) = P_n^{|k|}(\cos\theta)e^{ik\phi} = (\sin\theta)^{|k|}P_n^{(|k|)}(\cos\theta)e^{ik\phi}, \tag{8.5.1}$$

$n \in \mathbb{N}_0$, $-n \le k \le n$. One sometimes uses the corresponding real functions, $U_{nk} = \operatorname{Re} W_{nk}$, $0 \le k \le n$, and $V_{nk} = \operatorname{Im} W_{nk}$, $1 \le k \le n$. Functions W_{mj} and W_{nk} with $j \ne k$ are orthogonal to each other in $L^2(-\pi < \phi < \pi)$, while functions W_{mj} and W_{nk} with $m \ne n$ are orthogonal to each other in $L^2(0 < \theta < \pi; \sin\theta)$; cf. Theorem 7.2.1. One readily derives

PROPOSITION 8.5.1. *The functions W_{nk} form an orthogonal system in L^2 on the unit sphere S:*

$$L^2(S) = L^2(0 < \theta < \pi, -\pi < \phi < \pi;\, \sin\theta),$$

with inner product given by

$$(f,g) = \int_S f(\xi)\overline{g(\xi)}\,d\sigma(\xi) = \int_0^\pi \int_{-\pi}^\pi \tilde f(\theta,\phi)\,\overline{\tilde g(\theta,\phi)}\sin\theta\,d\theta\,d\phi.$$

Here $\xi = (\xi_1,\xi_2,\xi_3)$ stands for a unit vector, or a point of S; $\xi_1 = \sin\theta\cos\phi$, $\xi_2 = \sin\theta\sin\phi$, $\xi_3 = \cos\theta$. The area element $d\sigma(\xi)$ of S has the form $\sin\theta\,d\theta\,d\phi$. Finally

$$\tilde f(\theta,\phi) = f(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta),$$

and similarly for $\tilde g(\theta,\phi)$. In practice we will carelessly write $f(\theta,\phi)$ for $\tilde f(\theta,\phi)$.

PROOF OF THE PROPOSITION. By Fubini's theorem,

$$\begin{aligned}\int_0^\pi \int_{-\pi}^\pi & P_m^{|j|}(\cos\theta)e^{ij\phi}P_n^{|k|}(\cos\theta)e^{-ik\phi}\sin\theta\,d\theta\,d\phi \\ &= \int_0^\pi P_m^{|j|}(\cos\theta)P_n^{|k|}(\cos\theta)\sin\theta\,d\theta \int_{-\pi}^\pi e^{i(j-k)\phi}d\phi.\end{aligned}$$

Taking $(m,j) \ne (n,k)$, the answer will be zero if $j \ne k$ and also if $j = k$, since in the latter case, necessarily $m \ne n$. □

Observe that for fixed r, the double series in formula (8.4.12) is just the expansion of $u(r,\theta,\phi)$ with respect to the orthogonal system $\{W_{nk}\}$.

PROPOSITION 8.5.2. *The products*

$$\begin{aligned} u &= r^n W_{nk}(\theta,\phi) = r^n P_n^{|k|}(\cos\theta)e^{ik\phi} \\ &= r^n(\sin\theta)^{|k|}P_n^{(|k|)}(\cos\theta)\left(e^{\pm i\phi}\right)^{|k|}, \quad |k|\le n, \end{aligned} \tag{8.5.2}$$

can be written as homogeneous harmonic polynomials in x_1, x_2, x_3 of precise degree n.

Indeed, by Section 8.4, every product (8.5.2) satisfies Laplace's equation $\Delta_3 u = 0$; cf. equation (8.4.8) and Proposition 7.2.2. Observe now that $P_n^{(|k|)}(\cos\theta)$ with $|k|\le n$ is a polynomial in $\cos\theta$ of degree $n-|k|$. Next expanding $\left(e^{\pm i\phi}\right)^{|k|} = (\cos\phi \pm i\sin\phi)^{|k|}$, one finds that $r^n W_{nk}(\theta,\phi)$ can be represented as a sum of terms

$$\begin{aligned} &r^n(\sin\theta)^{|k|}(\cos\theta)^{n-|k|-2l}(\cos\phi)^{|k|-m}(\sin\phi)^m \\ &= r^{|k|-m}(\sin\theta)^{|k|-m}(\cos\phi)^{|k|-m}\cdot r^m(\sin\theta)^m(\sin\phi)^m \\ &\qquad \times r^{n-|k|-2l}(\cos\theta)^{n-|k|-2l}\cdot r^{2l} \\ &= x_1^{|k|-m}x_2^m x_3^{n-|k|-2l}\left(x_1^2+x_2^2+x_3^2\right)^l, \end{aligned}$$

with $|k|+2l\le n$ and $m\le |k|$. Thus $r^n W_{nk}(\theta,\phi)$ is equal to a homogeneous harmonic polynomial of degree n. It follows that $W_{nk}(\theta,\phi)$ is a spherical harmonic of order n:

DEFINITION 8.5.3. A *spherical harmonic* $Y_n = Y_n(\theta,\phi)$ of order n is the restriction to the unit sphere of a homogeneous harmonic polynomial of degree n in x_1, x_2, x_3. Cf. [**113**].

Examples of such harmonic polynomials are:

$$\text{degree } 0: \;\; 1; \quad \text{degree } 1: \;\; x_1,\ x_2,\ x_3;$$
$$\text{degree } 2: \;\; x_1^2-x_2^2,\ x_1x_2,\ x_2^2-x_3^2,\ x_1x_3,\ x_2x_3.$$

The relation $Y_n \leftrightarrow r^nY_n$ establishes a one to one correspondence between spherical harmonics of order n and homogeneous harmonic polynomials of degree n.

PROPOSITION 8.5.4. *The linear space $\mathcal{H}_n$ of the spherical harmonics of order n is rotation invariant and has dimension $2n+1$. The functions $W_{nk}(\theta,\phi)$, $-n\le k\le n$, form an orthogonal basis of $\mathcal{H}_n$. One has*

$$\|W_{nk}\|^2_{L^2(s)} = \int_S |W_{nk}(\xi)|^2 d\sigma(\xi) = \frac{(n+|k|)!}{(n-|k|)!}\,\frac{2\pi}{n+\frac12}.$$

PROOF. The Laplacian Δ_3 is rotation invariant [cf. Exercise 8.5.1], and so is the class of homogeneous polynomials of degree n. It follows that the linear space $\mathcal{K}_n$ of the homogeneous harmonic polynomials of degree n is rotation invariant, hence so is $\mathcal{H}_n$.

We will determine $\dim \mathcal{H}_n$ from $\dim \mathcal{K}_n$. Let

$$U(x_1, x_2, x_3) = \sum_{j+k\le n} \alpha_{jk}\, x_1^j x_2^k x_3^{n-j-k} \quad (\text{with } \ j, k \ge 0)$$

be an element of $\mathcal{K}_n$, that is,

$$\begin{aligned} 0 = \Delta_3 U = \sum_{j+k\le n} \Big\{ & j(j-1)\alpha_{jk} x_1^{j-2} x_2^k x_3^{n-j-k} + k(k-1)\alpha_{jk} x_1^j x_2^{k-2} x_3^{n-j-k} \\ & + (n-j-k)(n-j-k-1)\alpha_{jk} x_1^j x_2^k x_3^{n-j-k-2} \Big\} \\ = \sum_{j+k\le n-2} \{ & (j+2)(j+1)\alpha_{j+2,k} + (k+2)(k+1)\alpha_{j,k+2} \\ & + (n-j-k)(n-j-k-1)\alpha_{jk} \}\, x_1^j x_2^k x_3^{n-2-j-k}. \end{aligned}$$

In the final polynomial all coefficients must be equal to zero. Hence every coefficient α_{jk} with $j+k \le n-2$ can be expressed linearly in terms of $\alpha_{j+2,k}$ and $\alpha_{j,k+2}$. Continuing, we conclude that every α_{jk} with $j+k \le n-2$ can be expressed as a linear combination of coefficients α_{pq} with $p+q$ equal to n or $n-1$. The latter are the coefficients of products which either contain no factor x_3, or just one. These products are

$$x_1^n, x_1^{n-1}x_2, \cdots, x_1 x_2^{n-1}, x_2^n;\; x_1^{n-1}x_3, x_1^{n-2}x_2x_3, \cdots, x_1x_2^{n-2}x_3, x_2^{n-1}x_3;$$

there are $2n-1$ of them.

The $2n+1$ coefficients α_{pq} in U with $p+q \ge n-2$ can actually be selected arbitrarily. Indeed, when they are given, one can determine exactly one set of coefficients α_{jk} with $j+k \le n-2$ such that all coefficients in $\Delta_3 U$ become equal to zero; cf. Figure 8.3. If follows that $\dim \mathcal{K}_n = 2n+1$ and hence also $\dim \mathcal{H}_n = 2n+1$.

The $2n+1$ pairwise orthogonal elements W_{nk}, $-n \le k \le n$ of $\mathcal{H}_n$ must form an orthogonal basis. Finally, by Theorem 7.2.1,

$$\begin{aligned} \|W_{nk}\|^2_{L^2(s)} &= \int_0^\pi \int_{-\pi}^\pi \left| P_n^{|k|}(\cos\theta) e^{ik\phi} \right|^2 \sin\theta \, d\theta \, d\phi \\ &= \frac{(n+|k|)!}{(n-|k|)!} \frac{1}{n+\frac{1}{2}} 2\pi. \end{aligned}$$

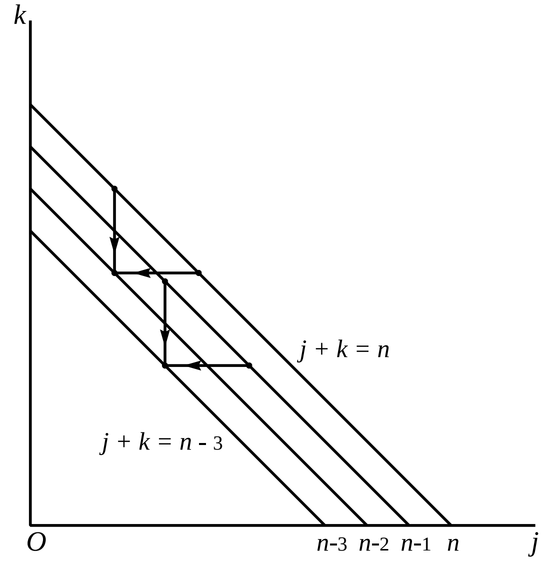

FIGURE 8.3

□

THEOREM 8.5.5. *The spherical harmonics W_{nk}, $n \in \mathbb{N}_0$, $-n \leq k \leq n$, form an orthogonal basis of $L^2(S)$. Every function $f \in L^2(S)$ has a unique representation as an L^2 convergent so-called* Laplace series*,*

$$f(\xi) = \sum_{n=0}^{\infty} Y_n[f](\xi) = \sum_{n=0}^{\infty} \sum_{-n\leq k\leq n} c_{nk} W_{nk}(\theta, \phi),$$

$\xi = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$*. Here $Y_n(\xi) = \sum_{-n\leq k\leq n} c_{nk} W_{nk}(\theta, \phi)$ represents the orthogonal projection of f onto the subspace $\mathcal{H}_n$ of the spherical harmonics of order n. One has the rotation invariant direct sum decomposition*

$$L^2(S) = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_n \oplus \cdots .$$

PROOF. By Proposition 8.5.1, the functions W_{nk} are pairwise orthogonal. We will show that they form a *maximal* orthogonal system in $L^2(S)$, hence, an orthogonal basis. To that end, suppose that $g \in L^2(S)$ is orthogonal to all functions W_{nk}. Then by Fubini's theorem,

$$\int_0^{\pi} \left\{ \int_{-\pi}^{\pi} g(\theta, \phi) e^{-ik\phi} d\phi \right\} P_n^{|k|}(\cos\theta) \sin\theta \, d\theta = 0 \tag{8.5.3}$$

for all $k \in \mathbb{Z}$ and all $n \geq |k|$. Also by Fubini's theorem, the finiteness of the integral

$$\int_0^\pi \int_{-\pi}^\pi |g(\theta, \phi)|^2 \sin\theta \, d\theta \, d\phi = \|g\|^2_{L^2(S)}$$

implies that $G(\theta) = \int_{-\pi}^\pi |g(\theta, \phi)|^2 d\phi$ exists (and is finite) for all θ's outside a set E of measure zero, and that $G(\theta) \in L^1(0, \pi; \sin\theta)$. It follows that $g(\theta, \phi)$ is in L^2 as a function of ϕ on $(-\pi, \pi)$ for $\theta \in (0, \pi) \setminus E$. Furthermore, by Cauchy–Schwarz, every function

$$g_k(\theta) = \frac{1}{2\pi} \int_{-\pi}^\pi g(\theta, \phi) e^{-ik\phi} d\phi$$

will be in $L^2(0, \pi; \sin\theta)$. Thus by (8.5.3) and the orthogonal basis property of the functions $P_n^{|k|}(\cos\theta)$, $n \geq |k|$ [Theorem 7.2.1], $g_k(\theta) = 0$ for all $\theta \in (0, \pi)$ outside some set E_k of measure zero. We now take $\theta \in (0, \pi)$ outside the union E^* of E and the sets E_k, $k \in \mathbb{Z}$, which is still a set of measure zero. Then all Fourier coefficients $g_k(\theta)$ of $g(\theta, \phi)$ are equal to zero, hence by Parseval's formula, $G(\theta) = \int_{-\pi}^\pi |g(\theta, \phi)|^2 d\phi = 0$, $\forall\, \theta \in (0, \pi) \setminus E^*$. Integration with respect to θ now shows that

$$\|g\|^2_{L^2(S)} = \int_0^\pi G(\theta) \sin\theta \, d\theta = 0.$$

Hence $g = 0$ in $L^2(S)$, so that the functions W_{nk} indeed form an orthogonal basis of $L^2(S)$.

Every function $f \in L^2(S)$ thus has a unique L^2 convergent representation $\sum_{n,k} c_{nk} W_{nk}$; here the numbers c_{nk} are simply the expansion coefficients of f with respect to the orthogonal basis $\{W_{nk}\}$. Combining the spherical harmonics of order n: $c_{nk} W_{nk}(\theta, \phi)$, $-n \leq k \leq n$, into a single term $Y_n[f]$, we obtain another orthogonal series with L^2 sum f, the Laplace series $\sum_0^\infty Y_n[f]$:

$$\left\| f - \sum_{n=0}^p Y_n[f] \right\|^2 = \sum_{n>p} \sum_{k=-n}^n |c_{nk}|^2 \|W_{nk}\|^2 \to 0 \quad \text{as } p \to \infty.$$

It is clear that f can have only one decomposition $\sum_0^\infty Y_n$ into spherical harmonics of different order: Y_n must be the orthogonal projection of f onto the subspace $\mathcal{H}_n$. Thus $L^2(S)$ is the direct sum of the subspaces $\mathcal{H}_n$. $\square$

There is an important integral representation for $Y_n[f]$:

PROPOSITION 8.5.6. *The orthogonal projection of f onto $\mathcal{H}_n$ may be written as*

$$Y_n[f](\xi) = \frac{n+\frac{1}{2}}{2\pi}\int_S f(\zeta)P_n(\xi\cdot\zeta)d\sigma(\zeta), \quad \xi \in S. \tag{8.5.4}$$

PROOF. Because the subspace $\mathcal{H}_n$ is rotation invariant, the orthogonal projection $Y_n[f]$ is independent of the choice of a rectangular coordinate system in $\mathbb{R}^3\,(=\mathbb{E}^3)$. In order to evaluate $Y_n[f](\xi)$, we temporarily choose our coordinate system such that $\xi = e_3$, in other words, ξ corresponds to $\theta = 0$. Observe now that all spherical harmonics $W_{nk}(\theta,\phi)$ with $k \neq 0$ vanish at the point $\theta = 0$, while $W_{n0}(\theta,\phi) \equiv P_n(1) = 1$. Hence

$$\begin{aligned} Y_n[f](\xi) &= \sum_{-n\le k\le n} c_{nk}W_{nk}(0,\phi) = c_{n0} = \frac{(f,W_{n0})}{(W_{n0},W_{n0})} \\ &= \frac{n+\frac{1}{2}}{2\pi}\int_0^\pi\int_{-\pi}^\pi \tilde{f}(\theta,\phi)P_n(\cos\theta)\sin\theta\,d\theta\,d\phi. \end{aligned}$$

We finally put the integral into a form independent of the coordinate system. To this end we replace the running point (θ,ϕ) in the integrand by $\zeta = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ on S and $\tilde{f}(\theta,\phi)$ by $f(\zeta)$. Observing that θ is the angle between the vectors $\xi = e_3$ and ζ, so that $\cos\theta = \xi\cdot\zeta$, one obtains formula (8.5.4). □

For the applications it is important to consider Abel summability of Laplace series.

PROPOSITION 8.5.7. *For integrable f on the unit sphere S, the* Abel mean *of the Laplace series,*

$$A_r[f](\xi) \stackrel{\text{def}}{=} \sum_{n=0}^\infty Y_n[f]r^n, \quad 0 \le r < 1,$$

where $Y_n[f]$ is given by (8.5.4)*, is equal to the* Poisson integral *of f for the unit ball B, cf.* [**94**]*:*

$$P[f](r\xi) \stackrel{\text{def}}{=} \int_S f(\zeta)\,\frac{1-r^2}{4\pi(1-2r\xi\cdot\zeta+r^2)^{\frac{3}{2}}}\,d\sigma(\zeta).$$

PROOF. Substituting the integral for $Y_n[f]$ into the definition of $A_r[f](\xi)$ and inverting the order of summation and integration, one obtains

$$A_r[f](\xi) = \int_S f(\zeta)\sum_{n=0}^\infty \frac{n+\frac{1}{2}}{2\pi}\,P_n(\xi\cdot\zeta)\,r^n\,d\sigma(\zeta).$$

The series in the integrand is readily summed with the aid of the generating function for the Legendre polynomials: for $0 \le r < 1$,

$$\begin{aligned} \sum_0^\infty (n+1/2)P_n(\cos\theta)r^n &= r^{\frac12}\frac{\partial}{\partial r}\sum_0^\infty P_n(\cos\theta)r^{n+\frac12} \\ &= r^{\frac12}\frac{\partial}{\partial r}\frac{r^{\frac12}}{(1-2r\cos\theta+r^2)^{\frac12}} = \frac12\frac{1-r^2}{(1-2r\cos\theta+r^2)^{\frac32}}. \end{aligned} \tag{8.5.5}$$

□

Corollary 8.5.8. *For continuous f on S, the Laplace series is uniformly Abel summable to f. The solution of the Dirichlet problem for Laplace's equation in the unit ball, with boundary function f, is given by the Abel mean $A_r[f](\xi)$ of the Laplace series, or equivalently, by the Possion integral $P[f](r\xi)$.*

The proof is similar to the proof of the corresponding result for Fourier series and the Dirichlet problem in the unit disc; cf. Section 3.6.

Exercises. 8.5.1. Prove that the Laplacian Δ_3 is invariant under rotations about the origin: if $x = Py$ with an orthogonal matrix P, then $\sum_j \partial^2/(\partial x_j^2)$ is equal to $\sum_k \partial^2/(\partial y_k^2)$.

8.5.2. Prove that the spherical harmonics $Y = Y_n(\theta,\phi)$ of order n are solutions of the boundary value problem

$$-\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2} = n(n+1)Y,$$
$$0<\theta<\pi,\ -\pi<\phi<\pi,$$

$Y(\theta,\phi)$ and $(\partial Y/\partial\theta)(\theta,\phi)$ remain bounded as
$\theta \searrow 0$ and $\theta \nearrow \pi$,
for fixed θ, $Y(\theta,\phi)$ can be extended to a $\mathcal{C}^1$ function
of ϕ of period 2π.

8.5.3. Use Exercise 8.5.2 to show that spherical harmonics Y_n and Y_p of different order are orthogonal to each other in $L^2(S)$.

8.5.4. Let $p(x)$ be any polynomial in x_1, x_2, x_3 of degree $\le n$. Prove that on the unit sphere S, $p(x)$ is equal to a harmonic polynomial of degree $\le n$.

Hint. One may use the Laplace series for $p(\xi)$, or prove directly that $p(x)$ is congruent to a harmonic polynomial of degree $\le n$, modulo the polynomial $(x_1^2+x_2^2+x_3^2-1)$.

8.5.5. Show that every spherical harmonic Y_n of order n satsfies the integral equation

$$Y_n(\xi) = \frac{n+\frac{1}{2}}{2\pi}\int_S Y_n(\zeta)P_n(\xi\cdot\zeta)d\sigma(\zeta), \quad \xi \in S.$$

8.5.6. Use Exercise 8.5.5 to prove the inequalities

$$\sup_{\xi\in S}|Y_n(\xi)| \le \sqrt{\frac{n+\frac{1}{2}}{2\pi}}\,\|Y_n\|_{L^2(s)},$$

$$\sup_{x\in[-1,1]}\left|P_n^{|k|}(x)\right| \le \sqrt{(n+1/2)}\,\left\|P_n^{|k|}\right\|_{L^2(-1,1)}.$$

8.5.7. Compute the coefficients c_{nk} in the Laplace series for f [Theorem 8.5.5] to show that

$$Y_n[f](\theta,\phi) = \sum_{k=-n}^{n} c_{nk}W_{nk}(\theta,\phi)$$

can be evaluated as

$$\frac{n+\frac{1}{2}}{2\pi}\int_0^{\pi}\int_{-\pi}^{\pi} f(\tilde\theta,\tilde\phi)\sum_{k=-n}^{n}\frac{(n-|k|)!}{(n+|k|)!}P_n^{|k|}(\cos\theta)P_n^{|k|}(\cos\tilde\theta)e^{ik(\phi-\tilde\phi)}\sin\tilde\theta\,d\tilde\theta\,d\tilde\phi.$$

8.5.8. Compare formula (8.5.4) and Exercise 8.5.7 to derive the so-called *Addition Theorem* for spherical harmonics:

$$P_n(\cos\gamma) = \sum_{k=-n}^{n}\frac{(n-|k|)!}{(n+|k|)!}P_n^{|k|}(\cos\theta)P_n^{|k|}(\cos\tilde\theta)e^{ik(\phi-\tilde\phi)}.$$

Here γ is the angle between the vectors $\xi = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ and $\zeta = (\sin\tilde\theta\cos\tilde\phi, \sin\tilde\theta\sin\tilde\phi, \cos\tilde\theta)$.

[For $n = 1$ one obtains an old formula of spherical trigonometry:

$$\cos\gamma = \xi\cdot\zeta = \cos\theta\cos\tilde\theta + \sin\theta\sin\tilde\theta\cos(\phi-\tilde\phi).]$$

8.5.9. Prove Corollary 8.5.8.

CHAPTER 9

Fourier transformation of well-behaved functions

An introduction to the theory of Fourier integrals was given in Section 1.7. There it was made plausible that under reasonable conditions, one has a *Fourier inversion theorem* as follows:

"If g is the Fourier transform of f, then

(9.0.1) $\quad f$ is equal to $\dfrac{1}{2\pi}$ times the reflected Fourier transform of g".

In this and the next chapters we will obtain precise conditions for Fourier inversion.

Another important fact about Fourier transformation is the following:

"Under Fourier transformation, *differentiation* goes over into

(9.0.2) $\quad$ *multiplication* by i times the (new) independent variable".

This property makes Fourier transformation very useful for solving certain ordinary and partial differential equations.

9.1. Fourier transformation on $L^1(\mathbb{R})$

Let f be integrable over $\mathbb{R}$ in the sense of Lebesgue, so that $|f|$ is also integrable over $\mathbb{R}$. A sufficient condition would be that the improper Riemann integrals of f and $|f|$ over $\mathbb{R}$ exist. The product $f(x)e^{-i\xi x}$ with $\xi \in \mathbb{R}$ will also be integrable over $\mathbb{R}$ since $e^{-i\xi x}$ is continuous and bounded.

Definitions 9.1.1. For f in $\mathcal{L}^1(\mathbb{R})$ [that is, for integrable f on $\mathbb{R}$], the *Fourier transform* $g = \mathcal{F}f = \widehat{f}$ is the function on $\mathbb{R}$ given by

$$g(\xi) = (\mathcal{F}f)(\xi) = \widehat{f}(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x)e^{-i\xi x}dx, \quad \xi \in \mathbb{R}.$$

Equivalently, using independent variable x for the transform,

$$(\mathcal{F}f)(x) = \int_{\mathbb{R}} f(t)e^{-ixt}dt, \quad x \in \mathbb{R}.$$

The *reflected Fourier transform* $h = \mathcal{F}_R f = \check{f}$ is simply the *reflection* $g_R(\xi) = g(-\xi)$ of the Fourier transform:

$$h(\xi) = (\mathcal{F}_R f)(\xi) = \check{f}(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x)e^{i\xi x}dx = g(-\xi) = g_R(\xi), \quad \xi \in \mathbb{R}.$$

Using independent variable x for the reflected transform, one has

$$\begin{aligned}(\mathcal{F}_R f)(x) &= \int_{-\infty}^{\infty} f(t)e^{ixt}dt = (\mathcal{F}f)(-x)\\ &= \int_{-\infty}^{\infty} f(-s)e^{-ixs}ds = (\mathcal{F}f_R)(x), \quad x \in \mathbb{R}.\end{aligned}$$

The *reflected* Fourier transform of f is also the *Fourier transform* of the reflection $f_R(x) = f(-x)$.

REMARKS 9.1.2. We use $e^{-i\xi x}$ [with a minus sign in the exponent] in the definition of $\widehat{f}(\xi)$; cf. the formula for Fourier coefficients. Some authors interchange the definitions of $\mathcal{F}f$ and $\mathcal{F}_R f$, but this has little effect on the theory. One sometimes puts a factor $1/\sqrt{2\pi}$ in front of the integrals for the Fourier transform and the reflected Fourier transform. Such normalization would give a more symmetric form to the Inversion theorem; cf. Section 9.2.

Integrable functions that differ only on a set of measure zero will have the same Fourier transform. Thus Fourier transformation on $\mathcal{L}^1(\mathbb{R})$ may be considered as a transformation on the normed space $L^1(\mathbb{R})$; cf. Examples 5.3.4.

EXAMPLES 9.1.3. The computations in Example 1.7.1 give the following Fourier pairs (where $a > 0$):

$f(x)$	$\widehat{f}(\xi)$	$\widehat{f}(x)$
$e^{-a\lvert x\rvert}$	$\dfrac{2a}{\xi^2+a^2}$	$\dfrac{2a}{x^2+a^2}$
$\dfrac{1}{2\pi}\,\dfrac{2a}{x^2+a^2}$	$e^{-a\lvert\xi\rvert}$	$e^{-a\lvert x\rvert}$
$e^{-ax}U(x) = \begin{cases} e^{-ax}, & x>0 \\ 0, & x<0 \end{cases}$	$\dfrac{1}{a+i\xi}$	$\dfrac{1}{a+ix}$

EXAMPLE 9.1.4. We will use Complex Analysis to obtain the useful pair

$$f(x) = e^{-ax^2}, \qquad \widehat{f}(\xi) = \sqrt{(\pi/a)}\, e^{-\xi^2/(4a)} \qquad (a > 0).$$

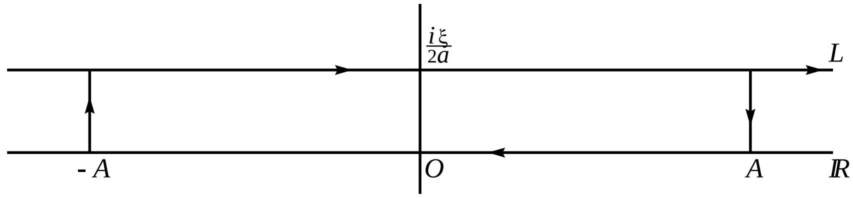

FIGURE 9.1

In particular for $a = 1/2$:

$$f(x) = e^{-\frac{1}{2}x^2}, \qquad \widehat{f}(x) = \sqrt{2\pi}\, e^{-\frac{1}{2}x^2}.$$

Thus the function $e^{-\frac{1}{2}x^2}$ is an *eigenfunction* of Fourier transformation. If one would define the Fourier transform with the normalizing factor $1/\sqrt{2\pi}$, the function $e^{-\frac{1}{2}x^2}$ would be invariant under Fourier transformation.

Derivation. For $f(x) = e^{-ax^2}$ one has

$$\begin{aligned} \widehat{f}(\xi) &= \int_{\mathbb{R}} e^{-ax^2} e^{-i\xi x} dx = \int_{\mathbb{R}} e^{-a\{x+i\xi/(2a)\}^2} e^{-\xi^2/(4a)} dx \\ &= e^{-\xi^2/(4a)} \int_L e^{-az^2} dz, \quad \text{where } L = \mathbb{R} + i\xi/(2a); \qquad (9.1.1) \end{aligned}$$

cf. Figure 9.1. (We are thinking of the case $\xi > 0$.) Now by Cauchy's theorem, the integral of the analytic function $f(z) = e^{-az^2}$ along the closed rectangular path with vertices $\pm A$, $\pm A + i\xi/(2a)$ is equal to zero. Also, the integrals along the vertical sides of the rectangle will tend to zero as $A \to \infty$:

$$\begin{aligned} &\left| \int_{\pm A}^{\pm A + i\xi/(2a)} e^{-az^2} dz \right| \le \max_{side} \left| e^{-az^2} \right| \cdot \xi/(2a) \\ &= \max_{0 \le y \le \xi/(2a)} e^{-a(x^2-y^2)} \cdot \xi/(2a) = e^{-aA^2+\xi^2/(4a)} \cdot \xi/(2a) \to 0 \end{aligned}$$

as $A \to \infty$. Hence by (9.1.1),

$$\begin{aligned} \widehat{f}(\xi) e^{\xi^2/(4a)} &= \lim_{A\to\infty} \int_{-A+i\xi/(2a)}^{A+i\xi/(2a)} e^{-az^2} dz = \lim_{A\to\infty} \int_{-A}^{A} e^{-az^2} dz \\ &= 2\int_0^\infty e^{-ax^2} dx = 2\int_0^\infty e^{-t} a^{-\frac{1}{2}} (1/2)\, t^{-\frac{1}{2}}\, dt = a^{-\frac{1}{2}} \Gamma(1/2) = \sqrt{\pi/a}. \end{aligned}$$

PROPERTIES 9.1.5. If f is even [or odd, respectively], so is $g = \widehat{f}$:

$$\begin{aligned} g(-\xi) &= \int_{-\infty}^{\infty} f(x)e^{i\xi x}dx = \int_{\infty}^{-\infty} f(-t)e^{-i\xi t}d(-t) \\ &= \int_{-\infty}^{\infty} \pm f(t)e^{-i\xi t}dt = \pm g(\xi). \end{aligned}$$

The Fourier transform g of $f \in L^1(\mathbb{R})$ is bounded on $\mathbb{R}$:

$$|g(\xi| \leq \int_{\mathbb{R}} \left|f(x)e^{-i\xi x}\right|dx = \int_{\mathbb{R}} |f| = \|f\|_1.$$

It is also *continuous*. Indeed, since for real a, b,

$$\left|e^{ib} - e^{ia}\right| = \left|\int_a^b e^{it}dt\right| \leq \min\{2, |b-a|\},$$

one has

$$\begin{aligned} g(\xi) - g(\xi_0) &= \int_{\mathbb{R}} f(x)\left(e^{-i\xi x} - e^{-i\xi_0 x}\right)dx = \int_{-A}^{A} \cdots + \int_{|x|>A} \cdots, \\ \left|\int_{|x|>A} \cdots\right| &\leq 2\int_{|x|>A} |f(x)|dx < \varepsilon \ \text{ for some } A = A(\varepsilon) \text{ and all } \xi, \\ \left|\int_{-A}^{A} \cdots\right| &\leq \max_{|x|\leq A} \left|e^{-i\xi x} - e^{-i\xi_0 x}\right| \cdot \int_{-A}^{A} |f| \\ &\leq |\xi - \xi_0|A\int_{\mathbb{R}} |f| < \varepsilon \quad \text{for } \ |\xi - \xi_0| < \delta. \end{aligned}$$

Furthermore, by the Riemann–Lebesgue Lemma 2.1.1 [which holds for unbounded intervals (a, b) as well as bounded intervals],

$$g(\xi) = \int_{-\infty}^{\infty} f(x)e^{-i\xi x}dx \to 0 \quad \text{as } \ \xi \to \pm\infty.$$

However, the Fourier transform g need not be in $L^1(\mathbb{R})$, as shown by the final Example 9.1.3:

$$\int_{\mathbb{R}} \left|\frac{1}{a+ix}\right| dx = \int_{\mathbb{R}} \frac{1}{\sqrt{a^2+x^2}}\, dx = +\infty.$$

Thus if $g = \mathcal{F}f$, one cannot expect that the inversion formula in (9.0.1): $f = \{1/(2\pi)\}\mathcal{F}_R g$, is valid without suitable interpretation; cf. Section 9.2.

The most important property of Fourier transformation is the way it acts on derivatives:

PROPOSITION 9.1.6. *Let f be continuous and piecewise smooth, or at any rate, let f be equal to an indefinite integral on $\mathbb{R}$. Suppose also that both f and its derivative f' are in $L^1(\mathbb{R})$. Then*

$$(\mathcal{F}f')(\xi) = i\xi(\mathcal{F}f)(\xi), \quad \forall\, \xi \in \mathbb{R}. \tag{9.1.2}$$

PROOF. We first remark that $f(x)$ tends to a finite limit as $x \to +\infty$:

$$f(x) = f(0) + \int_0^x f'(t)dt \to f(0) + \int_0^\infty f'(t)dt \quad \text{as } x \to \infty.$$

Calling the limit $f(\infty)$, we observe that $f(\infty)$ must be zero. Indeed, if $f(\infty) = c \neq 0$, we would have $|f(x)| > |c|/2$ for all x larger than some number A, and then f could not be in $L^1(\mathbb{R})$: $\int_A^\infty |f(x)|dx$ would be infinite. Similarly $f(x) \to f(-\infty) = 0$ as $x \to -\infty$. Integration by parts now gives

$$\begin{aligned}
(\mathcal{F}f')(\xi) &= \int_{-\infty}^\infty f'(x)e^{-i\xi x}dx = \lim_{A\to\infty} \int_{-A}^A \cdots \\
&= \lim_{A\to\infty} \left\{ \left[f(x)e^{-i\xi x}\right]_{-A}^A - \int_{-A}^A f(x)(-i\xi)e^{-i\xi x}dx \right\} \\
&= i\xi \int_{-\infty}^\infty f(x)e^{-i\xi x}dx = i\xi(\mathcal{F}f)(\xi).
\end{aligned}$$

□

COROLLARY 9.1.7. *Suppose that f is an indefinite integral on $\mathbb{R}$ of order $n \geq 1$, that is, f is an indefinite integral, f' is an indefinite integral, $\cdots$, $f^{(n-1)}$ is an indefinite integral. Suppose also that $f, f', \cdots, f^{(n-1)}, f^{(n)}$ are in $L^1(\mathbb{R})$. Then*

$$\begin{aligned}
&\left(\mathcal{F}f^{(k)}\right)(\xi) = (i\xi)^k(\mathcal{F}f)(\xi) \quad \textit{for } 0 \leq k \leq n, \\
&\left|(\mathcal{F}f)(\xi)\right| \leq \frac{A_k}{|\xi|^k} \quad \textit{on } \mathbb{R} \ \textit{for } 0 \leq k \leq n, \ \ A_k = \left\|f^{(k)}\right\|_1, \\
&\mathcal{F}\left[p(D)f\right](\xi) = p(i\xi)(\mathcal{F}f)(\xi), \quad D = d/dx,
\end{aligned}$$

for every polynomial $p(x)$ of degree $\leq n$.

Exercises. 9.1.1. Show that $f(x) = e^{-a|x|}\operatorname{sgn} x = e^{-a|x|}x/|x|$ (with $a > 0$) has Fourier transform $\widehat{f}(\xi) = -2i\xi/(\xi^2 + a^2)$,

(i) by direct computation; (ii) by application of Proposition 9.1.6 to the indefinite integral $f_0(x) = e^{-a|x|}$ on $\mathbb{R}$.

9.1.2. Show that the step function

$$\sigma_a(x) = U(x+a)U(a-x) = \begin{cases} 1 & \text{for } |x| < a \\ 0 & \text{for } |x| > a, \end{cases}$$

with $a > 0$, has Fourier transform

$$\widehat{\sigma}_a(\xi) = 2\,\frac{\sin a\xi}{\xi}.$$

9.1.3. Show that the "triangle function"

$$\Delta_a(x) = \begin{cases} 1 - |x|/a & \text{for } |x| < a \\ 0 & \text{for } |x| \geq a \end{cases} \qquad (a > 0)$$

has Fourier transform

$$\widehat{\Delta}_a(\xi) = \frac{\sin^2 a\xi/2}{a\xi^2/4}.$$

9.1.4. Let $f_k \to f$ in $L^1(\mathbb{R})$, that is, $\int_{\mathbb{R}} |f - f_k| \to 0$. Prove that $\widehat{f_k} \to \widehat{f}$ uniformly on $\mathbb{R}$.

9.1.5. Let f be an integrable function on $\mathbb{R}$ which vanishes for $|x| > a$. Prove that $\widehat{f}(\xi)$ can be extended to an *entire function* $\widehat{f}(\zeta) = f(\xi + i\eta)$, that is, a function analytic for all complex ζ. What can you say about the growth of $\widehat{f}(\zeta)$ as $\zeta \to \infty$ in different directions?

9.2. Fourier inversion

We have seen that the Fourier transform $g = \widehat{f}$ of an L^1 function f need not be in L^1. In fact, Fourier transforms $g(\xi)$ may go to zero very very slowly; cf. Exercises 9.2.6, 9.2.7. However, if f is locally well-behaved, the reflected Fourier transform $\mathcal{F}_R g$ will exist as a Cauchy *principal value* integral [here, a principal value at ∞]. More precisely,

$$\text{p.v.} \int_{\mathbb{R}} g(\xi)e^{ix\xi}d\xi \overset{\text{def}}{=} \lim_{A\to\infty} \int_{-A}^{A} g(\xi)e^{ix\xi}d\xi \tag{9.2.1}$$

will exist, and the limit will be equal to $2\pi f(x)$.

Example 9.2.1. For $f(x) = 1$ on $(-a, a)$, $= 0$ for $|x| > a$, one has $g(\xi) = \widehat{f}(\xi) = 2(\sin a\xi)/\xi$; cf. Exercise 9.1.2. For Fourier inversion we first observe that for $\lambda \in \mathbb{R}$,

$$\lim_{A\to\infty} \int_{-A}^{A} \frac{\sin \lambda\xi}{\xi}\,d\xi = 2 \cdot \lim_{A\to\infty} \int_{0}^{A} \frac{\sin \lambda\xi}{\xi}\,d\xi = \pi \operatorname{sgn} \lambda; \tag{9.2.2}$$

see Exercises 2.5.1 and (for sgn) 1.2.5. Now for our function g,

$$\int_{-A}^{A} g(\xi)e^{ix\xi}d\xi = \int_{-A}^{A} 2\,\frac{\sin a\xi}{\xi}\,e^{ix\xi}d\xi = \int_{-A}^{A} 2\,\frac{\sin a\xi}{\xi}\,\cos x\xi\,d\xi$$
$$= \int_{-A}^{A} \left\{\frac{\sin(a+x)\xi}{\xi} + \frac{\sin(a-x)\xi}{\xi}\right\}d\xi.$$

Hence by (9.2.2),

$$\lim_{A\to\infty} \int_{-A}^{A} g(\xi)e^{ix\xi}d\xi = \pi\,\mathrm{sgn}\,(a+x) + \pi\,\mathrm{sgn}\,(a-x)$$
$$= \begin{cases} 2\pi & \text{for } |x| < a, \\ \pi & \text{for } x = \pm a, \\ 0 & \text{for } |x| > a. \end{cases}$$

The result is equal to $2\pi f(x)$ for $x \neq \pm a$.

THEOREM 9.2.2. (First pointwise inversion theorem) *Let f be in $L^1(\mathbb{R})$ and differentiable at the point x, or at least, satisfy a Hölder–Lipschitz condition at the point x. That is, there should be constants M, α and $\delta > 0$ such that*

$$|f(x+t) - f(x)| \leq M|t|^\alpha \quad \textit{for} \quad -\delta < t < \delta.$$

Then

$$f(x) = \frac{1}{2\pi}\,\text{p.v.} \int_{\mathbb{R}} \widehat{f}(\xi)e^{ix\xi}d\xi = \frac{1}{2\pi} \lim_{A\to\infty} \int_{-A}^{A} \widehat{f}(\xi)e^{ix\xi}d\xi.$$

If in addition $\widehat{f}$ is in $L^1(\mathbb{R})$, then

$$f(x) = \frac{1}{2\pi}\left(\mathcal{F}_R\widehat{f}\right)(x) = \frac{1}{2\pi}\left(\mathcal{F}\widehat{f}_R\right)(x)$$

in the ordinary sense.

We will see in Theorem 9.2.5 below that *continuity* of f at the point x is sufficient for the inversion if $\widehat{f}$ is in $L^1(\mathbb{R})$.

Proof. Inverting order of integration, one finds that

$$\begin{aligned}\int_{-A}^{A} \widehat{f}(\xi)e^{ix\xi}d\xi &= \int_{-A}^{A} \left\{\int_{\mathbb{R}} f(u)e^{-i\xi u}du\right\} e^{ix\xi}d\xi \\ &= \int_{\mathbb{R}} f(u)du \int_{-A}^{A} e^{i(x-u)\xi}d\xi = \int_{\mathbb{R}} f(u)\,\frac{2\sin A(x-u)}{x-u}\,du \\ &= \int_{\mathbb{R}} f(x\pm t)\,\frac{2\sin At}{t}\,dt = \int_{-\delta}^{\delta}\cdots\,dt + \int_{|t|>\delta}\cdots\,dt.\end{aligned} \tag{9.2.3}$$

[The change in order of integration is justified by Fubini's theorem. Indeed, the first repeated integral is absolutely convergent:

$$\int_{-A}^{A} \left\{\int_{\mathbb{R}} \left|f(u)e^{-i\xi u}\right|du\right\} \left|e^{ix\xi}\right|d\xi = 2A\int_{\mathbb{R}} |f(u)|du,$$

which is finite.]

Now let $\varepsilon > 0$ be given. By the Hölder–Lipschitz condition,

$$\left|\int_{-\delta}^{\delta} \{f(x+t)-f(x)\}\,\frac{2\sin At}{t}\,dt\right| \le 4M\delta^{\alpha}/\alpha, \quad \forall\, A,$$

(9.2.4) and this is $< \varepsilon$, $\forall A$, if we take δ small enough.

Keeping δ fixed from here on, we also have

$$\int_{-\delta}^{\delta} f(x)\,\frac{2\sin At}{t}\,dt = 2f(x)\int_{-\delta A}^{\delta A} \frac{\sin v}{v}\,dv \to 2\pi f(x)$$

as $A\to\infty$; cf. (9.2.2). Hence

$$\left|\int_{-\delta}^{\delta} f(x)\,\frac{2\sin At}{t}\,dt - 2\pi f(x)\right| < \varepsilon \quad \text{for } A > A_1. \tag{9.2.5}$$

We finally remark that $f(x+t)/t$ is integrable over $\delta < |t| < \infty$ since $|1/t| < 1/\delta$ there. Thus by the Riemann–Lebesgue Lemma,

$$\left|\int_{|t|>\delta} f(x+t)\,\frac{2\sin At}{t}\,dt\right| < \varepsilon \quad \text{for } A > A_2. \tag{9.2.6}$$

Combining (9.2.3)–(9.2.6), we find that

$$\left|\int_{-A}^{A} \widehat{f}(\xi)e^{ix\xi}d\xi - 2\pi f(x)\right| < 3\varepsilon \quad \text{for } A > \max\{A_1, A_2\}.$$

Conclusion:

$$\lim_{A\to\infty}\int_{-A}^{A}\widehat{f}(\xi)e^{ix\xi}d\xi \quad \text{exists and} \quad = 2\pi f(x).$$

□

EXAMPLE 9.2.3. Applying Fourier inversion to the Fourier pair

$$\Delta_2(x) = \begin{cases} 1-|x|/2 & \text{for } |x|<2 \\ 0 & \text{for } |x|\geq 2 \end{cases}, \qquad \widehat{\Delta}_2(\xi) = 2\,\frac{\sin^2\xi}{\xi^2},$$

cf. Exercise 9.1.3, we obtain the formula

$$\frac{1}{2\pi}\int_{\mathbb{R}} 2\,\frac{\sin^2\xi}{\xi^2}\,e^{ix\xi}d\xi = \Delta_2(x).$$

In particular, for $x=0$:

$$\int_{\mathbb{R}}\frac{\sin^2\xi}{\xi^2}\,d\xi = \pi. \tag{9.2.7}$$

REMARK 9.2.4. For f in $L^1(\mathbb{R})$ and continuous at the point x one has a result analogous to the Cesàro summability of a Fourier series:

$$f(x) = \lim_{A\to\infty}\frac{1}{2\pi}\int_{-A}^{A}\left(1-\frac{|\xi|}{A}\right)\widehat{f}(\xi)e^{ix\xi}d\xi. \tag{9.2.8}$$

[One could say that, although the integral for $(\mathcal{F}_R\widehat{f})(x)$ need not converge, it is "Cesàro summable" to the value $2\pi f(x)$.] For the proof one would put the right-hand side into the form

$$\lim_{A\to\infty}\int_{\mathbb{R}} f(x\pm t)\,\frac{\sin^2 At/2}{2\pi At^2/4}\,dt;$$

cf. Exercise 9.1.3. The positive kernel here has integral one; cf. Example 9.2.3.

As a corollary one obtains

THEOREM 9.2.5. (Second pointwise inversion theorem) *For an L^1 function f on $\mathbb{R}$ which is continuous at the point x, and whose Fourier transform is also in $L^1(\mathbb{R})$, one has*

$$f(x) = \frac{1}{2\pi}\left(\mathcal{F}_R\hat{f}\right)(x).$$

Exercises. 9.2.1. Let f be in $L^1(\mathbb{R})$. Prove that $\mathcal{F}f \equiv 0$ implies $f \equiv 0$,

(i) if f satisfies a Hölder–Lipschitz condition at every point x;

(ii) if f is just continuous.

Carefully state the theorems which you have used.

9.2.2. Let f be an indefinite integral of order two on $\mathbb{R}$ such that f, f' and f'' are in $L^1(\mathbb{R})$. Prove that $g = \mathcal{F}f$ is in $L^1(\mathbb{R})$ and that $f = \frac{1}{2\pi}\mathcal{F}_R g$.

9.2.3. Prove Remark 9.2.4, on the Cesàro summability of $\int_{\mathbb{R}} \widehat{f}(\xi)e^{ix\xi}d\xi$ to $2\pi f(x)$, when $f \in L^1(\mathbb{R})$ is continuous at the point x.

9.2.4. Prove Theorem 9.2.5.

9.2.5. Apply Fourier inversion to the Fourier pair of Exercise 9.1.3. Deduce that for $\lambda \in \mathbb{R}$, $\Delta_a(\xi - \lambda)$ is the Fourier transform of

$$\frac{1}{2\pi}\, e^{i\lambda x}\, \frac{\sin^2 ax/2}{ax^2/4}.$$

9.2.6. Let Δ_a be the triangle function of Exercise 9.1.3 and suppose that $\varepsilon_n > 0$, $\rho_n > 0$ and $\lambda_n \in \mathbb{R}$, $n = 1, 2, \cdots$, while $\sum_1^\infty \varepsilon_n < \infty$. Verify that the function

$$g(\xi) = \sum_1^\infty \varepsilon_n \Delta_{\rho_n}(\xi - \lambda_n)$$

is the Fourier transform of an L^1 function f. Next show that $\int_{-A}^{A} g(\xi)d\xi$ may be almost of the same order of magnitude as A for a sequence of A's tending to ∞.

Hint. One has $\int_{-A}^{A} g(\xi)d\xi \geq \sum' \varepsilon_n\rho_n$, where the summation extends over those integers n for which $-A \leq \lambda_n - \rho_n$, $\lambda_n + \rho_n \leq A$.

9.2.7. (i) Functions f in $L^1(\mathbb{R})$ need not tend to zero as $x \to \infty$. Give an example.

(ii) Fourier transforms $\widehat{f}(\xi)$ of L^1 functions f may tend to zero quite slowly as $\xi \to \infty$. Indeed, prove that for any positive decreasing function $\varepsilon(\xi)$ with limit zero as $\xi \to \infty$, there exist a sequence $\lambda_n \to \infty$ and an L^1 function f such that $\widehat{f}(\lambda_n) \geq \varepsilon(\lambda_n)$, $n = 1, 2, \cdots$.

9.3. Operations on functions and Fourier transformation

In the following it is assumed as a minimum that the function f is integrable over $\mathbb{R}$; the letters g and h are used for the Fourier transform, and the reflected Fourier transform, respectively:

$$g(\xi) = \int_{\mathbb{R}} f(x)e^{-i\xi x}dx, \quad h(\xi) = g_R(\xi) = \int_{\mathbb{R}} f(x)e^{i\xi x}dx.$$

	Original	Fourier trf	Reflected trf	Remark
(i)	$f(\lambda x)$	$\frac{1}{\lvert\lambda\rvert}\, g\Big(\frac{\xi}{\lambda}\Big)$	$\frac{1}{\lvert\lambda\rvert}\, h\Big(\frac{\xi}{\lambda}\Big)$	λ real, $\neq 0$
(ii)	$f(x+\lambda)$	$e^{i\lambda\xi}g(\xi)$	$e^{-i\lambda\xi}h(\xi)$	λ real
(iii)	$e^{i\lambda x}f(x)$	$g(\xi-\lambda)$	$h(\xi+\lambda)$	λ real
(iv)	$Df(x)$	$i\xi g(\xi)$	$-i\xi h(\xi)$	$D = d/dx$
(v)	$xf(x)$	$iDg(\xi)$	$-iDh(\xi)$	$D = d/d\xi$
(vi)	$p(D)f(x)$	$p(i\xi)g(\xi)$	$p(-i\xi)h(\xi)$	$p(x) =$
(vii)	$p(x)f(x)$	$p(iD)g(\xi)$	$p(-iD)h(\xi)$	$\sum_0^n a_k x^k$
(viii)	$(f_1 * f_2)(x)$	$g_1(\xi)g_2(\xi)$	$h_1(\xi)h_2(\xi)$	Section 9.4
(ix)	$f_1(x)f_2(x)$	$\frac{1}{2\pi}\,(g_1 * g_2)(\xi)$	$\frac{1}{2\pi}\,(h_1 * h_2)(\xi)$	Section 9.4

For suitably matched locally integrable functions f_j on $\mathbb{R}$ one defines the *convolution* by the formula

$$(f_1 * f_2)(x) = \int_{\mathbb{R}} f_1(x-y)f_2(y)dy = \int_{\mathbb{R}} f_1(y)f_2(x-y)dy. \tag{9.3.1}$$

For integrable f_1 on $\mathbb{R}$ and bounded f_2, the convolution is defined for all $x \in \mathbb{R}$. If both f_j are in $L^1(\mathbb{R})$, the convolution integral will exist almost everywhere, and $f_1 * f_2$ will be integrable over $\mathbb{R}$; see Section 9.4.

Discussion of rules (i)–(vii) below. Rules (i)–(iii) follow immediately from the defining integrals. For example, if $\lambda < 0$,

$$\begin{aligned}\int_{-\infty}^{\infty} f(\lambda x)e^{-i\xi x}dx &= \int_{\infty}^{-\infty} f(t)e^{-i\xi t/\lambda}\, dt/\lambda \\ &= -\frac{1}{\lambda}\int_{-\infty}^{\infty} f(t)e^{-i(\xi/\lambda)t}dt = \frac{1}{|\lambda|}\, g\left(\frac{\xi}{\lambda}\right),\end{aligned}$$

while for any $\lambda \in \mathbb{R}$,

$$\int_{-\infty}^{\infty} f(x+\lambda)e^{-i\xi x}dx = \int_{-\infty}^{\infty} f(t)e^{-i\xi(t-\lambda)}dt = e^{i\lambda\xi}g(\xi).$$

Sufficient conditions for the validity of rules (iv) and (vi) have been stated in Proposition 9.1.6 and Corollary 9.1.7.

Rule (v) will be valid under the natural condition that both f and xf are in $L^1(\mathbb{R})$. Indeed, by (iii),

$$\int_{\mathbb{R}} \left\{ \frac{e^{i\lambda x}-1}{i\lambda} - x \right\} f(x)e^{-i\xi x}dx = \frac{g(\xi-\lambda)-g(\xi)}{i\lambda} - \mathcal{F}[xf(x)](\xi). \tag{9.3.2}$$

Now observe that

$$\begin{aligned} \left|\frac{e^{i\lambda x}-1}{i\lambda} - x\right| &= \left|\int_0^x \left(e^{i\lambda t}-1\right)dt\right| \\ &\leq \int_0^{|x|} \min\{2, |\lambda t|\}dt \leq \min\{2|x|, |\lambda|x^2/2\}. \end{aligned} \tag{9.3.3}$$

Denoting the integrand in (9.3.2) by $F_\lambda(x)$, we will show that $\int_{\mathbb{R}} F_\lambda \to 0$ as $\lambda \to 0$. To that end split $\int_{\mathbb{R}} F_\lambda$ as $\int_{-A}^{A} + \int_{|x|>A}$. For given $\varepsilon > 0$ one first chooses A so large that $\int_{|x|>A} 2|xf(x)|dx < \varepsilon$, so that by (9.3.3), $\left|\int_{|x|>A} F_\lambda\right| < \varepsilon$, $\forall\, \lambda,\, \xi$. Keeping A fixed, one next takes $\delta > 0$ so small that $|\lambda/2| \int_{-A}^{A} x^2|f(x)|dx < \varepsilon$ for $|\lambda| < \delta$, so that $\left|\int_{-A}^{A} F_\lambda\right| < \varepsilon$ for $|\lambda| < \delta$ and all ξ. Then $\left|\int_{\mathbb{R}} F_\lambda\right| < 2\varepsilon$ for $|\lambda| < \delta$ and all ξ. Conclusion from (9.3.2):

$$\lim_{\lambda\to 0} \frac{g(\xi-\lambda)-g(\xi)}{i\lambda} = ig'(\xi) \quad \text{exists and} \quad = \mathcal{F}[xf(x](\xi).$$

Apparently, $g'(\xi)$ may here be obtained by *differentiation under the integral sign.* Repeated application of rule (v) gives rule (vii) when the functions $f, xf, \cdots, x^n f$ are all in $L^1(\mathbb{R})$.

EXAMPLE 9.3.1. One may compute $\mathcal{F}\left[e^{-ax^2}\right]$ (with $a > 0$) by observing that $y = e^{-ax^2}$ satisfies the differential equation $Dy = -2axy$, so that by Fourier transformation, $i\xi\,\widehat{y}(\xi) = -2aiD\widehat{y}(\xi)$; see rules (iv) and (v). Integrating the equation $(1/\widehat{y})D\widehat{y} = -\xi/(2a)$ one obtains $\log \widehat{y}(\xi) - \log \widehat{y}(0) = -\xi^2/(4a)$, so that $\widehat{y}(\xi) = \widehat{y}(0)e^{-\xi^2/(4a)}$. Here $\widehat{y}(0) = \int_{\mathbb{R}} e^{-ax^2}dx = \sqrt{\pi/a}$; cf. Example 9.1.4.

EXAMPLE 9.3.2. (*Hermite functions*) By Proposition 7.3.6,

$$h_n(x) = \rho_n H_n(x) e^{-\frac{1}{2}x^2} = \rho_n (x-D)^n e^{-\frac{1}{2}x^2}.$$

The Fourier transform of $e^{-\frac{1}{2}x^2}$ is $\sqrt{(2\pi)}\, e^{-\frac{1}{2}\xi^2}$. Thus

$$\begin{aligned}(\mathcal{F}h_1)(\xi) &= \mathcal{F}\left[\rho_1(x-D)e^{-\frac{1}{2}x^2}\right](\xi) = \rho_1(iD-i\xi)\mathcal{F}\left[e^{-\frac{1}{2}x^2}\right](\xi)\\ &= -i\rho_1(\xi-D)\sqrt{(2\pi)}\, e^{-\frac{1}{2}\xi^2} = -i\sqrt{(2\pi)}\, h_1(\xi).\end{aligned}$$

In general,

$$\begin{aligned}(\mathcal{F}h_n)(\xi) &= \mathcal{F}\left[\rho_n(x-D)^n e^{-\frac{1}{2}x^2}\right](\xi) = \rho_n(iD-i\xi)^n\mathcal{F}\left[e^{-\frac{1}{2}x^2}\right](\xi)\\ &= (-i)^n\rho_n(\xi-D)^n\sqrt{(2\pi)}\, e^{-\frac{1}{2}\xi^2} = (-i)^n\sqrt{(2\pi)}\, h_n(\xi).\end{aligned}$$

Using x as independent variable, we may write

$$(\mathcal{F}h_n)(x) = (-i)^n\sqrt{(2\pi)}\, h_n(x) : \tag{9.3.4}$$

the Hermite functions are *eigenfunctions* of Fourier transformation.

Exercises. 9.3.1. Given that $\mathcal{F}\left[e^{-\frac{1}{2}x^2}\right](\xi) = ce^{-\frac{1}{2}\xi^2}$, deduce the value of c from the Fourier inversion theorem. Use the answer to compute $\mathcal{F}\left[e^{-ax^2}\right]$ by rule (i).

9.3.2. Given that $\mathcal{F}[e^{-x}U(x)](\xi) = 1/(1+i\xi)$, compute

$$\mathcal{F}[e^{x}U(-x)],\quad \mathcal{F}[e^{-ax}U(x)],\quad \mathcal{F}[xe^{-ax}U(x)],\quad \mathcal{F}[e^{-ax}U(x-b)]$$

by using appropriate rules.

9.3.3. Prove that the operators $\mathcal{F}$ and (x^2-D^2) commute when applied to "good" functions. Deduce without identifying them that all "smooth and small" eigenfunctions for the operator (x^2-D^2) must also be eigenfunctions for $\mathcal{F}$.

9.4. Products and convolutions

We begin with the important

PROPOSITION 9.4.1. *For f and ϕ in $L^1(\mathbb{R})$ one has*

$$\int_{\mathbb{R}}\mathcal{F}f\cdot\phi = \int_{\mathbb{R}} f\cdot\mathcal{F}\phi,\qquad \int_{\mathbb{R}}\mathcal{F}_R f\cdot\phi = \int_{\mathbb{R}} f\cdot\mathcal{F}_R\phi. \tag{9.4.1}$$

This proposition will later become the basis for an operational definition of extended Fourier transformation. It says that

"$\mathcal{F}f$ does to ϕ whatever f does to $\mathcal{F}\phi$",

and similarly for $\mathcal{F}_R$.

PROOF. The formulas are direct applications of Fubini's theorem. Inverting order of integration, one finds that

$$\begin{aligned}\int_{\mathbb{R}}(\mathcal{F}f)(\xi)\phi(\xi)d\xi &= \int_{\mathbb{R}}\left\{\int_{\mathbb{R}} f(x)e^{-i\xi x}dx\right\}\phi(\xi)d\xi \\ &= \int_{\mathbb{R}} f(x)dx\int_{\mathbb{R}}\phi(\xi)e^{-ix\xi}d\xi = \int_{\mathbb{R}} f(x)(\mathcal{F}\phi)(x)dx.\end{aligned}$$

The second step is justified by the absolute convergence of one of the repeated integrals:

$$\int_{\mathbb{R}}|f(x)|dx\int_{\mathbb{R}}|\phi(\xi)e^{-ix\xi}|d\xi = \int_{\mathbb{R}}|f|\int_{\mathbb{R}}|\phi| \quad \text{is finite.}$$

□

Rule (ix) in the table of Section 9.3 may be obtained by the same method. In the derivation below, the convolution will show up in a natural manner.

PROPOSITION 9.4.2. *Let f_1 and g_2 be arbitrary functions in $L^1(\mathbb{R})$ and set $g_1 = \mathcal{F}f_1$, $f_2 = (1/2\pi)\mathcal{F}_R g_2$* [so that formally, $g_2 = \mathcal{F}f_2$]. *Then*

$$\int_{\mathbb{R}} f_1(x)f_2(x)e^{-i\xi x}dx = \frac{1}{2\pi}\int_{\mathbb{R}} g_1(\xi-t)g_2(t)dt = \frac{1}{2\pi}(g_1 * g_2)(\xi). \tag{9.4.2}$$

PROOF. Inverting order of integration, we obtain

$$\begin{aligned}\mathcal{F}[f_1f_2](\xi) &= \int_{\mathbb{R}} f_1(x)\left\{\frac{1}{2\pi}\int_{\mathbb{R}} g_2(t)e^{ixt}dt\right\}e^{-i\xi x}dx \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} g_2(t)dt\int_{\mathbb{R}} f_1(x)e^{-i(\xi-t)x}dx = \frac{1}{2\pi}\int_{\mathbb{R}} g_2(t)g_1(\xi-t)dt.\end{aligned}$$

□

The "dual result", rule (viii) in the table of Section 9.3, is more directly relevant for the applications:

PROPOSITION 9.4.3. *Suppose that f_1 and f_2 belong to $L^1(\mathbb{R})$. Then the convolution*

$$(f_1 * f_2)(x) = \int_{\mathbb{R}} f_1(y)f_2(x-y)dy$$

exists for almost all $x \in \mathbb{R}$. Giving it arbitrary values for the exceptional x, the resulting function $f_1 * f_2$ belongs to $L^1(\mathbb{R})$, and

$$\mathcal{F}[f_1 * f_2](\xi) = (\mathcal{F}f_1)(\xi)(\mathcal{F}f_2)(\xi) = g_1(\xi)g_2(\xi). \tag{9.4.3}$$

PROOF. We apply Fubini's theorem to $F(x,y) = f_1(y)f_2(x-y)$ on the "rectangle" $\mathbb{R}\times\mathbb{R}$. Since the repeated integral

$$\begin{aligned}\int_{\mathbb{R}} dy \int_{\mathbb{R}} |F(x,y)|dx &= \int_{\mathbb{R}} |f_1(y)|dy \int_{\mathbb{R}} |f_2(x-y)|dx \\ &= \int_{\mathbb{R}} |f_1(y)|dy \int_{\mathbb{R}} |f_2(t)|dt\end{aligned}$$

is finite, Fubini's theorem says that $F(x,y)$ is integrable over $\mathbb{R}^2$ and that

$$\left[\int_{\mathbb{R}^2} F(x,y)dxdy =\right] \int_{\mathbb{R}} dy \int_{\mathbb{R}} F(x,y)dx = \int_{\mathbb{R}} dx \int_{\mathbb{R}} F(x,y)dy.$$

More precisely, $G(x) = \int_{\mathbb{R}} F(x,y)dy$ – in our case, $\int_{\mathbb{R}} f_1(y)f_2(x-y)dy$ – will exist for almost all x, the function $G\,(= f_1 * f_2)$ will be in $L^1(\mathbb{R})$, and

$$\begin{aligned}&\int_{\mathbb{R}} G(x)dx \left[= \int_{\mathbb{R}} f_1 * f_2\right] = \int_{\mathbb{R}} dx \int_{\mathbb{R}} F(x,y)dy \\ &= \int_{\mathbb{R}} dy \int_{\mathbb{R}} F(x,y)dx = \int_{\mathbb{R}} f_1(y)dy \int_{\mathbb{R}} f_2(x-y)dx = \int_{\mathbb{R}} f_1 \int_{\mathbb{R}} f_2.\end{aligned}$$

For any $\xi \in \mathbb{R}$, the final argument will also give (9.4.3), that is, rule (viii):

$$\begin{aligned}&\int_{\mathbb{R}} (f_1 * f_2)(x)e^{-i\xi x}dx = \int_{\mathbb{R}} \left\{\int_{\mathbb{R}} f_1(y)f_2(x-y)dy\right\} e^{-i\xi x}dx \\ &\quad = \int_{\mathbb{R}} \left\{\int_{\mathbb{R}} f_1(y)e^{-i\xi y} \cdot f_2(x-y)e^{-i\xi(x-y)}dy\right\} dx \\ &\quad = \int_{\mathbb{R}} f_1(y)e^{-i\xi y}dy \int_{\mathbb{R}} f_2(x-y)e^{-i\xi(x-y)}d(x-y) = g_1(\xi)g_2(\xi).\end{aligned}$$

□

EXAMPLE 9.4.4. Let f be in $L^1(\mathbb{R})$ and $t > 0$. Problem: Determine the function $u(x)$ that has Fourier transform $\widehat{u}(\xi) = \widehat{f}(\xi)e^{-t\xi^2}$.

Solution. By rule (viii), $u(x)$ will be the convolution of $f(x)$ and the inverse Fourier transform $h(x)$ of $e^{-t\xi^2}$. Now since $e^{-t\xi^2}$ is even, we obtain from Example 9.1.4 or 9.3.1 that

$$h(x) = \frac{1}{2\pi}\mathcal{F}_R\left[e^{-t\xi^2}\right](x) = \frac{1}{2\pi}\mathcal{F}\left[e^{-t\xi^2}\right](x) = \frac{1}{2\sqrt{\pi t}}\, e^{-x^2/(4t)}.$$

Conclusion:

$$u(x) = f(x) * \frac{1}{2\sqrt{\pi t}}\, e^{-x^2/(4t)} = \frac{1}{2\sqrt{\pi t}} \int_{\mathbb{R}} f(x-y) e^{-y^2/(4t)} dy$$
$$= \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} f\left(x - 2\sqrt{t}\, w\right) e^{-w^2} dw.$$

Exercises. 9.4.1. Let $\sigma_1(x) = 1$ on $(-1, 1)$, $= 0$ for $|x| > 1$. Determine $(\sigma_1 * \sigma_1)(x)$:

(i) by direct computation,

(ii) with the aid of Fourier transformation and Exercises 9.1.2, 9.1.3.

9.4.2. Which functions have Fourier transforms $(\sin^3 \xi)/\xi^3$, $(\sin^4 \xi)/\xi^4$? Compute the integrals of these Fourier transforms over $\mathbb{R}$.

9.4.3. Let $f(x) = |x|^{-\frac{1}{2}}$ for $0 < x \leq 1$, $= 0$ for all other $x \in \mathbb{R}$. Prove that the convolution $(f * f)(x)$ does not exist at the point $x = 0$.

9.4.4. Let $p(x)$ be a polynomial in x of degree n. Use Fourier transformation to determine a (formal) solution of the differential equation $p(D)u = f$ when f is in $L^1(\mathbb{R})$.

9.4.5. Use Fourier transformation to obtain a solution of the differential equation $u'' - u = f$ when f is in $L^1(\mathbb{R})$. Why are there difficulties in the case of the equation $u'' + u = f$? And in the case of $u' = f$?

9.4.6. Determine a solution of the integral equation

$$u(x) + 4 \int_{\mathbb{R}} e^{-|x-y|} u(y) dy = f(x) \quad \text{on } \mathbb{R} \quad \text{for } f \in L^1(\mathbb{R}).$$

9.4.7. Determine the convolution

$$\frac{1}{2\sqrt{\pi s}}\, e^{-x^2/(4s)} * \frac{1}{2\sqrt{\pi t}}\, e^{-x^2/(4t)} \qquad (s,\, t > 0).$$

9.4.8. Let f be in $L^1(\mathbb{R})$ and $y > 0$. Determine the function $u(x)$ with Fourier transform $\widehat{u}(\xi) = \widehat{f}(\xi) e^{-y|\xi|}$.

9.5. Applications in mathematics

We will apply the preceding theory to some mathematical questions; cf. Exercises 3.3.4 and 3.4.3.

THEOREM 9.5.1. *Fourier transformation on $L^1(\mathbb{R})$ is one to one: if $\mathcal{F} f_1 = \mathcal{F} f_2$ for integrable functions f_j on $\mathbb{R}$, then $f_1(x) = f_2(x)$ almost everywhere, and hence $f_1 = f_2$ in the sense of the normed space $L^1(\mathbb{R})$.*

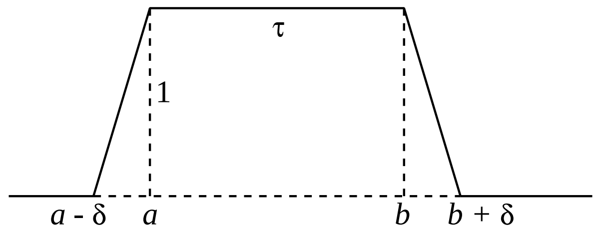

FIGURE 9.2

PROOF. Setting $f_1 - f_2 = f$, let f be in $\mathcal{L}^1(\mathbb{R})$ and $g = \widehat{f} = 0$. We now introduce a trapezoidal function $\tau(x)$ as follows; cf. Figure 9.2. For $a < b$ and $\delta > 0$,

$$\tau(x) = \tau(x; a, b, \delta) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{for } a \leq x \leq b \\ 0 & \text{for } x \leq a - \delta \text{ and } x \geq b + \delta \\ \text{linear} & \text{for } a - \delta \leq x \leq a \text{ and } b \leq x \leq b + \delta. \end{cases}$$

Since τ is the difference of two triangle functions, the Fourier transform $\widehat{\tau}$ is in $L^1(\mathbb{R})$. Hence by Theorem 9.2.2 on pointwise Fourier inversion, $\tau = (1/2\pi)\,(\mathcal{F}_R\widehat{\tau})$ can be written as the Fourier transform $\widehat{\omega}$ of an L^1 function ω. [Just take $\omega = (1/2\pi)\tau_R$.] Thus by Proposition 9.4.1,

$$\int_{a-\delta}^{b+\delta} f\tau = \int_{\mathbb{R}} f\tau = \int_{\mathbb{R}} f\widehat{\omega} = \int_{\mathbb{R}} \widehat{f}\omega = 0.$$

It follows that

$$\left|\int_a^b f\right| = \left|\int_a^b f\tau\right| = \left|\int_{a-\delta}^a f\tau + \int_b^{b+\delta} f\tau\right| \leq \int_{a-\delta}^a |f| + \int_b^{b+\delta} |f|.$$

As $\delta \searrow 0$, the right-hand member will tend to zero: indefinite integrals $\int_{x_0}^x |f|$ are continuous. Thus $\int_a^b f = 0$. This holds for all intervals (a, b). In particular then

$$F(x) \stackrel{\text{def}}{=} \int_0^x f(t)dt = 0, \quad \forall\, x.$$

Since by Integration Theory $f(x) = F'(x)$ almost everywhere [cf. the proof of Theorem 4.1.1], the conclusion is that $f(x) = 0$ a.e. □

THEOREM 9.5.2. (Moment theorem) *Let $f(x)e^{b|x|}$ be integrable over $\mathbb{R}$ for some number $b > 0$, and suppose that all power moments of f are equal*

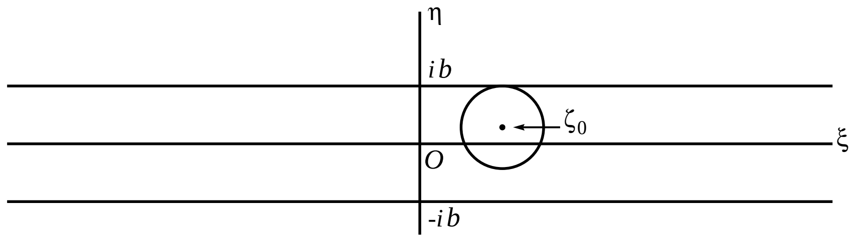

FIGURE 9.3

to zero:

$$\int_{\mathbb{R}} f(x)x^n dx = 0, \quad n = 0, 1, 2, \cdots . \tag{9.5.1}$$

Then $f(x) = 0$ *almost everywhere on* $\mathbb{R}$.

PROOF. Let $g(\xi) = \int_{\mathbb{R}} f(x)e^{-i\xi x}dx$ be the Fourier transform of f. By the hypothesis $x^n f(x)$ is in $L^1(\mathbb{R})$ for every $n \in \mathbb{N}_0$. Hence by the proof of rule (v) in Section 9.3, $g(\xi)$ is infinitely differentiable and

$$g^{(n)}(\xi) = (-i)^n \int_{\mathbb{R}} x^n f(x)e^{-i\xi x}dx.$$

Thus by (9.5.1),

$$g^{(n)}(0) = 0, \quad n = 0, 1, 2, \cdots . \tag{9.5.2}$$

If we would know that g is analytic in a complex neighborhood Ω of the real axis, it would now follow that $g = 0$ in a neighborhood of the origin and hence on Ω. [Uniqueness Theorem for analytic functions; see Complex Analysis.] In particular $g = \mathcal{F}f$ would vanish on $\mathbb{R}$ and hence $f(x) = 0$ a.e. [Theorem 9.5.1]. □

The desired analyticity will follow from

PROPOSITION 9.5.3. *Suppose that* $f(x)e^{b|x|}$ *with* $b > 0$ *is integrable over* $\mathbb{R}$. *Then the Fourier transform* $g(\xi)$ *of* f *has an analytic extension* $g(\zeta) = g(\xi + i\eta)$ *to the strip* $\{|\eta| = |\mathrm{Im}\,\zeta| < b\}$ [see Figure 9.3].

PROOF. By the hypothesis the "complex Fourier transform"

$$g(\zeta) = \int_{\mathbb{R}} f(x)e^{-i\zeta x}dx = \int_{\mathbb{R}} f(x)e^{\eta x}e^{-i\xi x}dx$$

is well-defined for $|\eta| \le b$. We expand the integrand according to powers of $\zeta - \zeta_0$:

$$\begin{aligned} f(x)e^{-i\zeta x} &= f(x)e^{-i\zeta_0 x}e^{-i(\zeta-\zeta_0)x} \\ &= \sum_{n=0}^{\infty} f(x)e^{-i\zeta_0 x}\,\frac{(-ix)^n(\zeta-\zeta_0)^n}{n!}. \end{aligned} \tag{9.5.3}$$

For $\zeta_0 = \xi_0 + i\eta_0$ with $|\eta_0| < b$ and $|\zeta - \zeta_0| < b - |\eta_0|$, this series may be integrated term by term to obtain

$$g(\zeta) = \int_{\mathbb{R}} f(x)e^{-i\zeta x}dx = \sum_{n=0}^{\infty} c_n(\zeta-\zeta_0)^n, \tag{9.5.4}$$

where

$$c_n = \frac{1}{n!}\int_{\mathbb{R}}(-ix)^n f(x)e^{-i\zeta_0 x}dx, \quad n = 0, 1, 2, \cdots . \tag{9.5.5}$$

The justification is by norm convergence of the series (9.5.3) in $L^1(\mathbb{R})$, cf. Examples 5.4.6:

$$\begin{aligned} \sum_{n=0}^{N}\int_{\mathbb{R}}\left|f(x)e^{-i\zeta_0 x}\,\frac{(-ix)^n(\zeta-\zeta_0)^n}{n!}\right|dx &= \int_{\mathbb{R}}\left\{\sum_{0}^{N}\left|\cdots\right|\right\}dx \\ \le \int_{\mathbb{R}}\left\{\sum_{0}^{\infty}|f(x)|e^{\eta_0 x}\,\frac{|x|^n|\zeta-\zeta_0|^n}{n!}\right\}dx &= \int_{\mathbb{R}}|f(x)|e^{\eta_0 x+|\zeta-\zeta_0||x|}\,dx \\ \le \int_{\mathbb{R}}|f(x)|e^{b|x|}dx \quad \text{[a finite constant]}, \ \ \forall N. \end{aligned}$$

Since ζ_0 was arbitrary in the strip $\{|\mathrm{Im}\,\zeta| < b\}$, it follows from (9.5.4) that $g(\zeta)$ is analytic in that strip. One may also observe that by (9.5.5),

$$g^{(n)}(\zeta_0) = \int_{\mathbb{R}}(-ix)^n f(x)e^{-i\zeta_0 x}dx.$$

□

Returning to the Moment Theorem 9.5.2, we remark that the growth condition "$f(x)e^{b|x|}$ in $L^1(\mathbb{R})$ for some number $b > 0$" cannot be relaxed very much; cf. Exercise 9.5.2. Theorem 9.5.2 has an important corollary:

THEOREM 9.5.4. *The normalized Hermite functions*

$$h_n(x) = \rho_n H_n(x)e^{-\frac{1}{2}x^2}, \quad n = 0, 1, 2, \cdots$$

[Definition 7.3.5] *form an orthonormal basis of* $L^2(\mathbb{R})$.

Proof. We will show that $\{h_n\}$ is a maximal orthogonal system. To that end, suppose that $g \in L^2(\mathbb{R})$ is orthogonal to all functions h_n. Since $H_n(x)$ is a polynomial of precise degree n, it will follow that g is orthogonal to all products $x^n e^{-\frac{1}{2}x^2}$:

$$\int_{\mathbb{R}} g(x) e^{-\frac{1}{2}x^2} x^n dx = 0, \quad \forall\, n \in \mathbb{N}_0.$$

Now as the product of two L^2 functions, the function

$$g(x)e^{-\frac{1}{2}x^2} \cdot e^{b|x|} = g(x) \cdot e^{-\frac{1}{2}x^2} e^{b|x|}$$

is integrable over $\mathbb{R}$ for every constant b [use Cauchy–Schwarz]. Hence by Theorem 9.5.2, $g(x)e^{-\frac{1}{2}x^2} = 0$ a.e., so that $g = 0$ in the sense of $L^2(\mathbb{R})$. $\square$

Exercises. 9.5.1. Let $\sum_0^\infty c_n h_n = \sum_0^\infty c_n[f] h_n$ be the Hermite expansion of f in $L^2(\mathbb{R})$. Prove that

(i) $s_k = \sum_0^k c_n h_n \to f$ in $L^2(\mathbb{R})$ as $k \to \infty$;

(ii) $\sum_0^\infty |c_n|^2 = \int_{\mathbb{R}} |f|^2$;

(iii) The series $\sum_0^\infty c_n \mathcal{F} h_n = \sum_0^\infty (-i)^n \sqrt{(2\pi)}\, c_n h_n$ converges to a function g in $L^2(\mathbb{R})$. [This g will be the "generalized Fourier transform" of f as defined in Section 10.2 below].

9.5.2. For $0 < p_k \nearrow \infty$, $\sum_1^\infty 1/p_k < \infty$, define

$$f(x) = \prod_{k=1}^{\infty} (1 - ix/p_k)^{-1}.$$

(i) Taking $p_k = k^2$, show that $|f(x)| \approx e^{-c\sqrt{|x|}}$ (with $c > 0$) as $|x| \to \infty$.

Hint. If $n(t)$ is the number of $p_k \le t$ [here $n(t) \approx \sqrt{t}$], one has

$$\log |f(x)| = -\int_0^\infty \frac{x^2}{x^2 + t^2} \frac{n(t)}{t}\, dt.$$

(ii) Show that $f(z) = f(x + iy)$ is analytic for $y > -p_1 = -1$, and that $|f(x + iy)| \le |f(x)|$ for $y \ge 0$.

(iii) Prove that $g(\xi) = \widehat{f}(\xi)$ is of class $\mathcal{C}^\infty$ on $\mathbb{R}$ and use Cauchy's theorem to show that $g(\xi) = 0$ for $\xi < 0$.

(iv) Show that $\int_{\mathbb{R}} x^n f(x) dx = 0$, $n = 0, 1, 2, \cdots$.

(v) Prove corresponding results for the case where $p_k = k^{1+\delta}$ with $\delta > 0$.

9.6. The test space $\mathcal{S}$ and Fourier transformation

In order to extend Fourier transformation to a large class of functions and generalized functions, one needs suitable test functions; cf. Chapter 4. In the years 1945–1950, Laurent Schwartz introduced the test space $\mathcal{S}$ of "rapidly decreasing functions with rapidly decreasing derivatives"; cf. [**110**]. It consists of the $\mathcal{C}^\infty$ functions ϕ on $\mathbb{R}$ with the following property: $\phi(x)$ and its derivatives $\phi'(x)$, $\phi''(x)$, $\cdots$ tend to zero faster than every negative power of x as $x \to \pm\infty$. Equivalently one has

Definition 9.6.1. $\mathcal{S}$ consists of the $\mathcal{C}^\infty$ functions ϕ on $\mathbb{R}$ for which each of the so-called *seminorms*

$$M_{pq}(\phi) \stackrel{\text{def}}{=} \sup_{x\in\mathbb{R}} \left|x^p \phi^{(q)}(x)\right|, \quad p,\, q = 0, 1, 2, \cdots \tag{9.6.1}$$

is finite.

Important members of $\mathcal{S}$ are the functions e^{-ax^2} $(a > 0)$ and the Hermite functions $h_n(x) = \rho_n H_n(x) e^{-\frac{1}{2}x^2}$ [Definition 7.3.5].

Proposition 9.6.2. (Fourier inversion on $\mathcal{S}$) *Let ϕ be in $\mathcal{S}$. Then $\widehat{\phi} = \mathcal{F}\phi$ is also in $\mathcal{S}$, and*

$$\phi = \frac{1}{2\pi}\,\mathcal{F}_R\widehat{\phi} = \mathcal{F}\,\frac{1}{2\pi}\,\widehat{\phi}_R.$$

Thus in operator sense,

$$\mathcal{F}_R\mathcal{F} = \mathcal{F}\mathcal{F}_R = 2\pi \times \text{identity} \ \ \textit{on} \ \ \mathcal{S}.$$

Proof. (i) For ϕ in $\mathcal{S}$ the functions ϕ, $x\phi$, $x^2\phi$, $\cdots$ are in $L^1(\mathbb{R})$; see (9.6.1) with $q = 0$. Hence by rule (v) in Section 9.3 and its proof, $\widehat{\phi}$ is differentiable, $iD\widehat{\phi} = \mathcal{F}[x\phi]$ is differentiable, $(iD)^2\widehat{\phi} = \mathcal{F}[x^2\phi]$ is differentiable, etc. Thus $\widehat{\phi}$ will be of class $\mathcal{C}^\infty$ on $\mathbb{R}$.

The $\mathcal{C}^\infty$ functions $x^q\phi$, $D(x^q\phi)$, $\cdots, D^p(x^q\phi)$, $\cdots$ will also be in $L^1(\mathbb{R})$; cf. (9.6.1). Hence the $\mathcal{C}^\infty$ functions

$$(i\xi)^p (iD)^q \widehat{\phi} = \mathcal{F}[D^p(x^q\phi)]$$

are bounded on $\mathbb{R}$. Thus by (9.6.1) the function $\widehat{\phi}$ is in $\mathcal{S}$.

(ii) By Inversion Theorem 9.2.2 one has

$$\phi = \frac{1}{2\pi}\,\mathcal{F}_R\widehat{\phi} = \frac{1}{2\pi}\,\mathcal{F}_R\mathcal{F}\phi, \ \text{ and } \ \phi = \frac{1}{2\pi}\,\mathcal{F}\widehat{\phi}_R = \frac{1}{2\pi}\,\mathcal{F}\mathcal{F}_R\phi.$$

□

For later use we define a strong notion of convergence in $\mathcal{S}$ with the aid of the seminorms (9.6.1):

DEFINITION 9.6.3. One says that $\phi_j \to \phi$ in $\mathcal{S}$ if $M_{pq}(\phi - \phi_j) \to 0$ as $j \to \infty$ for every p and q. In other words,

$$x^p \phi_j^{(q)}(x) \to x^p \phi^{(q)}(x) \ \text{ uniformly on } \mathbb{R}, \quad \forall\, p,\, q \in \mathbb{N}_0.$$

We will see in Section 10.3 that for ϕ in $\mathcal{S}$, the Hermite series converges to ϕ in this strong sense.

PROPOSITION 9.6.4. *Fourier transformation defines a one to one continuous linear map of $\mathcal{S}$ onto itself.*

PROOF. That Fourier transformation $\mathcal{F}$ restricted to $\mathcal{S}$ is both injective [that is, one to one] and surjective [that is, onto] follows from Proposition 9.6.2. Indeed, if $\widehat{\phi} = 0$ then $\phi = (1/2\pi)\mathcal{F}_R\widehat{\phi} = 0$. Moreover, every ϕ in $\mathcal{S}$ is the image of an element in $\mathcal{S}$: $\phi = \mathcal{F}(1/2\pi)\widehat{\phi}_R$.

Suppose now that $\phi_j \to \phi$ in $\mathcal{S}$. Then for fixed p, $q \in \mathbb{N}_0$,

$$(x^2+1)D^p\bigl\{x^q(\phi - \phi_j)(x)\bigr\} \to 0 \ \text{ uniformly on } \ \mathbb{R}.$$

Hence for given $\varepsilon > 0$,

$$\bigl|D^p\bigl\{x^q(\phi - \phi_j)(x)\bigr\}\bigr| < \frac{\varepsilon}{x^2+1} \ \text{ on } \ \mathbb{R} \ \text{ for } \ j > j_0 = j_0(\varepsilon).$$

It follows that for $j > j_0$,

$$\begin{aligned} M_{pq}(\widehat{\phi} - \widehat{\phi}_j) &= \sup_\xi \left|(i\xi)^p(iD)^q(\widehat{\phi} - \widehat{\phi}_j)(\xi)\right| \\ &= \sup_\xi \left|\int_{\mathbb{R}} D^p\left\{x^q(\phi - \phi_j)(x)\right\} \cdot e^{-i\xi x} dx\right| < \int_{\mathbb{R}} \frac{\varepsilon}{x^2+1}\, dx = \pi\varepsilon. \end{aligned}$$

□

Exercises. 9.6.1. Derive the *Parseval formula* for Fourier transformation on $\mathcal{S}$:

$$\int_{\mathbb{R}} \left|\widehat{\phi}(\xi)\right|^2 d\xi = 2\pi \int_{\mathbb{R}} |\phi(x)|^2 dx, \quad \forall\, \phi \in \mathcal{S}.$$

9.6.2. Prove that on $\mathcal{S}$,

$$\mathcal{F}^2 = 2\pi \times \text{reflection}, \quad \mathcal{F}^4 = 4\pi^2 \times \text{identity}.$$

What are the possible eigenvalues of $\mathcal{F}$ on $\mathcal{S}$? Do all possibilities occur?

9.6.3. Prove that differentiation and multiplication by x are continuous linear operations on $\mathcal{S}$: if $\phi_j \to \phi$ in $\mathcal{S}$, then $D\phi_j \to D\phi$ and $x\phi_j \to x\phi$ in $\mathcal{S}$.

9.7. Application: the linear harmonic oscillator

The linear harmonic oscillator in quantum mechanics leads to the following eigenvalue problem:

$$\mathcal{H}y \equiv (x^2 - D^2)y = \lambda y, \quad y \in L^2(\mathbb{R}); \tag{9.7.1}$$

cf. the article [**97**]. The condition $y \in L^2(\mathbb{R})$ is a boundary condition at $\pm\infty$. It comes from the fact that $|y|^2$ is a probability density: $\int_a^b |y(x)|^2 dx$ represents the probability to find the oscillating particle on the interval (a, b) when the energy is equal to λ. The values of λ for which the problem has a nonzero solution y represent the possible energy levels of the particle in quantum mechanics. Thus one expects the eigenvalues to be real and positive.

We begin by making it plausible that the eigenfunctions belong to the class $\mathcal{S}$. The solution of the differential equation

$$y'' = (x^2 - \lambda)y \tag{9.7.2}$$

will be of class $\mathcal{C}^\infty$. Indeed, if a solution y is locally integrable, so is y'', hence y will be an indefinite integral of order two, etc. [In fact, by Proposition 8.1.2 the solutions will be analytic on $\mathbb{R}$.]

How do the solutions behave at $+\infty$? Multiplying equation (9.7.2) by $2y'$, one finds

$$2y'y'' = (x^2 - \lambda)2yy',$$

hence by integration one expects

$$\{y'(x)\}^2 = \int^x (t^2 - \lambda)dy^2(t) \approx (x^2 - \lambda)y^2(x)$$

as $x \to \infty$. This gives

$$\frac{y'}{y}(x) \approx \pm x\left(1 - \frac{\lambda}{x^2}\right)^{\frac{1}{2}} \approx \pm x \mp \frac{\lambda}{2x},$$

and hence

$$\log y \approx \pm\frac{1}{2}\,x^2 \mp \frac{1}{2}\,\lambda \log x, \quad y \approx e^{\pm\frac{1}{2}x^2} x^{\mp\frac{1}{2}\lambda}. \tag{9.7.3}$$

More precisely, one expects the differential equation to have a 'large' solution, one that behaves like $e^{\frac{1}{2}x^2} x^{-\frac{1}{2}\lambda}$ as $x \to +\infty$, and a 'small' solution that behaves like $e^{-\frac{1}{2}x^2}x^{\frac{1}{2}\lambda}$ at $+\infty$. (The above argument can be made rigorous; cf. Korevaar [**67**].)

A similar reasoning applies to $-\infty$. What we need is a solution that becomes small at both ends. Such a solution can be expected only for special values of λ, the eigenvalues of problem (9.7.1). Finally observe that a $\mathcal{C}^\infty$ solution that behaves like

$$e^{-\frac{1}{2}x^2}x^{\frac{1}{2}\lambda} \quad \text{at} \quad \pm\infty$$

would be in $\mathcal{S}$. Thus our eigenvalue problem may be restated in the form

$$\mathcal{H}y \equiv (x^2 - D^2)y = \lambda y \quad \text{on } \mathbb{R}, \quad y \in \mathcal{S}. \tag{9.7.4}$$

Without prior knowledge of Hermite functions, this eigenvalue problem may be solved by the so-called *factorization method*. Observe that on $\mathcal{S}$,

$$\mathcal{H}y = \{(x-D)(x+D)+1\}y = \{(x+D)(x-D)-1\}y.$$

Thus the operator $\mathcal{H}$ may be "factored" as follows:

$$\mathcal{H} = (x-D)(x+D)+1 = (x+D)(x-D)-1,$$

where 1 now stands for the identity operator. Using the inner product of $L^2(\mathbb{R})$, integration by parts shows that on $\mathcal{S}$,

$$\begin{aligned}
&([x+D]f, g) = (xf, g) + (Df, g) = (f, xg) - (f, Dg) = (f, [x-D]g),\\
&([x-D]f, g) = (f, [x+D]g).
\end{aligned}$$

Suppose now that γ, ϕ is a characteristic pair of problem (9.7.4). Then

$$\begin{aligned}
([x+D]\phi, [x+D]\phi) &= (\phi, [x-D][x+D]\phi) = (\phi, [\mathcal{H}-1]\phi)\\
&= (\overline{\gamma}-1)(\phi, \phi),\\
([x-D]\phi, [x-D]\phi) &= (\overline{\gamma}+1)(\phi, \phi).
\end{aligned}$$

It follows that all possible eigenvalues γ of $\mathcal{H}$ must be real and ≥ 1. Moreover, if 1 is an eigenvalue, the corresponding eigenfunctions ϕ must satisfy the equation $(x+D)y = 0$, or $y' = -xy$.

Conclusion: 1 *is indeed* an eigenvalue, with corresponding eigenfunctions $\beta e^{-\frac{1}{2}x^2}$ $(\beta \neq 0)$.

Continuing with an arbitrary characteristic pair γ, ϕ, one finds that

$$\begin{aligned}\mathcal{H}(x+D)\phi &= \{(x+D)(x-D)-1\}(x+D)\phi \\ &= (x+D)\{(x-D)(x+D)-1\}\phi \\ &= (x+D)(\mathcal{H}-2)\phi = (\gamma-2)(x+D)\phi, \\ \mathcal{H}(x-D)\phi &= (\gamma+2)(x-D)\phi.\end{aligned}$$

If $\gamma > 1$ then $(x+D)\phi \neq 0$ and hence the pair $\gamma - 2$, $(x+D)\phi$ is also a characteristic pair. In that case $\gamma \geq 3$, and *either* $\gamma = 3$ and $(x+D)^2\phi = 0$, *or* $\gamma > 3$ and $(x+D)^2\phi \neq 0$, in which case $\cdots$. Continuing, one finds that γ must have the form $2n+1$ for some $n \in \mathbb{N}_0$ and then $(x+D)^{n+1}\phi = 0$.

Another conclusion is that $\gamma+2$, $(x-D)\phi$ is a characteristic pair whenever γ, ϕ is one. Thus the characteristic pairs of problem (9.7.4) are the pairs

$$2n+1, \quad \beta_n(x-D)^n e^{-\frac{1}{2}x^2} \quad \text{with } n \in \mathbb{N}_0 \text{ and } \beta_n \neq 0. \tag{9.7.5}$$

With all this information it is not difficult to verify that eigenfunctions belonging to different eigenvalues are orthogonal to each other. From this one may derive that $(x-D)^n e^{-\frac{1}{2}x^2}$ must be equal to $H_n(x)e^{-\frac{1}{2}x^2}$, where $H_n(x)$ is the Hermite polynomial of degree n introduced in Definition 7.3.1.

Remark 9.7.1. The factorization method, also called ladder method, goes back to Dirac [**23**]. It was extended to a variety of classical eigenvalue problems by the physicists Leopold Infeld (Poland–Canada, 1898–1968; [**55**]) and Tom E. Hull (Canada, 1922–1996; [**53**]); see [**56**].

Exercises. 9.7.1. Let ϕ be an eigenfunction of problem (9.7.4) belonging to the eigenvalue $2k+1$. Show that $(x+D)^{k+1}\phi = 0$. Next taking $n > k$, show that $\psi = (x-D)^n e^{-\frac{1}{2}x^2}$ is orthogonal to ϕ. Use this fact to derive that ψ must be equal to $H_n(x)e^{-\frac{1}{2}x^2}$.

9.7.2. First show for $n = 0$ and then for $n \in \mathbb{N}$ that

$$H_n(x)e^{-x^2} = (2i)^n \pi^{-\frac{1}{2}} \int_{\mathbb{R}} e^{-t^2} t^n e^{-2ixt} dt.$$

9.8. More applications in mathematical physics

We continue with a few applications to boundary value problems for partial differential equations. Fourier transformation is a very powerful tool for problems involving (practically) infinite media. It is standard procedure to begin by applying the rules without worrying about questions of existence or convergence. One thus tries to arrive at a plausible answer. In the end, the *answer* should of course be *verified*.

EXAMPLE 9.8.1. (*Heat equation*) What can one say about the temperature distribution $u(x) = u(x,t)$ at time $t > 0$ in an "infinite medium" [for example, a very thick wall]? We assume that there is heat transport only in the X-direction, and that the initial temperature distribution is known. The problem is to solve the one-dimensional heat equation

$$u_{xx} = u_t, \quad -\infty < x < \infty, \;\; t > 0,$$

subject to the initial condition $u(x,0) = f(x)$, $-\infty < x < \infty$.

Introducing the Fourier transform of u relative to x,

$$v(\xi,t) = \mathcal{F}^x[u(x,t)](\xi) = \int_{\mathbb{R}} u(x,t)e^{-i\xi x}dx, \tag{9.8.1}$$

where t is treated as a parameter, we obtain

$$\begin{aligned}&\mathcal{F}^x[u_{xx}(x,t)](\xi) = (i\xi)^2 v(\xi,t), \quad \mathcal{F}^x[u_t(x,t)](\xi) = v_t(\xi,t),\\ &\mathcal{F}^x[u(x,0)](\xi) = v(\xi,0) = \widehat{f}(\xi), \quad \xi \in \mathbb{R}, \;\; t > 0.\end{aligned}$$

The transformation rules for the partial derivatives correspond to differentiation under an integral sign, and this is permitted if $u(x,t)$ is a nice enough function.

Thus by Fourier transformation, our problem takes the simpler form

$$v_t(\xi,t) = -\xi^2 v(\xi,t), \quad v(\xi,0) = \widehat{f}(\xi), \quad \xi \in \mathbb{R},\, t > 0.$$

We now have an ordinary differential equation for v as a function of t in which ξ occurs as a parameter! The solution of the new initial value problem is

$$v(\xi,t) = v(\xi,0)e^{-\xi^2 t} = \widehat{f}(\xi)e^{-\xi^2 t}.$$

We know already which function $u(x) = u(x,t)$ has the final product as its Fourier transform: by Example 9.4.4, for $t > 0$,

$$\begin{aligned} u(x,t) &= \frac{1}{2\sqrt{\pi t}} \int_{\mathbb{R}} f(y)e^{-(x-y)^2/(4t)}dy \\ &= \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} f\left(x - 2\sqrt{t}\,w\right) e^{-w^2}dw. \end{aligned} \tag{9.8.2}$$

Verification. Assuming f locally integrable and bounded on $\mathbb{R}$, the function u in (9.8.2) will satisfy the heat equation for $t > 0$ [use the first integral].

Also, for bounded continuous f, $u(x,t) \to f(x)$ as $t \searrow 0$, uniformly on every bounded interval $-A \le x \le A$ [use the second integral]. It will follow that

$$u(x,t) \to f(x_0) \quad \text{as} \;\; (x,t) \to (x_0, 0), \quad \text{that is,} \;\; x \to x_0, \; t \searrow 0.$$

EXAMPLE 9.8.2. (*Wave equation*) What can one say about the displacements $u(x) = u(x,t)$ at time t in an "infinite vibrating medium", assuming that the displacements are only in the "vertical" direction? We assume for simplicity that the displacements at time $t = 0$ are known, and that the velocities at that instant are equal to zero. The problem then is to solve the one-dimensional wave equation

$$u_{xx} = \frac{1}{c^2}\, u_{tt}, \;\; -\infty < x < \infty, \; t > 0 \quad (\text{or } -\infty < t < \infty),$$

subject to the "initial conditions" $u(x,0) = f(x)$, $u_t(x,0) = 0$, $-\infty < x < \infty$.

Introducing the Fourier transform v of u relative to x as in (9.8.1), our problem takes the simpler form

$$\begin{aligned} &v_{tt}(\xi,t) + c^2\xi^2 v(\xi,t) = 0, \\ &v(\xi,0) = \widehat{f}(\xi), \;\; v_t(\xi,0) = 0, \;\; -\infty < \xi < \infty, \; t > 0. \end{aligned}$$

The solution of the new problem is

$$v(\xi,t) = v(\xi,0)\cos c\,\xi t = \frac{1}{2}\,\widehat{f}(\xi)e^{ic\xi t} + \frac{1}{2}\,\widehat{f}(\xi)e^{-ic\xi t}.$$

By rule (ii) in the table of Section 9.3 the corresponding function u is

$$u(x,t) = \frac{1}{2}\, f(x+ct) + \frac{1}{2}\, f(x-ct). \tag{9.8.3}$$

Physically speaking, the candidate solution (9.8.3) makes sense for arbitrary (continuous) f, both for $t > 0$ and for $t < 0$. The displacement function $u(x)$ at time t appears as the superposition of two disturbances, two "waves"; the first moves to the left with velocity c, the second to the right with velocity c. [The displacement $\frac{1}{2}\, f(x-ct)$ remains equal to $\frac{1}{2}\, f(x_0)$ if $x - ct = x_0$ or $x = x_0 + ct$.] For $t = 0$ the two waves jointly produce the given displacements $u(x,0) = f(x)$ and (formally) the velocities $u_t(x,0) = 0$. However, mathematically speaking there is some question whether the physically acceptable function (9.8.3) satisfies the wave equation in more than just a formal way.

For the above and other reasons it is desirable to develop a more general theory of derivatives, convergence and Fourier transformation which can justify the various formal operations. The theory of *tempered distributions* will provide an appropriate framework; see Chapter 10.

Exercises. 9.8.1. (*Wave equation*) Use Fourier transformation to treat the initial value problem

$$\begin{aligned} &u_{xx} = \frac{1}{c^2}\, u_{tt}, \quad -\infty < x < \infty,\ t > 0, \\ &u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad -\infty < x < \infty. \end{aligned}$$

Draw pictures of the solution at time t, (i) if $f(x) = \Delta_1(x)$, $g(x) = 0$, and (ii) if $f(x) = 0$, $g(x) = \sigma_1(x)$, where Δ_1 and σ_1 are as in Exercises 9.1.3, 9.1.2.

9.8.2. Let $u(x,t)$ be the function given by (9.8.2) with bounded continuous f. Prove that $u(x,t) \to f(x)$ uniformly for $-A \le x \le A$ as $t \searrow 0$.

9.8.3. (*Dirichlet problem for upper half-plane*) Use Fourier transformation to treat the boundary value problem

$$\begin{aligned} &\Delta u \equiv u_{xx} + u_{yy} = 0 \quad \text{on } \ H = \{(x,y) \in \mathbb{R}^2 : -\infty < x < \infty,\, y > 0\}, \\ &u(x,0) = f(x), \quad -\infty < x < \infty. \end{aligned}$$

(i) Determine the general form of $v(\xi, y) = \mathcal{F}^x[u(x,y)](\xi)$, $(\xi, y) \in H$. It is reasonable to require boundedness of the "potential" u on H, or to require finiteness of the "energy" $\int_H \left(u_x^2 + u_y^2\right) dxdy$, in which case the integral $\int_H \left(\xi^2 |v|^2 + |v_y|^2\right) d\xi dy$ must also be finite; cf. Section 10.2 below.

(ii) Determine $v(\xi, y)$ under the additional condition that v must remain bounded as $y \to \infty$. [Consider $\xi > 0$ and $\xi < 0$ separately.]

(iii) Show that the corresponding function u is given by the *Poisson integral* for the upper half-plane,

$$\begin{aligned} u(x,y) &\stackrel{\text{def}}{=} \int_{\mathbb{R}} f(t)\, \frac{1}{\pi}\, \frac{y}{(x-t)^2 + y^2}\, dt \\ &= \int_{\mathbb{R}} f(x - yw)\, \frac{1}{\pi}\, \frac{1}{1+w^2}\, dw, \quad y > 0. \end{aligned}$$

(iv) Verify that $u(x,y)$ in (iii) satisfies Laplace's equation whenever f is bounded on $\mathbb{R}$ and locally integrable.

(v) Taking f bounded and continuous, prove that $u(x,y) \to f(x_0)$ as $x \to x_0$, $y \searrow 0$.

(vi) Determine $u(x, y)$ explicitly if $f(x) = 1$ for $a < x < b$, $f(x) = 0$ for $x < a$ and for $x > b$.

CHAPTER 10

Generalized functions of slow growth: tempered distributions

It is useful to extend the theory of Fourier integrals beyond the class of well-behaved functions that are integrable over the whole line. Also, in order to facilitate the use of Fourier theory in applications, the rules in Section 9.3 should be widely applicable. It would in particular be desirable that differentiation should be always possible. Following Laurent Schwartz [**110**], we show that such ends can be achieved. The theory uses an operational definition of Fourier transformation as suggested by Proposition 9.4.1: the action of the transform $\mathcal{F}f$ on "test functions" ϕ shall be the same as the action of f on the transform $\mathcal{F}\phi$. We will employ Schwartz's test-function space $\mathcal{S}$ described in Section 9.6.

If one limits oneself to L^1 functions on $\mathbb{R}$ and their transforms, or to L^2 functions, the operational definition succeeds within the class of locally integrable functions. However, for a really general theory one has to admit a larger class of objects, the so-called *tempered distributions*. These are defined as continuous linear functionals on the test space $\mathcal{S}$. More concretely, tempered distributions turn out to be locally integrable functions of at most polynomial growth, together with their generalized derivatives of any order. The class $\mathcal{S}'$ of tempered distributions is closed under multiplication by x and differentiation. It will also be closed under Fourier transformation; see Chapter 11.

In our development of Fourier theory, the Hermite functions h_n [Definition 7.3.5] will play a special role; cf. Korevaar [**67**]. For the case of L^2, the use of Hermite functions goes back to Wiener [**124**].

10.1. Initial Fourier theory for the class $\mathcal{P}$

Functions of (at most) *polynomial growth* on $\mathbb{R}$ frequently occur in applications. We begin with a limited Fourier theory for such functions; the general theory will come later.

DEFINITION 10.1.1. We say that a locally integrable function f on $\mathbb{R}$ is of class $\mathcal{P}$ if there is an integer $q \geq 0$ such that

$$\frac{f(x)}{(x+i)^q} \quad \text{is in} \quad L^1(\mathbb{R}). \tag{10.1.1}$$

Equality in $\mathcal{P}$ shall mean equality almost everywhere on $\mathbb{R}$.

Functions in $\mathcal{P}$ are uniquely determined by their action on test functions:

PROPOSITION 10.1.2. *Let f in $\mathcal{P}$ be such that $\int_{\mathbb{R}} f\phi = 0$ for all functions ϕ in $\mathcal{S}$. Then $f = 0$.*

PROOF. It will be enough to use the Hermite functions $\phi = h_n$ of Section 7.3. Indeed, suppose that f has Hermite expansion 0:

$$c_n[f] \stackrel{\text{def}}{=} \int_{\mathbb{R}} f h_n = \rho_n \int_{\mathbb{R}} f(x) H_n(x) e^{-\frac{1}{2}x^2} dx = 0, \quad n = 0, 1, 2, \cdots. \tag{10.1.2}$$

Then

$$\int_{\mathbb{R}} f(x) x^n e^{-\frac{1}{2}x^2} dx = 0, \quad \forall\, n \in \mathbb{N}_0.$$

Thus one may apply Moment Theorem 9.5.2 to conclude that $f = 0$. [The function f in that theorem should then be taken equal to the present f times $e^{-\frac{1}{2}x^2}$.] □

We now define a generalized Fourier transform for certain functions in $\mathcal{P}$ in accordance with Proposition 9.4.1.

DEFINITION 10.1.3. A function g in $\mathcal{P}$ is called the Fourier transform $\mathcal{F}f = \widehat{f}$ of the function f in $\mathcal{P}$ if

$$\int_{\mathbb{R}} g\phi = \int_{\mathbb{R}} f\mathcal{F}\phi, \quad \forall\, \phi \in \mathcal{S}.$$

Similarly for the reflected Fourier transform $h = \mathcal{F}_R f$: $\int_{\mathbb{R}} h\phi = \int_{\mathbb{R}} f\mathcal{F}_R\phi$, $\forall\, \phi$.

If $\mathcal{F}f = g$ exists in $\mathcal{P}$, it is unique, and $\mathcal{F}_R f = g_R$:

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{F}_R f \cdot \phi &= \int_{\mathbb{R}} f\mathcal{F}_R\phi = \int_{\mathbb{R}} f\mathcal{F}\phi_R \\ &= \int_{\mathbb{R}} \mathcal{F}f \cdot \phi_R = \int_{\mathbb{R}} g\phi_R = \int_{\mathbb{R}} g_R\phi, \quad \forall\, \phi. \end{aligned} \tag{10.1.3}$$

As a consequence of the definition we have general validity of Fourier inversion:

THEOREM 10.1.4. *Suppose f in $\mathcal{P}$ has Fourier transform g in $\mathcal{P}$. Then g has a reflected Fourier transform in $\mathcal{P}$, and $f = (1/2\pi)\mathcal{F}_R g$.*

Indeed, $h = \mathcal{F}_R g = \mathcal{F}_R \mathcal{F} f$ should be a function in $\mathcal{P}$ such that

$$\int_{\mathbb{R}} h\phi = \int_{\mathbb{R}} \mathcal{F}_R \mathcal{F} f \cdot \phi = \int_{\mathbb{R}} \mathcal{F} f \cdot \mathcal{F}_R \phi = \int_{\mathbb{R}} f \cdot \mathcal{F}\mathcal{F}_R \phi, \quad \forall\, \phi \in \mathcal{S}.$$

However, by the Inversion Theorem 9.6.2 for $\mathcal{S}$, the last integral is equal to $\int_{\mathbb{R}} f \cdot 2\pi\phi$, so that $h = 2\pi f$ or $f = (1/2\pi)h = (1/2\pi)\mathcal{F}_R g$.

Convergence relative to test functions is defined as one would expect, cf. Definition 4.1.3:

DEFINITION 10.1.5. Functions f_λ in $\mathcal{P}$ converge to f in $\mathcal{P}$ relative to the test class $\mathcal{S}$ as $\lambda \to \lambda_0$, and [in accordance with later notation] we write

$$\mathcal{S}' \lim f_\lambda = f \quad \text{as} \;\; \lambda \to \lambda_0,$$

if

$$\int_{\mathbb{R}} f_\lambda \phi \to \int_{\mathbb{R}} f\phi, \quad \forall\, \phi \in \mathcal{S}.$$

THEOREM 10.1.6. *Suppose f in $\mathcal{P}$ has Fourier transform g in $\mathcal{P}$. Then*

$$g(\xi) = (\mathcal{F} f)(\xi) = \mathcal{S}' \lim_{A\to\infty} \int_{-A}^{A} f(x) e^{-i\xi x} dx,$$

and conversely,

$$f(x) = \frac{1}{2\pi} (\mathcal{F}_R g)(x) = \frac{1}{2\pi} \cdot \mathcal{S}' \lim_{A\to\infty} \int_{-A}^{A} g(\xi) e^{ix\xi} d\xi.$$

Indeed, introducing the *truncated* function

$$f_A(x) = \begin{cases} f(x) & \text{for } |x| < A \\ 0 & \text{for } |x| > A \end{cases}$$

[not to be confused with the reflection $f_R(x) = f(-x)$], we have

$$\int_{-A}^{A} f(x) e^{-i\xi x} dx = (\mathcal{F} f_A)(\xi).$$

Hence by Proposition 9.4.1 for L^1 functions,

$$\begin{aligned} &\int_{\mathbb{R}} \mathcal{F} f_A \cdot \phi = \int_{\mathbb{R}} f_A \cdot \mathcal{F}\phi = \int_{-A}^{A} f \cdot \mathcal{F}\phi \\ &\to \int_{\mathbb{R}} f \cdot \mathcal{F}\phi = \int_{\mathbb{R}} \mathcal{F} f \cdot \phi = \int_{\mathbb{R}} g\phi, \quad \forall\, \phi \in \mathcal{S}. \end{aligned} \tag{10.1.4}$$

The convergence follows from the integrability of $f \cdot \mathcal{F}\phi$ over $\mathbb{R}$ when f is in $\mathcal{P}$.

The proof in the other direction is similar.

As a corollary we obtain

THEOREM 10.1.7. (Extended inversion theorem for L^1) *In the extended theory, Fourier inversion is valid for every function f in $L^1(\mathbb{R})$. More precisely, setting $\mathcal{F}f = g$, one will have*

$$f(x) = \frac{1}{2\pi}(\mathcal{F}_R g)(x) = \frac{1}{2\pi}\mathcal{S}' \lim_{A\to\infty} \int_{-A}^{A} g(\xi)e^{ix\xi}d\xi.$$

In particular, if g is also in $L^1(\mathbb{R})$, then

$$f(x) = \frac{1}{2\pi}(\mathcal{F}_R g)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} g(\xi)e^{ix\xi}d\xi \quad a.e. \ on \ \mathbb{R}.$$

We need comment only on the last part: if g is in $L^1(\mathbb{R})$, the integrals $\int_{-A}^{A} g(\xi)e^{ix\xi}d\xi$ will converge uniformly to $(\mathcal{F}_R g)(x)$ as $A \to \infty$. The uniform limit will agree with the $\mathcal{S}'$ limit in the sense of $\mathcal{P}$, hence a.e.

Exercises. 10.1.1. Given $f \in \mathcal{P}$ and $\phi \in \mathcal{S}$, prove the existence of

$$\int_{\mathbb{R}} f\phi = \int_{\mathbb{R}} \frac{f(x)}{(x+i)^q}(x+i)^q\phi(x)dx.$$

10.1.2. Prove that L^2 functions f on $\mathbb{R}$ and Fourier transforms $g = \widehat{f}$ of L^1 functions f belong to $\mathcal{P}$.

10.1.3. Prove that locally integrable functions f_λ converge to f relative to $\mathcal{S}$ as $\lambda \to \lambda_0$ if one of the following conditions is satisfied:

(i) $f_\lambda \to f$ in $L^1(\mathbb{R})$,

(ii) $f_\lambda \to f$ in $L^2(\mathbb{R})$,

(iii) $f_\lambda \to f$ uniformly on every bounded interval and $|f_\lambda(x)| \le Q(x)$, a polynomial, $\forall\, x$, λ.

10.1.4. Let p and q be in $\mathbb{N}_0$. Prove that

$$\lambda^p x^q e^{i\lambda x} \to 0 \quad \text{as } \lambda \to \infty$$

relative to the test class $\mathcal{S}$.

10.2. Fourier transformation in $L^2(\mathbb{R})$

Functions f in $L^2(\mathbb{R})$ are in $\mathcal{P}$ [Exercise 10.1.2]. We will prove

PROPOSITION 10.2.1. *For any f in $L^2 = L^2(\mathbb{R})$, the generalized Fourier transform $\widehat{f}$ exists and belongs to L^2.*

PROOF. We will use Hermite series, recalling that the normalized Hermite functions h_n form an orthonormal basis of L^2 [Theorem 9.5.4]. Thus every function f in L^2 is equal to the sum of its Hermite expansion $\sum c_n h_n$, where $c_n = c_n[f] = \int_{\mathbb{R}} f h_n$. Also, a series $\sum d_n h_n$ will converge to an element g in L^2 if and only if the series $\sum |d_n|^2$ converges [see Corollaries 6.3.5].

Suppose now that $f \in L^2$ has a Fourier transform $\widehat{f}$ in the class $\mathcal{P}$. Then by Definition 10.1.3,

$$\int \widehat{f}\phi = \int f\widehat{\phi}, \quad \forall\, \phi \in \mathcal{S}, \tag{10.2.1}$$

where $\int$ stands for $\int_{\mathbb{R}}$ (also below). Taking $\phi = h_n$, we find that

$$\begin{aligned} c_n\left[\widehat{f}\right] &\stackrel{\text{def}}{=} \int \widehat{f} h_n = \int f \widehat{h}_n \\ &= \sqrt{2\pi}(-i)^n \int f h_n = \sqrt{2\pi}(-i)^n c_n[f]; \end{aligned}$$

cf. formula (9.3.4). Thus by Parseval's formula for the Hermite coefficients $c_n[f]$, cf. Theorem 6.3.1,

$$\sum \left| c_n\left[\widehat{f}\right] \right|^2 = 2\pi \sum |c_n[f]|^2 = 2\pi \int |f|^2. \tag{10.2.2}$$

It follows that there is a function g in L^2 (and hence in $\mathcal{P}$) with Hermite series $\sum c_n \left[\widehat{f}\right] h_n$:

$$g \stackrel{\text{def}}{=} \sum c_n\left[\widehat{f}\right] h_n = \sum \sqrt{2\pi}(-i)^n c_n[f] h_n. \tag{10.2.3}$$

We will show that this function g is indeed the Fourier transform $\widehat{f}$ of f in the sense of (10.2.1). To that end we apply the extended Parseval formula [Theorem 6.3.1]:

$$\begin{aligned} \int g\phi &= \lim_{k\to\infty} \int \sum_{n=0}^{k} c_n[g] h_n \cdot \phi = \lim \sum_{0}^{k} c_n[g] c_n[\phi] \\ &= \sum_{0}^{\infty} c_n[g] c_n[\phi] = \sum \sqrt{2\pi}\,(-i)^n c_n[f] c_n[\phi] \\ &= \sum c_n[f] c_n\left[\widehat{\phi}\right] = \int f\widehat{\phi}, \quad \forall\, \phi \in \mathcal{S}, \end{aligned}$$

hence $g = \widehat{f}$. □

Observe also that by (10.2.3),

$$\int |g|^2 = \sum \left| c_n \left[\widehat{f} \right] \right|^2 = 2\pi \sum |c_n[f]|^2 = 2\pi \int |f|^2.$$

We thus obtain the *Parseval formula* for Fourier transformation:

$$\int_{\mathbb{R}} \left| \widehat{f} \right|^2 = \int_{\mathbb{R}} |g|^2 = 2\pi \int_{\mathbb{R}} |f|^2. \tag{10.2.4}$$

Corollary 10.2.2. *Fourier transformation defines a one to one continuous linear map of L^2 onto itself.*

Indeed, if $f \in L^2$ and $g = \mathcal{F}f$, then $f = (1/2\pi)\mathcal{F}_R g = (1/2\pi)\mathcal{F} g_R$ [Theorem 10.1.4]. The continuity follows from (10.2.4): if $f - f_\lambda \to 0$ in L^2, then $\widehat{f} - \widehat{f}_\lambda \to 0$ in L^2.

Fourier transformation on L^2 is almost an *isometry*. In fact, under our transformation $\mathcal{F}$, all norms and distances are multiplied by $\sqrt{2\pi}$. If one would redefine Fourier transformation as $\mathcal{F}^* = (1/\sqrt{2\pi})\mathcal{F}$:

$$(\mathcal{F}^* f)(\xi) = \mathcal{S}' \lim_{A\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-A}^{A} f(x) e^{-i\xi x} dx,$$

then the transformation would preserve all norms and distances in L^2.

We finally remark that for f in L^2, the limits in Theorem 10.1.6 will exist in the much stronger sense of L^2. The following result was proved by the Swiss mathematician Michel Plancherel (1885–1967; [**91**]), cf. [**92**]:

Theorem 10.2.3. (Plancherel's theorem) *Let f be in $L^2(\mathbb{R})$ and let $f_A(x) = f(x)$ for $|x| < A$, $f_A(x) = 0$ for $|x| > A$. Then one may define a (the) Fourier transform of f by the formula*

$$g(\xi) = (\mathcal{F}f)(\xi) = \lim_{A\to\infty} (\mathcal{F}f_A)(\xi) = \lim_{A\to\infty} \int_{-A}^{A} f(x) e^{-i\xi x} dx,$$

where the limit is taken in the sense of $L^2(\mathbb{R})$. Conversely,

$$f(x) = \frac{1}{2\pi} (\mathcal{F}_R g)(x) = \frac{1}{2\pi} \lim_{A\to\infty} \int_{-A}^{A} g(\xi) e^{ix\xi} d\xi,$$

also in the sense of L^2. The functions f and g satisfy the Parseval formula (10.2.4).

PROOF. Let f be in L^2 and let g be its Fourier transform $\widehat{f}$ in the sense of Proposition 10.2.1; similarly $g^A \stackrel{\text{def}}{=} \widehat{f}_A$. We know already that f and g satisfy the Parseval formula (10.2.4), and so will $f - f_A$ and $g - g^A$. Thus

$$\frac{1}{2\pi} \int \left|g - g^A\right|^2 = \int |f - f_A|^2 = \left(\int_{-\infty}^{-A} + \int_A^{\infty}\right) |f|^2.$$

The right-hand side will tend to zero as $A \to \infty$ since $|f|^2$ is integrable over $\mathbb{R}$.

The proof that $\mathcal{F}_R g = \lim \mathcal{F}_R g_A$ is similar. □

EXAMPLES 10.2.4. The function $f(x) = (1/2)\sigma_1(x)$ has Fourier transform $g(\xi) = (\sin \xi)/\xi$; see Exercise 9.1.2. Hence by the Parseval formula,

$$\int_{\mathbb{R}} \frac{\sin^2 \xi}{\xi^2}\, d\xi = 2\pi \int_{\mathbb{R}} (1/4)\sigma_1^2(x) dx = 2\pi \int_{-1}^{1} (1/4) dx = \pi.$$

For $f(x) = 1/(x + i)$ one may use the Residue Theorem to show that

$$\lim_{A\to\infty} \int_{-A}^{A} \frac{1}{x+i}\, e^{-i\xi x} dx = -2\pi i e^{-\xi} U(\xi) \tag{10.2.5}$$

for all $\xi \neq 0$. Indeed, let C_A with $A > 1$ denote the semi-circle $z = Ae^{i\theta}$, $0 \leq \theta \leq \pi$ [with center 0 and radius A] in the upper half-plane if $\xi < 0$, and the semi-circle $z = Ae^{i\theta}$, $0 \geq \theta \geq -\pi$ in the lower half-plane if $\xi > 0$. Then

$$\left(\int_{-A}^{A} + \int_{C_A}\right) \frac{1}{z+i}\, e^{-i\xi z} dz = \begin{cases} 0 & \text{if } \xi < 0, \\ -2\pi i\, e^{-\xi} & \text{if } \xi > 0; \end{cases}$$

cf. Example 1.7.1. Careful estimates show that as $A \to \infty$,

$$\rho_A(\xi) \stackrel{\text{def}}{=} \int_{C_A} \frac{1}{z+i}\, e^{-i\xi z} dz \tag{10.2.6}$$

tends to zero pointwise for $\xi \in \mathbb{R} \setminus 0$, as well as boundedly. Hence by Lebesgue's Theorem on Dominated Convergence, $\rho_A(\xi) \to 0$ relative to test functions. Conclusion: the limit in (10.2.5) exists relative to test functions, and it has the value given there.

We verify the Parseval formula for the present Fourier pair:

$$\int_{\mathbb{R}} \left|2\pi i e^{-\xi} U(\xi)\right|^2 d\xi = 4\pi^2 \int_0^{\infty} e^{-2\xi} d\xi = 2\pi^2,$$

$$2\pi \int_{\mathbb{R}} \left|\frac{1}{x+i}\right|^2 dx = 2\pi \int_{\mathbb{R}} \frac{1}{x^2+1}\, dx = 2\pi^2.$$

Exercises. 10.2.1. Taking α, β real, $\beta > 0$, compute the Fourier transforms of

$$\frac{1}{x - i\beta}, \quad \frac{1}{x + i\beta}, \quad \frac{x}{x^2 + \beta^2}, \quad \frac{1}{x - \alpha - i\beta}.$$

10.2.2. Use the Parseval formula to compute

$$\int_{\mathbb{R}} \frac{\sin^4 \xi}{\xi^4} \, d\xi.$$

10.2.3. Prove the L^2 convergence in the second formula of Plancherel's theorem.

10.2.4. One can show that $\rho_A(\xi)$ in (10.2.6) is the L^2 Fourier transform of the function $f_A^*(x) = f(x) - f_A(x)$ that is equal to $1/(x+i)$ for $|x| > A$ and equal to 0 for $|x| < A$. Deduce that

$$\int_{\mathbb{R}} |\rho_A(\xi)|^2 d\xi = 4\pi\{(\pi/2) - \arctan A\},$$

a quantity that tends to 0 as $A \to \infty$.

10.2.5. Let f be in $L^2(\mathbb{R})$. Prove Plancherel's formulas

$$\begin{aligned}
\widehat{f}(\xi) &= \frac{d}{d\xi} \int_{\mathbb{R}} f(x) \, \frac{e^{-i\xi x} - 1}{-ix} \, dx \\
&= \int_{-1}^{1} f(x) e^{-i\xi x} + \frac{d}{d\xi} \int_{|x|>1} f(x) \, \frac{e^{-i\xi x} - 1}{-ix} \, dx
\end{aligned}$$

for almost all $\xi \in \mathbb{R}$, and similarly for $f(x)$ in terms of $\widehat{f}(\xi)$.

Hint. Compute $\int_0^\xi \widehat{f}_A(t) dt$ and let $A \to \infty$.

10.3. Hermite series for test functions

In our study of tempered distributions, Hermite series play a role analogous to Fourier series in the case of periodic distributions (distributions on the unit circle), cf. Chapter 4. As preparation for the general theory we will characterize the functions ϕ in $\mathcal{S}$ [Section 9.6] by their Hermite series.

Recall that the normalized Hermite functions have the form

$$h_n(x) = \rho_n H_n(x) e^{-\frac{1}{2}x^2} = \rho_n (x - D)^n e^{-\frac{1}{2}x^2}, \quad n \in \mathbb{N}_0,$$

where $\rho_n = 2^{-\frac{1}{2}n} (n!)^{-\frac{1}{2}} \pi^{-\frac{1}{4}}$ [Section 7.3]. They satisfy the relations

$$\begin{aligned}
&(x + D) h_n = \sqrt{2n} \, h_{n-1}, \quad (x - D) h_n = \sqrt{2n+2} \, h_{n+1}, \\
&\mathcal{H} h_n \equiv (x^2 - D^2) h_n = (2n+1) h_n.
\end{aligned}$$

By combination one finds that

$$xh_n = \sqrt{(n/2)}\, h_{n-1} + \sqrt{(n+1)/2}\, h_{n+1}, \tag{10.3.1}$$

$$Dh_n = \sqrt{(n/2)}\, h_{n-1} - \sqrt{(n+1)/2}\, h_{n+1}. \tag{10.3.2}$$

LEMMA 10.3.1. *There are absolute constants* α *and* C [for example, $\alpha = 1$ and $C = 6$] *such that*

$$|h_n(x)| \le Cn^\alpha, \quad \forall\, x \in \mathbb{R},\ \forall\, n \in \mathbb{N}. \tag{10.3.3}$$

PROOF. Since $|h_n(x)|$ is even we may take $x \ge 0$. If n is odd, $h_n(0) = 0$ while for even n, using (10.3.1) with $x = 0$ and $n-1$ instead of n,

$$|h_n(0)| \le |h_0(0)| = \pi^{-\frac{1}{4}} < 1.$$

Integration of (10.3.2) [with $n = k \ge 1$] from 0 to x, application of Cauchy–Schwarz and the relation $\int_0^\infty h_j^2 = 1/2$ now give

$$\begin{aligned} |h_k(x)| &\le |h_k(0)| + \sqrt{(k/2)} \int_0^x |h_{k-1}| + \sqrt{(k+1)/2} \int_0^x |h_{k+1}| \\ &< 1 + (1/2)\sqrt{k}\sqrt{x} + (1/2)\sqrt{k+1}\sqrt{x} < \begin{cases} 3\sqrt{k} & \text{for } 0 \le x \le 1, \\ 3\sqrt{kx} & \text{for } x > 1. \end{cases} \end{aligned}$$

Combination with (10.3.1) for $x > 1$ gives (10.3.3) with $\alpha = 1$. □

[Remark. Using more sophisticated tools one can prove an inequality (10.3.3) with $\alpha = 0$, and even $\alpha = -1/12$; see [**117**], formula (8.19.10).]

COROLLARY 10.3.2. *There are absolute constants* C_{pq} *such that*

$$\left|x^p D^q h_n(x)\right| \le C_{pq} n^{\frac{1}{2}p + \frac{1}{2}q + 1}, \quad \forall\, x \in \mathbb{R},\ \forall\, n \in \mathbb{N}. \tag{10.3.4}$$

[Use (10.3.1)–(10.3.3) with $\alpha = 1$.]

THEOREM 10.3.3. (Characterization of $\mathcal{S}$ by Hermite series) *Let* ϕ *be a continuous* L^2 *function on* $\mathbb{R}$. *Then* ϕ *is in* $\mathcal{S}$ *if and only if for every nonnegative integer* r*, there is a constant* $B_r = B_r(\phi)$ *such that the Hermite coefficients* $c_n[\phi] = \int \phi h_n$ *satisfy the inequalities*

$$|c_n[\phi]| \le \frac{B_r}{n^r}, \quad \forall\, n \in \mathbb{N}_0. \tag{10.3.5}$$

PROOF. (i) Let ϕ be in $\mathcal{S}$. Then $\mathcal{H}^r\phi = (x^2 - D^2)^r\phi$ is also in $\mathcal{S}$ [Section 9.6], and hence in L^2. It follows that $\sum_0^\infty \left|c_n[\mathcal{H}^r\phi]\right|^2 < \infty$, and thus

$$\left|c_n[\mathcal{H}^r\phi]\right| \le C = C_r(\phi), \quad \forall\, n \in \mathbb{N}_0.$$

Now for any $\psi \in \mathcal{S}$,

$$\int \mathcal{H}\psi \cdot h_n = \int (x^2 - D^2)\psi \cdot h_n = \int \psi \cdot \mathcal{H}h_n = (2n+1) \int \psi h_n.$$

Hence

$$c_n[\mathcal{H}^r \phi] = (2n+1)^r c_n[\phi],$$

which by the boundedness of the sequence $\{|c_n[\mathcal{H}^r \phi]|\}$ implies (10.3.5).

(ii) Let ϕ be an L^2 function such that (10.3.5) holds for every r. Because of (10.3.4) with $p = 0$, the series

$$\sum c_n[\phi]h_n, \;\; \sum c_n[\phi]Dh_n, \;\; \sum c_n[\phi]D^2 h_n, \;\; \cdots$$

then converge uniformly on $\mathbb{R}$. By the uniform convergence of $\sum c_n[\phi]h_n$ on every interval and its L^2 convergence to ϕ, the sum $\sum c_n[\phi]h_n$ is continuous and a.e. equal to ϕ. If necessary, we modify ϕ on a set of measure zero to make it equal to $\sum c_n[\phi]h_n$ everywhere. It now follows from the uniform convergence of $\sum c_n[\phi]Dh_n$ that ϕ is of class $\mathcal{C}^1(\mathbb{R})$ and that $D\phi = \sum c_n[\phi]Dh_n$. Continuing in this way, one finds that ϕ is of class $\mathcal{C}^q(\mathbb{R})$ and that $D^q\phi = \sum c_n[\phi]D^q h_n$ for every r, hence

$$x^p D^q \phi = \sum c_n[\phi] x^p D^q h_n.$$

Relations (10.3.4) and (10.3.5) with $r > (p/2) + (q/2) + 2$ now show that $x^p D^q \phi$ is bounded on $\mathbb{R}$ for every choice of p and q, so that ϕ is in $\mathcal{S}$. $\square$

Corollary 10.3.4. *For every function ϕ in $\mathcal{S}$, the Hermite series $\sum c_n[\phi]h_n$ converges to ϕ in the sense of $\mathcal{S}$* [Definition 9.6.3].

Indeed, (10.3.4) and (10.3.5) with $r > (p/2) + (q/2) + 2$ imply that for all $x \in \mathbb{R}$,

$$\begin{aligned} \left| x^p D^q (\phi - s_k[\phi]) \right| &= \Big| \sum_{n>k} c_n[\phi] x^p D^q h_n \Big| \\ &\leq B_r C_{pq} \sum_{n>k} \frac{n^{(p/2)+(q/2)+1}}{n^r} \to 0 \quad \text{as} \;\; k \to \infty. \end{aligned}$$

Thus $x^p D^q s_k[\phi] \to x^p D^q \phi$ uniformly on $\mathbb{R}$. Since this holds for all $p, q \in \mathbb{N}_0$ we conclude that $s_k[\phi] \to \phi$ in $\mathcal{S}$.

Exercises. 10.3.1. (*A challenge!*) It would be nice to have a simple proof for the uniform boundedness of the family $\{h_n\}$ on $\mathbb{R}$. Try to determine a constant C such that $|h_n(x)| \leq C$, $\forall\, x$, n.

Hint. Next to the relations (10.3.1), (10.3.2), the Exercises 7.3.7, 7.3.8 and 9.7.2 may be useful.

10.4. Tempered distributions

We are now ready to introduce distributions of slow or polynomial growth on $\mathbb{R}$ – so-called tempered distributions. As before, let $\mathcal{S}$ be the test space of "rapidly decreasing functions with rapidly decreasing derivatives", with the associated notion of convergence [Section 9.6].

DEFINITION 10.4.1. (*Laurent Schwartz*, about 1950; cf. Schwartz [**110**], Hörmander [**52**]) A tempered distribution T on $\mathbb{R}$ is a continuous linear functional on the test space $\mathcal{S}$. Thus in particular

$$< T, \phi_j > \to < T, \phi > \quad \text{whenever} \;\; \phi_j \to \phi \;\text{ in }\; \mathcal{S}.$$

EXAMPLES 10.4.2. (i) Every function f of at most *polynomial growth* on $\mathbb{R}$ [Definition 10.1.1] defines a tempered distribution T_f by the formula

$$< T_f, \phi > = \int_{\mathbb{R}} f\phi, \qquad \forall\, \phi \in \mathcal{S}.$$

Indeed, if $q \geq 0$ is so large that (10.1.1) holds and $\phi_j \to \phi$ in $\mathcal{S}$, then

$$\begin{aligned} \big| < T_f, \phi > \,-\, < T_f, \phi_j > \big| &\leq \int_{\mathbb{R}} \left| \frac{f(x)}{(x+i)^q} \right| \cdot \big|(x+i)^q\{\phi(x) - \phi_j(x)\}\big| dx \\ \leq \int_{\mathbb{R}} \left| \frac{f(x)}{(x+i)^q} \right| dx \cdot \sup_{x\in\mathbb{R}} &\big|(x+i)^q\{\phi(x) - \phi_j(x)\}\big| \to 0 \quad \text{as } \; j \to \infty. \end{aligned}$$

The correspondence $f \leftrightarrow T_f$ is one to one for $f \in \mathcal{P}$ [see Proposition 10.1.2]. We identify T_f with f, and usuallly write $< T_f, \phi >$ simply as $< f, \phi >$.

(ii) Certain functions that are locally integrable, apart from simple singularities, also have representatives in the class of tempered distributions. Thus the function $f(x) = 1/x$ leads to the *principal value distribution* $\mathrm{pv}\, 1/x$ through the formula

$$\begin{aligned} \left\langle \mathrm{pv}\, \frac{1}{x}, \phi(x) \right\rangle &\stackrel{\text{def}}{=} \text{p.v.} \int_{\mathbb{R}} \frac{1}{x}\, \phi(x) dx \stackrel{\text{def}}{=} \lim_{\varepsilon \searrow 0} \int_{|x|>\varepsilon} \frac{1}{x}\, \phi(x) dx \\ &= \lim_{\varepsilon \searrow 0} \left\{ \Big[\phi(x) \log|x|\Big]_{-\infty}^{-\varepsilon} + \Big[\,\cdots\Big]_{\varepsilon}^{\infty} - \int_{|x|>\varepsilon} \log|x| \cdot \phi'(x) dx \right\} \\ &= -\int_{\mathbb{R}} \log|x| \cdot \phi'(x) dx, \qquad \forall\, \phi \in \mathcal{S}; \end{aligned}$$

cf. Example 4.2.6. Other examples may be found in Exercises 11.2.10 and 11.2.11.

(iii) The *delta distribution*, δ, on $\mathbb{R}$ is defined by the formula

$$< \delta, \phi > = \phi(0), \qquad \forall\, \phi \in \mathcal{S}.$$

The translate "$\delta(x-a)$" is denoted by δ_a: $< \delta_a, \phi > = \phi(a)$.

(iv) With the delta distribution δ_Γ on the unit circle Γ we may associate the *periodic delta distribution* $\delta_{2\pi}^{\rm per}$ on $\mathbb{R}$ with period 2π. It is given by the formula

$$\left\langle \delta_{2\pi}^{\rm per}, \phi \right\rangle = \sum_{n=-\infty}^{\infty} \phi(2\pi n), \qquad \forall\, \phi \in \mathcal{S}.$$

Other periods also occur, for example, $\left\langle \delta_1^{\rm per}, \phi \right\rangle = \sum_{-\infty}^{\infty} \phi(n)$. We will verify the continuity of the functional $\delta_{2\pi}^{\rm per}$. Let $\phi_j \to \phi$ in $\mathcal{S}$. Then

$$(x^2+1)|\phi(x) - \phi_j(x)| < \varepsilon \;\text{ for } j > j_0(\varepsilon) \text{ and all } x \in \mathbb{R}.$$

Hence

$$\begin{aligned} \left|\left\langle \delta_{2\pi}^{\rm per}, \phi \right\rangle - \left\langle \delta_{2\pi}^{\rm per}, \phi_j \right\rangle\right| &\leq \sum_{-\infty}^{\infty} |\phi(2\pi n) - \phi_j(2\pi n)| \\ < \varepsilon \sum_{-\infty}^{\infty} \frac{1}{4\pi^2 n^2 + 1} &< \varepsilon \left(1 + 2\sum_{1}^{\infty} \frac{1}{4\pi^2 n^2}\right) = (13/12)\varepsilon, \quad \forall\, j > j_0. \end{aligned}$$

Every distribution on the unit circle may be interpreted as a tempered distribution of period 2π; cf. Exercises 10.6.1, 10.6.2.

Definition 10.4.3. (Simple operations) Linear combinations $\lambda_1 T_1 + \lambda_2 T_2$, translates $T_c(x) = T(x-c)$, the reflection $T_R(x) = T(-x)$, and products $\omega T = T\omega$ of T and $\mathcal{C}^\infty$ functions ω are defined in the obvious manner; cf. Section 4.2. For the definition $< T\omega, \phi > = < T, \omega\phi >$ one has to require that ω and its derivatives ω', ω'', $\cdots$ are bounded by polynomials.

Important is the notion of *equality* $T_1 = T_2$ on an open set $\Omega \subset \mathbb{R}$:

$$(10.4.1) \qquad T = T_1 - T_2 = 0 \quad \text{on } \Omega \quad \text{if} \quad < T, \phi > = 0$$

for all test functions ϕ with support in Ω; cf. Definition 4.2.8. The *support* of T is the smallest closed set outside of which T is equal to zero.

Examples 10.4.4. The distribution δ is even: $\delta(-x) = \delta(x)$; the distribution $\operatorname{pv} 1/x$ is odd. One has $\delta(x) = 0$ on $(0,\infty)$ and on $(-\infty, 0)$: the support of δ is the origin. It follows that δ cannot be equal to a function in $\mathcal{P}$; cf. Section 4.2. Other properties are $x \cdot \delta(x) = 0$ and $x \cdot \operatorname{pv} 1/x = 1$.

DEFINITION 10.4.5. (*Convergence* of tempered distributions; the *space* $\mathcal{S}'$) We say that tempered distributions T_λ converge [or converge weakly] to the tempered distribution T on $\mathbb{R}$ for $\lambda \to \lambda_0$ if

$$< T_\lambda, \phi > \to < T, \phi > \quad \text{as } \lambda \to \lambda_0, \quad \forall\, \phi \in \mathcal{S}.$$

With this notion of convergence the tempered distributions on $\mathbb{R}$ form the space $\mathcal{S}'$, the space *dual* to $\mathcal{S}$.

EXAMPLES 10.4.6. (i) Let f_λ, with $\lambda \to \lambda_0$, and f be functions of the class $\mathcal{P}$ such that $f_\lambda(x) \to f(x)$ for almost all x. Suppose, moreover, that there is a fixed polynomial $Q(x)$ such that $|f_\lambda(x)| \le Q(x)$, $\forall\, x$, λ. Or more generally, suppose that there is an integer $q \ge 0$ such that

$$\frac{f_\lambda(x)}{(x+i)^q} \to \frac{f(x)}{(x+i)^q} \quad \text{in } L^1(\mathbb{R}). \tag{10.4.2}$$

Then $f_\lambda \to f$ in $\mathcal{S}'$:

$$\left| \int_{\mathbb{R}} \{f(x) - f_\lambda(x)\}\phi(x)dx \right| \le \int_{\mathbb{R}} \left| \frac{f(x) - f_\lambda(x)}{(x+i)^q} \right| |(x+i)^q \phi(x)| dx.$$

[Under the first condition one may also use Lebesgue's Dominated Convergence Theorem to prove that $\int f_\lambda \phi \to \int f\phi$.]

(ii) In applications one encounters all sorts of *delta families* $\{f_\lambda\}$ on $\mathbb{R}$, that is, families of functions in $\mathcal{P}$ that converge to the delta distribution in $\mathcal{S}'$. Specific examples are

$$\frac{\sin Ax}{\pi x} \to \delta(x) \quad \text{and} \quad \frac{\sin^2 \frac{1}{2}Ax}{2\pi A(\frac{1}{2}x)^2} \to \delta(x) \quad \text{as } A \to \infty,$$

$$\frac{e^{-x^2/(4t)}}{2\sqrt{\pi t}} \to \delta(x) \;\text{ as }\; t \searrow 0, \qquad \frac{y}{\pi(x^2+y^2)} \to \delta(x) \;\text{ as }\; y \searrow 0.$$

For proofs that $\int f_\lambda \phi \to \phi(0)$ in these cases for all functions ϕ in $\mathcal{S}$, cf. the proof of Theorem 9.2.2, Remark 9.2.4, Example 9.8.1 and Exercise 9.8.3.

(iii) The operations of translation, reflection, multiplication by $e^{i\lambda x}$ or by $\omega(x)$ as in Definition 10.4.3 are continuous on $\mathcal{S}'$. For example, if $T_k \to T$ then $e^{i\lambda x}T_k \to e^{i\lambda x}T$.

(iv) The infinite series $\sum_{\infty}^{\infty} \delta(x - 2\pi n)$ converges in $\mathcal{S}'$ to $\delta_{2\pi}^{\text{per}}(x)$.

(v) If a Hermite series $\sum c_n h_n$ converges to T in $\mathcal{S}'$, then

$$< T, h_n > = \lim_{k\to\infty} \left\langle \sum_1^k c_j h_j, h_n \right\rangle = c_n, \quad \forall\, n.$$

Definition 10.4.7. For a tempered distribution T one defines the Hermite series as

$$T \sim \sum_{n=0}^{\infty} c_n[T] h_n, \quad \text{with } c_n[T] = < T, h_n > .$$

Proposition 10.4.8. *The Hermite series of a tempered distribution T converges to T in $\mathcal{S}'$.*

Proof. For ϕ in $\mathcal{S}$ one has $s_k[\phi] = \sum_0^k c_n[\phi] h_n \to \phi$ in $\mathcal{S}$ [see Corollary 10.3.4]. Hence for a continuous linear functional T on $\mathcal{S}$,

$$< T, \phi > = \lim < T, s_k[\phi] > = \lim \sum_0^k c_n[\phi] < T, h_n >$$

$$= \lim \sum_0^k c_n[T] c_n[\phi] = \sum_0^{\infty} c_n[T] c_n[\phi]. \tag{10.4.3}$$

Thus for every ϕ in $\mathcal{S}$,

$$< s_k[T], \phi > = \sum_0^k c_n[T] < h_n, \phi > = \sum_0^k c_n[T] c_n[\phi] \to < T, \phi > .$$

□

It will follow from Section 10.6 that the series in the final member of (10.4.3) is absolutely convergent.

Example 10.4.9. One has

$$1 \sim \sum_0^{\infty} a_n h_n, \quad \text{where } a_n = \int_{\mathbb{R}} h_n.$$

Integrating formula (10.3.2) with $n-1$ instead of n, one finds that $\sqrt{n/2} \int_{\mathbb{R}} h_n$ is equal to $\sqrt{(n-1)/2} \int_{\mathbb{R}} h_{n-2}$, so that

$$a_n = \sqrt{\frac{n-1}{n}}\, a_{n-2} = \cdots = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ \binom{2k}{k}^{\frac{1}{2}} 2^{\frac{1}{2}-k} \pi^{\frac{1}{4}} & \text{if } n = 2k. \end{cases}$$

Exercises. 10.4.1. Prove that for $\mathcal{C}^{\infty}$ functions ω as in Definition 10.4.3, one has $\omega(x)\delta_a(x) = \omega(a)\delta_a(x)$.

10.4.2. Prove that

(i) $\left\langle \text{pv}\,\frac{1}{x}, \phi(x) \right\rangle = \int_0^\infty \frac{1}{x}\,\{\phi(x) - \phi(-x)\}dx, \quad \forall\, \phi \in \mathcal{S};$

(ii) $\text{pv}\,\frac{1}{x} = \frac{1}{x}$ on $(0, \infty)$ and on $(-\infty, 0)$;

(iii) $x \cdot \text{pv}\,\frac{1}{x} = 1$ on $\mathbb{R}$.

10.4.3. Let $g_\varepsilon(x) = \frac{1}{\varepsilon}$ for $|x| < \frac{1}{2}\varepsilon$, $g_\varepsilon(x) = 0$ for $|x| > \frac{1}{2}\varepsilon$. Compute $\lim_{\varepsilon\searrow 0}\, g_\varepsilon$ in $\mathcal{S}'$.

10.4.4. Verify the convergence results in Examples 10.4.6.

10.4.5. Let g be an integrable function on $\mathbb{R}$ with $\int_{\mathbb{R}} g = 1$ and let f be a bounded uniformly continuous function on $\mathbb{R}$. For $0 < \varepsilon \le 1$ one sets $g_\varepsilon(x) = \frac{1}{\varepsilon}\, g\left(\frac{x}{\varepsilon}\right)$. Prove that for $\varepsilon \searrow 0$,

(i) $\quad g_\varepsilon \to \delta$ in $\mathcal{S}'$;

(ii) $$(g_\varepsilon * f)(x) = \int_{\mathbb{R}} g_\varepsilon(y) f(x-y)dy = \int_{\mathbb{R}} f(x - \varepsilon y) g(y) dy$$
$$\to \text{“}(\delta * f)(x)\text{”} = f(x), \quad \text{uniformly on } \mathbb{R}.$$

10.4.6. Prove that $\delta = \sum_0^\infty b_n h_n$ with

$$b_n = h_n(0) = -\sqrt{\frac{n-1}{n}}\, b_{n-2} = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ (-1)^k \binom{2k}{k}^{\frac{1}{2}} 2^{-k}\pi^{-\frac{1}{4}} & \text{if } n = 2k. \end{cases}$$

10.4.7. Solve the equation $xT(x) = 0$ in $\mathcal{S}'$.

10.4.8. Determine all solutions of the equation $xT(x) = 1$.

10.4.9. Show that the Hermite series $\sum_0^\infty a_n h_n$ for $f = 1$ converges to 1 at the origin.

10.4.10. Prove that $\mathcal{S}$ lies dense in $\mathcal{S}'$. That is, every tempered distribution T is $\mathcal{S}'$-limit of test functions ϕ_k.

10.5. Derivatives of tempered distributions

Suppose first that T is equal to a function f on $\mathbb{R}$ which is bounded by a polynomial and can be written as an indefinite integral: $f(t) = c + \int_a^t f'(v)dv$. Then integration by parts gives

$$< f', \phi > = \int_{\mathbb{R}} f'\phi = \big[f\phi\big]_{-\infty}^{\infty} - \int_{\mathbb{R}} f\phi' = - < f, \phi' > .$$

For an arbitrary tempered distribution T one defines the distributional derivative DT by a corresponding *formal* integration by parts, as in the case of periodic distributions; cf. Section 4.5.

DEFINITION 10.5.1. For a tempered distribution T, the *derivative* DT is the tempered distribution given by the formula

$$< DT, \phi > \stackrel{\text{def}}{=} - < T, \phi' >, \quad \forall\, \phi \in \mathcal{S}. \tag{10.5.1}$$

EXAMPLES 10.5.2. One may consider δ as the *derivative* of the *unit step function*

$$U(x) = 1_+(x) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x < 0. \end{cases} \tag{10.5.2}$$

Indeed, since $U \in \mathcal{P}$, one has

$$\begin{aligned} < DU, \phi > &\stackrel{\text{def}}{=} - < U, \phi' > = - \int_{\mathbb{R}} U\phi' \\ &= \; - \int_0^\infty \phi' = \phi(0) = < \delta, \phi >, \quad \forall\, \phi \in \mathcal{S}. \end{aligned}$$

Example 10.4.2 (ii) shows thatindexderivative!of $\log |x|$

$$D \log |x| = \mathrm{pv}\, \frac{1}{x}.$$

For *any function* f in $\mathcal{P}$, whether differentiable or not, the class $\mathcal{S}'$ contains distributional *derivatives* Df, $D^2 f$, $\cdots$ of every order.

For T in $\mathcal{S}'$ and $\mathcal{C}^\infty$ functions ω as in Definition 10.4.3, one has the *Product Rule*

$$D(T\omega) = DT \cdot \omega + T \cdot \omega'.$$

The distributional derivative DT determines T up to a constant:

PROPOSITION 10.5.3. *Let $DT = 0$ on $(a, b) \subset \mathbb{R}$. Then $T = C$ on (a, b), a constant function.*

PROOF. For the case $(a, b) = \mathbb{R}$ one can give a quick proof with the aid of Hermite series. Indeed, by Definition 10.5.1 and formula (10.3.2),

$$\begin{aligned} < DT, h_n > &= - < T, h_n' > \\ &= - < T, \sqrt{(n/2)}\, h_{n-1} - \sqrt{(n+1)/2}\, h_{n+1} > \\ &= \sqrt{(n+1)/2}\, c_{n+1}[T] - \sqrt{(n/2)}\, c_{n-1}[T]. \end{aligned}$$

This holds for all $n \geq 0$ if we set $c_{-1}[T] = 0$. Now suppose $DT = 0$ on $\mathbb{R}$. Then $< DT, h_n >= 0$, $\forall\, n$, hence

$$c_n[T] = \sqrt{\frac{n-1}{n}}\, c_{n-2}[T], \quad \forall\, n \geq 1.$$

This is the same recurrence relation as is satisfied by the Hermite coefficients a_n of the constant function 1; see Example 10.4.9. Thus

$$\begin{gathered} c_n[T]/c_n[1] = c_0[T]/c_0[1] \quad \text{for } n = 2, 4, \cdots, \\ c_n[T] = c_n[1] = 0 \quad \text{for } n = 1, 3, \cdots . \end{gathered}$$

It follows that

$$T = \sum_0^\infty c_n[T] h_n = \frac{c_0[T]}{c_0[1]} \sum_0^\infty c_n[1] h_n = \frac{c_0[T]}{c_0[1]} \cdot 1 = C.$$

For the case $(a, b) \neq \mathbb{R}$, cf. Exercises 4.5.6, 4.5.7. □

THEOREM 10.5.4. *Differentiation is a* continuous *linear operation on* $\mathcal{S}'$*: if* $T_\lambda \to T$ *in* $\mathcal{S}'$ *as* $\lambda \to \lambda_0$*, then* $DT_\lambda \to DT$ *in* $\mathcal{S}'$*. In particular convergent series in* $\mathcal{S}'$ *may be differentiated term by term.*

Indeed, if $< T_\lambda, \psi > \to < T, \psi >$ for all test functions ψ, then

$$< DT_\lambda, \phi > = \, - < T_\lambda, \phi' > \to \, - < T, \phi' > = < DT, \phi >$$

for all ϕ in $\mathcal{S}$.

COROLLARY 10.5.5. (*Test for convergence in* $\mathcal{S}'$) The following condition is sufficient for convergence $T_\lambda \to T$ in $\mathcal{S}'$ as $\lambda \to \lambda_0$: There exist functions f_λ and f in $\mathcal{P}$ and a pair of nonnegative integers s and q such that

$$T_\lambda = D^s f_\lambda, \quad T = D^s f, \quad \frac{f_\lambda(x)}{(x+i)^q} \to \frac{f(x)}{(x+i)^q} \tag{10.5.3}$$

uniformly on $\mathbb{R}$ or in $L^1(\mathbb{R})$ as $\lambda \to \lambda_0$; cf. Example 10.4.6 (i). [Uniform convergence in (10.5.3) implies L^1 convergence when q is replaced by $q+2$.]

It will follow from Theorem 10.6.3 below that the above it strong convergence $T_\lambda \to T$ is also necessary for (weak) convergence $T_\lambda \to T$ in $\mathcal{S}'$.

Exercises. 10.5.1. Verify that the functional DT on $\mathcal{S}$ defined by formula (10.5.1) is linear and continuous.

10.5.2. Let g be an indefinite integral on $\mathbb{R}$ with g' in $\mathcal{P}$. Prove that $|g|$ is bounded by a polynomial, and that $Dg = g'$ on $\mathbb{R}$.

10.5.3. Compute $D|x|$ and $D^2|x|$.

10.5.4. Prove that for $f \in \mathcal{P}$ and $\phi \in \mathcal{S}$,

$$< D^s f, \phi > = (-1)^s \int_{\mathbb{R}} f \phi^{(s)}.$$

10.5.5. Verify the Product Rule in Examples 10.5.2.

10.5.6. One defines the 'principal value functions'

$$\text{p.v. } \log(x \pm i0) \quad \text{as} \quad \lim_{\varepsilon \searrow 0} \text{p.v. } \log(x \pm i\varepsilon),$$

and the distributions

$$\frac{1}{x \pm i0} \quad \text{as} \quad \lim_{\varepsilon \searrow 0} \frac{1}{x \pm i\varepsilon} \quad \text{in} \quad \mathcal{S}'.$$

Prove that in distributional sense,

$$\frac{1}{x \pm i0} = D \text{ p.v. } \log(x \pm i0) = \text{pv}\,\frac{1}{x} \mp \pi i \delta(x).$$

10.5.7. Treat the eigenvalue problem $(x^2 - D^2)T = \lambda T$ in $\mathcal{S}'$.

10.6. Structure of tempered distributions

We begin with an auxiliary result for Hermite series.

PROPOSITION 10.6.1. *A Hermite series* $\sum_0^\infty d_n h_n$ *converges in* $\mathcal{S}'$ [hence, converges to a tempered distribution] *if and only if there are constants* B *and* β *such that*

$$|d_n| \le B n^\beta, \quad n = 1, 2, \cdots. \tag{10.6.1}$$

PROOF. The proof is similar to that of Proposition 4.6.1 for Fourier series of periodic distributions, but here we use the operator $\mathcal{H} = x^2 - D^2$ instead of D. Note that $\mathcal{H}$ is continuous on $\mathcal{S}'$.

(i) Supposing that (10.6.1) is satisfied, let q be a nonnegative integer greater than $\beta + 1/2$. Then

$$\left| \frac{d_n}{(2n+1)^q} \right|^2 \le \frac{B^2}{n^{2q-2\beta}} = \frac{B^2}{n^{1+\delta}}, \quad n = 1, 2, \cdots,$$

where $\delta = 2q - 1 - 2\beta > 0$. Thus the series $\sum_0^\infty |d_n/(2n+1)^q|^2$ is convergent, and hence the Hermite series

$$\sum_0^\infty \frac{d_n}{(2n+1)^q} h_n$$

will converge to a function g in $L^2(\mathbb{R})$. It follows that $\sum d_n h_n = \mathcal{H}^q g$ in $\mathcal{S}'$.

(ii) Suppose now that $\sum_0^\infty d_n h_n = T$ in $\mathcal{S}'$. Then the series $\sum_0^\infty d_n c_n[\phi]$ converges to $< T, \phi >$ for every function ϕ in $\mathcal{S}$; cf. (10.4.3). It now follows as in the proof of Proposition 4.6.1 that the coefficients d_n must satisfy a set of inequalities (10.6.1); see also Theorem 10.3.3 on Hermite series of test functions. □

THEOREM 10.6.2. (Structure theorem) *The following three assertions are equivalent:*

(i) T is in $\mathcal{S}'$;

(ii) There exist g in $L^2(\mathbb{R})$ and q in $\mathbb{N}_0$ such that $T = \mathcal{H}^q g$;

(iii) There exist f in $\mathcal{P}$ and s in $\mathbb{N}_0$ such that $T = D^s f$.

PROOF. It will be enough to prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii). Hence let T be in $\mathcal{S}'$. Then $T = \sum d_n h_n$ in $\mathcal{S}'$ where $d_n = c_n[T] =< T, h_n >$ [see Proposition 10.4.8]. Thus by Proposition 10.6.1 the coefficients d_n satisfy a set of inequalities (10.6.1), and by part (i) of the proof for that proposition, T can be represented in the form $\mathcal{H}^q g$ with $g \in L^2$.

One may next prove inductively that $\mathcal{H}^q g$ can be written as $D^{2q} f$, with $f \in \mathcal{P}$, for any $q \in \mathbb{N}_0$. Indeed, it is correct for $q = 0$ since $L^2 \subset \mathcal{P}$. Suppose now that the result has been proved for a certain $q \geq 0$. Then

$$\mathcal{H}^{q+1} g = (x^2 - D^2) D^{2q} f = x^2 D^{2q} f - D^{2q+2} f, \quad \text{with } f \in \mathcal{P}.$$

In order to write $x^2 D^{2q} f$ as a derivative of order $2q + 2$ of a function in $\mathcal{P}$ one may use two steps as follows. Observe that $x D^p f_1$ with $f_1 \in \mathcal{P}$ is equal to

$$D^p (x f_1) - p D^{p-1} f_1 = D^p f_2 = D^{p+1} f_3,$$

where

$$f_2 = x f_1 - p \int_0^x f_1 \quad \text{and} \quad f_3 = \int_0^x f_2 \quad \text{are in } \mathcal{P}.$$

□

Refinement of the above method as in Section 4.6 gives the following

THEOREM 10.6.3. (Characterization of convergence in $\mathcal{S}'$) *The following four statements about tempered distributions T_λ and T, where $\lambda \to \lambda_0$, are equivalent:*

(i) $T_\lambda \to T$ (weakly) in $\mathcal{S}'$;

(ii) $c_n[T_\lambda] \to c_n[T]$ as $\lambda \to \lambda_0$ for each $n \in \mathbb{N}_0$ and there are constants B and β such that

$$|c_n[T_\lambda]| \leq B n^\beta, \quad n = 1, 2, \cdots, \quad \forall \lambda \text{ close to } \lambda_0;$$

(iii) There are L^2 functions g_λ and g, and a nonnegative integer q, such that for all λ close to λ_0,

$$T_\lambda = \mathcal{H}^q g_\lambda, \quad T = \mathcal{H}^q g, \quad g_\lambda \to g \ \text{ in } L^2(\mathbb{R}) \text{ as } \lambda \to \lambda_0;$$

(iv) There are functions f_λ and f in $\mathcal{P}$, and nonnegative integers s and q, such that (10.5.3) *holds for all λ close to λ_0.*

One may say that (ii), (iii) and (iv) all describe "strong convergence" of T_λ to T.

Theorem 10.6.3 may be used to show that the space $\mathcal{S}'$ is complete; cf. Theorem 4.6.5.

Remark 10.6.4. The space of tempered distributions may, in fact, be obtained by *completion* of the space of integrable functions on $\mathbb{R}$ under the concept of convergence relative to test functions of class $\mathcal{S}$; cf. [**68**].

Exercises. 10.6.1. Let T be a distribution on the unit circle Γ. Prove that there is a 2π-periodic distribution T^{per} on $\mathbb{R}$ as follows. The restriction of T^{per} to any open interval (a,b) of length $\leq 2\pi$ is equal to the restriction of T to the subarc γ of Γ which corresponds to (a,b) modulo 2π. Here 'equality' is of course defined with the aid of test functions ϕ whose support belongs to (a,b) or γ.

Hint. Use the representation of Theorem 4.6.2.

10.6.2. Let T^{per} be a tempered distribution of period 2π. Prove that there exist a periodic integrable function f_0 of period 2π, a nonnegative integer s and a constant c such that $T^{\text{per}} = c + D^s f_0$. Deduce that to every distribution T^{per} there is a distribution T on the unit circle Γ such that T and T^{per} are related as in Exercise 10.6.1.

CHAPTER 11

Fourier transformation of tempered distributions

The class $\mathcal{P}$ of Section 10.1 is too small for a good theory of Fourier transformation. For example, the function $f(x) = 1$ cannot have a Fourier transform within $\mathcal{P}$. If it did, the prescription of Definition 10.1.3 would require that

$$\begin{aligned} < \mathcal{F}1, \phi > &= \int_{\mathbb{R}} \mathcal{F}1 \cdot \phi = \int_{\mathbb{R}} 1 \cdot \mathcal{F}\phi = \int_{\mathbb{R}} \widehat{\phi}(\xi) d\xi \\ &= 2\pi\phi(0) = < 2\pi\delta, \phi >, \quad \forall\, \phi \in \mathcal{S}. \end{aligned} \tag{11.0.2}$$

However, the distribution $2\pi\delta$ is not in $\mathcal{P}$! We have to extend $\mathcal{P}$ to the class $\mathcal{S}'$ of so-called tempered distributions in order to get a symmetric theory. Cf. books such as [**110**], [**111**], [**27**].

11.1. Fourier transformation in $\mathcal{S}'$

The following definition extends Definition 10.1.3 for $\mathcal{P}$ and says that always, "The effect of $\mathcal{F}T$ on ϕ must be the same as the effect of T on $\mathcal{F}\phi$".

Definition 11.1.1. Let T be in $\mathcal{S}'$. Then $\mathcal{F}T = \widehat{T}$ and $\mathcal{F}_R T = \check{T}$ are the tempered distributions given by

$$< \mathcal{F}T, \phi > = < T, \mathcal{F}\phi >, \quad < \mathcal{F}_R T, \phi > = < T, \mathcal{F}_R \phi >, \quad \forall\, \phi \in \mathcal{S}.$$

These formulas define $\mathcal{F}T$ and $\mathcal{F}_R T$ as continuous linear functionals on $\mathcal{S}$, because $\mathcal{F}$ and $\mathcal{F}_R$ are continuous linear operators $\mathcal{S} \mapsto \mathcal{S}$. Observe also that

$$\mathcal{F}_R T = (\mathcal{F}T)_R = \mathcal{F}T_R,$$

since

$$\begin{aligned} < \mathcal{F}_R T, \phi > &= < T, \mathcal{F}_R \phi > = < T, \mathcal{F}\phi_R > \\ &= < \mathcal{F}T, \phi_R > = < (\mathcal{F}T)_R, \phi >, \end{aligned}$$

and similarly for the second equality.

THEOREM 11.1.2. (Inversion and continuity) *Fourier transformation defines a one to one continuous linear map of $\mathcal{S}'$ onto itself. In $\mathcal{S}'$,*

$$\mathcal{F}^{-1} = \frac{1}{2\pi}\,\mathcal{F}_R, \quad \mathcal{F}_R^{-1} = \frac{1}{2\pi}\,\mathcal{F}.$$

PROOF. (i) Let T be in $\mathcal{S}'$. We will prove that

$$\frac{1}{2\pi}\,\mathcal{F}_R\mathcal{F}T = \frac{1}{2\pi}\,\mathcal{F}\mathcal{F}_R T = T.$$

Indeed, the inversion theorem for $\mathcal{S}$ [Proposition 9.6.2 shows that

$$< \mathcal{F}_R\mathcal{F}T, \phi > = < \mathcal{F}T, \mathcal{F}_R\phi > = < T, \mathcal{F}\mathcal{F}_R\phi > = 2\pi < T\phi >,$$

and similarly with $\mathcal{F}\mathcal{F}_R$ instead of $\mathcal{F}_R\mathcal{F}$. In particular $\mathcal{F}$ will be injective [$\mathcal{F}T = 0$ implies $T = 0$] and surjective [$T = \mathcal{F}(1/2\pi)\mathcal{F}_R T$].

(ii) Suppose $T_\lambda \to T$ in $\mathcal{S}'$ as $\lambda \to \lambda_0$. Then $\mathcal{F}T_\lambda \to \mathcal{F}T$:

$$< \mathcal{F}T_\lambda, \phi > = < T_\lambda, \mathcal{F}\phi > \to < T, \mathcal{F}\phi > = < \mathcal{F}T, \phi >, \quad \forall\, \phi \in \mathcal{S}.$$

□

EXAMPLES 11.1.3. The computation in (11.0.2) and inversion give

$$\mathcal{F}1 = \mathcal{F}_R 1 = 2\pi\delta, \quad \mathcal{F}_R\delta = \mathcal{F}\delta = 1.$$

The second formula is in agreement with formal calculation:

$$(\mathcal{F}\delta)(\xi) = \text{``}\int \delta(x)e^{-i\xi x}dx\text{''} = e^{-i\xi x}\Big|_{x=0} = 1,$$

as well as with Definition 11.1.1:

$$< \mathcal{F}\delta, \phi > = \left\langle \delta, \widehat{\phi} \right\rangle = \widehat{\phi}(0) = \int_{\mathbb{R}} \phi(x)dx = < 1, \phi > .$$

One could also use Hermite series [cf. Exercise 10.4.6 and Example 10.4.9]:

$$\mathcal{F}\sum c_n h_n = \sum c_n \mathcal{F}h_n = \sum \sqrt{2\pi}\,(-i)^n c_n h_n. \tag{11.1.1}$$

PROPOSITION 11.1.4. *In $\mathcal{S}'$, rules (i)–(vii) of the Table in Section 9.3 are valid without any restrictions.*

Proofs may be derived from the corresponding rules for $\mathcal{S}$ and Definition 11.1.1. However, it is more elegant to argue by continuity: the various operations, including $\mathcal{F}$, are continuous on $\mathcal{S}'$, and every tempered distribution is a limit of test functions. [Think of the Hermite series for T.]

EXAMPLES 11.1.5. (i) For $p \in \mathbb{N}_0$,

$$(\mathcal{F}x^p)(\xi) = (iD)^p(\mathcal{F}1)(\xi) = 2\pi i^p D^p \delta(\xi).$$

(ii) Let $T(x) = \mathrm{pv}\,(1/x)$, so that $xT(x) = 1$. Thus by Fourier transformation,

$$iD\widehat{T}(\xi) = 2\pi\delta(\xi), \quad \text{and hence} \quad \widehat{T}(\xi) = -2\pi i U(\xi) + C.$$

Here U is the unit step function [Examples 4.5.3]. Since T is odd, so is $\widehat{T}$, hence $C = \pi i$. Thus the signum function of Exercise 1.2.5 will show up:

$$\widehat{T}(\xi) = -\pi i \,\mathrm{sgn}\, \xi.$$

Exercises. 11.1.1. Derive the rule $\mathcal{F}DT = i\xi\mathcal{F}T$ for tempered distributions T by computation of $< \mathcal{F}DT, \phi >$.

11.1.2. Use Fourier transformation to determine all tempered solutions T of the equation $xT = 0$. Also discuss the equation $x^2T = 0$.

11.1.3. Successively compute the Fourier transforms of

$$e^{i\lambda x}, \quad \cos\lambda x, \quad \sin\lambda x, \quad \frac{\sin\lambda x}{x} \quad (\lambda \in \mathbb{R}).$$

11.1.4. Verify relation (11.1.1) on termwise Fourier transformation of Hermite series and use it to compute $\mathcal{F}\delta$.

11.1.5. Compute the Fourier transforms of

$$\frac{1}{x+i0}, \quad \frac{1}{x-i0}, \quad U = 1_+, \quad U_R = 1_-, \quad \mathrm{sgn}\, x.$$

11.1.6. Compute the Fourier transforms of pv $1/(x-a)$ and pv $1/(x^2-a^2)$ for $a \in \mathbb{R}$, $a \neq 0$.

11.1.7. Use Fourier transformation to determine all tempered solutions $I = I(t)$ of the differential equation $LDI + RI = \delta(t)$ on $\mathbb{R}$. [Cf. Exercise 4.5.10.]

11.2. Some applications

As a nice application of Fourier transformation we will obtain *Poisson's summation formula*. By way of preparation we prove

LEMMA 11.2.1. *In $\mathcal{S}'$ one has*

$$\sum_{n=-\infty}^{\infty} e^{inx} = 2\pi\delta_{2\pi}^{\mathrm{per}}.$$

PROOF. Leaving out the constant term and integrating twice, one obtains a uniformly convergent series, hence the given series converges to a distribution T in $\mathcal{S}'$. One integration of $T-1$ gives the familiar series

$$\sum_{n\neq 0} \frac{e^{inx}}{in} = \sum_{n=1}^{\infty} 2\,\frac{\sin nx}{n}.$$

The sum of this series is equal to $\pi - x$ on $(0, 2\pi)$; cf. formula (1.1.3). By periodicity, the sum function on $\mathbb{R}$ will have a jump 2π at each point $2\pi k$. Thus, by differentiation,

$$T - 1 = \sum_{k=-\infty}^{\infty} 2\pi\delta(x - 2\pi k) - 1 = 2\pi\delta_{2\pi}^{\text{per}} - 1.$$

The final term -1 is the classical derivative of $\pi - x$. □

THEOREM 11.2.2. (Poisson's sum formula) *For every test function ϕ in $\mathcal{S}$ and its Fourier transform $\widehat{\phi}$,*

$$\sum_{n=-\infty}^{\infty} \widehat{\phi}(2\pi n) = \sum_{n=-\infty}^{\infty} \phi(n). \tag{11.2.1}$$

PROOF. The left-hand side of (11.2.1) is equal to

$$\begin{aligned} \langle \delta_{2\pi}^{\text{per}}, \mathcal{F}\phi\rangle \;&=\; \langle \mathcal{F}\delta_{2\pi}^{\text{per}}, \phi\rangle \;=\; \left\langle \mathcal{F}\left[\frac{1}{2\pi}\sum_{-\infty}^{\infty} e^{inx}\right], \phi \right\rangle \\ &= \sum_{-\infty}^{\infty} \left\langle \frac{1}{2\pi}\mathcal{F}[e^{inx}\cdot 1], \phi \right\rangle = \sum_{-\infty}^{\infty} < \delta(\xi - n), \phi(\xi) > = \sum_{-\infty}^{\infty} \phi(n). \end{aligned}$$

□

EXAMPLE 11.2.3. We apply (11.2.1) to $\phi(x) = e^{-ax^2}$ with $a > 0$. In this case $\widehat{\phi}(\xi) = \sqrt{(\pi/a)}\, e^{-\xi^2/(4a)}$, so that one obtains the identity

$$\sum_{-\infty}^{\infty} e^{-an^2} = \sqrt{\frac{\pi}{a}} \sum_{-\infty}^{\infty} e^{-\pi^2 n^2/a}.$$

Poisson's formula actually holds for a class of well-behaved functions considerably larger than $\mathcal{S}$; cf. [**95**].

Fourier transforms can often be computed by using continuity.

THEOREM 11.2.4. (Evaluation theorem) *Let f be in $\mathcal{P}$ and let f_A be the truncated function equal to f for $|x| < A$ and equal to 0 for $|x| > A$. Also, let $\varepsilon > 0$. Then*

$$\begin{aligned}\widehat{f}(\xi) &= \mathcal{S}' \lim_{A\to\infty} \widehat{f}_A(\xi) = \mathcal{S}' \lim_{A\to\infty} \int_{-A}^{A} f(x)e^{-i\xi x}dx \\ &= \mathcal{S}' \lim_{\varepsilon\searrow 0} \int_{\mathbb{R}} f(x)e^{-\varepsilon|x|}e^{-i\xi x}dx.\end{aligned}$$

Indeed, by Definition 10.4.5,

$$f(x) = \mathcal{S}' \lim f_A(x) = \mathcal{S}' \lim f(x)e^{-\varepsilon|x|},$$

and Fourier transformation is continuous on $\mathcal{S}'$.

EXAMPLE 11.2.5. For $\alpha > -1$, the function $f(x) = |x|^\alpha$ is in $\mathcal{P}$ and

$$\begin{aligned}(\mathcal{F}|x|^\alpha)(\xi) &= \mathcal{S}' \lim_{\varepsilon\searrow 0} \int_{\mathbb{R}} |x|^\alpha e^{-\varepsilon|x|}e^{-i\xi x}dx \\ &= 2\cdot\mathcal{S}' \lim_{\varepsilon\searrow 0} \int_0^\infty x^\alpha e^{-\varepsilon x}\cos\xi x\,dx.\end{aligned}$$

For the evaluation of the limit we will use the *Laplace transform*

$$\int_0^\infty x^\alpha e^{-sx}dx = \Gamma(\alpha+1)\,\text{p.v.}\,s^{-\alpha-1}, \quad \text{Re}\,s > 0. \tag{11.2.2}$$

A proof of (11.2.2) may be obtained from Cauchy's theorem and the integral for the Gamma function. Set $s = \rho e^{i\theta}$ with $\rho > 0$ and $|\theta| < \pi/2$, so that $\text{p.v.}\,s = \log\rho + i\theta$. Writing $\rho e^{i\theta}x = z$ one then has

$$\begin{aligned}\int_0^\infty x^\alpha e^{-sx}dx &= \int_0^\infty x^\alpha e^{-\rho e^{i\theta}x}dx = (\rho e^{i\theta})^{-\alpha-1}\int_0^{\infty e^{i\theta}} z^\alpha e^{-z}dz \\ &= (\rho e^{i\theta})^{-\alpha-1}\int_0^\infty t^\alpha e^{-t}dt = \Gamma(\alpha+1)\text{p.v.}\,s^{-\alpha-1}.\end{aligned}$$

Returning to our Fourier transform, we may take $\xi > 0$ to obtain

$$\begin{aligned}(\mathcal{F}|x|^\alpha)(\xi) &= 2\Gamma(\alpha+1)\,\mathcal{S}' \lim_{\varepsilon\searrow 0} \text{Re}\,\text{p.v.}\,(\varepsilon + i\xi)^{-\alpha-1} \\ &= 2\Gamma(\alpha+1)\xi^{-\alpha-1}\text{Re}\,e^{-(\alpha+1)(\pi/2)i}.\end{aligned}$$

For the final step we impose the condition $\alpha < 0$, so that $\xi^{-\alpha-1}$ is integrable from 0 on. Since the complete answer for $\xi \in \mathbb{R}$ must be even, we conclude that

$$(\mathcal{F}|x|^\alpha)(\xi) = -2\Gamma(\alpha+1)\sin(\alpha\pi/2)\cdot|\xi|^{-\alpha-1}, \quad -1 < \alpha < 0. \tag{11.2.3}$$

The special case $\alpha = -1/2$ reveals another *eigenfunction* of Fourier transformation in $\mathcal{S}'$.

Exercises. 11.2.1. Use the relations

$$1 = \lim 1_A = \lim e^{-\varepsilon|x|} = \lim e^{-\varepsilon x^2}$$

in $\mathcal{S}'$ for the computation of $\mathcal{F}1$.

11.2.2. Use the relation $\mathrm{pv}\, 1/x = \lim\, x/(x^2+\varepsilon^2)$ for the computation of $\mathcal{F}[\mathrm{pv}\, 1/x]$.

11.2.3. For small $\delta > 0$ one will expect that

$$\sum_{n=-\infty}^{\infty} \frac{\sin^2 n\delta}{n^2\delta^2}\,\delta \approx \int_{\mathbb{R}} \frac{\sin^2 \xi}{\xi^2}\, d\xi = \pi.$$

(i) Prove that the approximation is exact for $0 < \delta < \pi$.

(ii) Determine the sum of the series for other values of δ.

11.2.4. Does one get the right answer if one applies Poisson's sum formula to $\phi(x) = e^{-a|x|}$?

11.2.5. Compute $\mathcal{F}\delta_a^{\mathrm{per}} = \mathcal{F}\left[\sum_{n=-\infty}^{\infty} \delta(x-an)\right]$ for $a > 0$. What happens if $a = \sqrt{2\pi}$?

11.2.6. Compute $\mathcal{F}\left[|x|^{-\frac{1}{2}}\right]$ and $\mathcal{F}\left[|x|^{-\frac{1}{2}}\mathrm{sgn}\, x\right]$.

11.2.7. Determine all eigenvalues of $\mathcal{F}$ as an operator on $\mathcal{S}'$. Characterize the 'eigendistributions' by their Hermite series.

*11.2.8. The Bessel function $J_0(x)$ is even and tends to zero as $x \to \infty$; cf. Examples 8.1.3, 8.1.6 and Exercise 8.1.10. Use the power series for $J_0(x)$ to derive that its Fourier transform is given by

$$\begin{aligned} g(\xi) &= \mathcal{S}' \lim_{\varepsilon \searrow 0} \int_0^\infty J_0(x) e^{-\varepsilon x}\left(e^{-i\xi x} + e^{i\xi x}\right) dx \\ &= \mathcal{S}' \lim_{\varepsilon \searrow 0} \left[\mathrm{p.v.}\left\{1 + (\varepsilon + i\xi)^2\right\}^{-\frac{1}{2}} + \mathrm{p.v.}\left\{1 + (\varepsilon - i\xi)^2\right\}^{-\frac{1}{2}}\right] \\ &= \begin{cases} 2(1-\xi^2)^{-\frac{1}{2}} & \text{for } |\xi| < 1, \\ 0 & \text{for } |\xi| > 1. \end{cases} \end{aligned}$$

Hint. One may start with $\varepsilon > 1$ for termwise integration; cf. Exercise 12.1.4.

*11.2.9. Show that the Fourier transform $g(\xi)$ of $f(x) = J_0(x)\,\mathrm{sgn}\, x$ is (for $\xi > 0$) given by

$$\begin{aligned} g(\xi) &= \mathcal{S}' \lim_{\varepsilon\searrow 0} \int_0^\infty J_0(x) e^{-\varepsilon x}\big(e^{-i\xi x} - e^{i\xi x}\big)dx \\ &= \begin{cases} 0 & \text{for } 0 < \xi < 1, \\ -2i(\xi^2-1)^{-\frac{1}{2}} & \text{for } \xi > 1. \end{cases} \end{aligned}$$

11.2.10. For $\mathrm{Re}\, a > -1$ one has the following formula in $\mathcal{P}$:

$$|x|^a = \frac{1}{(a+1)(a+2)}\, D^2 |x|^{a+2}. \tag{11.2.4}$$

For $\mathrm{Re}\, a \leq -1$ (but $a \neq -1, -2, \cdots$), tempered distributions $|x|^a$ may be defined recursively by formula (11.2.4). Prove that the resulting family of distributions $\{|x|^a\}$ depends analytically on a for $a \neq -1, -2, \cdots$. More precisely, the function $f_\phi(a) = \langle |x|^a, \phi(x)\rangle$ is analytic for every test function ϕ. Deduce that the family of Fourier transforms $\{\mathcal{F}|x|^a\}$ also depends analytically on a. Can one conclude that the formula for $\mathcal{F}|x|^a$, obtained in Example 11.2.5 for $-1 < a < 0$, is valid for all complex a with $a \neq 0, \pm 1, \pm 2, \cdots$?

11.2.11. A similar story applies to the functions or distributions $|x|^a \mathrm{sgn}\, x$. Show that

$$|x|^a \mathrm{sgn}\, x = \frac{1}{a+1}\, D|x|^{a+1} \;\text{ and }\; x|x|^a = |x|^{a+1}\mathrm{sgn}\, x.$$

Deduce that for $a \notin \mathbb{Z}$,

$$\mathcal{F}\big[|x|^a \mathrm{sgn}\, x\big](\xi) = -2i\Gamma(a+1)\cos(a\pi/2)\cdot|\xi|^{-a-1}\mathrm{sgn}\,\xi.$$

11.3. Convolution

Convolution is important for many applications, but it can be defined only under certain restrictions. For tempered distributions T and test functions ϕ one sets

$$(T*\phi)(x) \overset{\text{def}}{=} \text{“}\int_{\mathbb{R}} T(y)\phi(x-y)dy\text{”} = < T(y), \phi(x-y) > . \tag{11.3.1}$$

The result will be a $\mathcal{C}^\infty$ function ω of the type described in Definition 10.4.3. [Write $T = D^s f$ to verify this.] If T has compact support, $T*\phi$ is again a test function. Example:

$$\delta * \phi = \phi. \tag{11.3.2}$$

In the case of L^1 functions f and g, with g of compact support, Fubini's theorem gives

$$\begin{aligned}< f * g, \phi > &= \int \left\{ \int f(y) g(x-y) dy \right\} \phi(x) dx \\ &= \int f(y) dy \int g_R(y-x)\phi(x) dx = < f, g_R * \phi > .\end{aligned}$$

One may use an analog of this formula to define a convolution for special S and T:

Definition 11.3.1. For S and T in $\mathcal{S}'$ with S of compact support, one sets

$$< S * T, \phi > = < T * S, \phi > = < T, S_R * \phi > . \tag{11.3.3}$$

Using (11.3.3) one finds in particular that $< \delta * T, \phi > = < T, \delta * \phi >$ is equal to $< T, \phi >$ for all ϕ, hence

$$\delta * T = T * \delta = T, \quad \forall T \in \mathcal{S}'. \tag{11.3.4}$$

In words, δ plays the role of a *unit* under convolution in $\mathcal{S}'$.

Proposition 11.3.2. *In the cases described by* (11.3.1) *and* (11.3.3),

$$\mathcal{F}(T * \phi) = \widehat{T}\widehat{\phi}, \quad \mathcal{F}(S * T) = \widehat{S}\widehat{T}. \tag{11.3.5}$$

For T of compact support, one can show that the Fourier transform $\widehat{T}$ is a polynomially bounded $\mathcal{C}^\infty$ function, as are the derivatives of $\widehat{T}$.

Exercises. 11.3.1. Verify that δ acts as unit element relative to convolution in $\mathcal{S}'$. More precisely, prove formulas (11.3.2) and (11.3.4).

11.3.2. Let f be an integrable function on $\mathbb{R}$ with compact support. Prove that

(i) $\widehat{f}(\xi)$ can be extended to an entire function (a function that is analytic everywhere) $\widehat{f}(\zeta) = \widehat{f}(\xi + i\eta)$;

(ii) $\widehat{f}(\xi)$ and its derivatives are polynomially bounded on $\mathbb{R}$.

11.3.3. Let f and g be integrable functions on $\mathbb{R}$ such that $f * g = 0$. Given that f has compact support and that g is not the zero function, what can you conclude about f?

11.3.4. Let f be any integrable function on $\mathbb{R}$ – compactly supported or not – such that $\int_{\mathbb{R}} f = 0$. Prove that $f * 1 = 0$.

11.3.5. Prove the formulas (11.3.5):

(i) for the case $T * \phi$ with $T = f \in \mathcal{P}$, so that $T * \phi = \mathcal{S}' \lim (f_A * \phi)$;

(ii) for the case $T * \phi$ with $T = D^s f$, $f \in \mathcal{P}$, so that $T * \phi = D^s(f * \phi)$;
(iii) for the case $S * T$ with S of compact support.

In the following exercises $\delta(x)$ appears as a convenient idealization of either a large displacement around the point $x = 0$ of a system at time $t = 0$, or of a unit impulse transmitted to the system at the origin at time $t = 0$, or of a high temperature peak in the immediate vicinity of the origin at time $t = 0$.

If one has a solution corresponding to boundary 'function' $\delta(x)$, one can obtain a solution with boundary function $f(x)$ with the aid of convolution.

11.3.6. Solve the boundary value problem

$$\begin{aligned} &u_{xx} = \frac{1}{c^2}\, u_{tt}, \quad -\infty < x < \infty, \ \ t > 0; \\ &u(x,0) = \delta(x), \ \ u_t(x,0) = 0, \quad -\infty < x < \infty. \end{aligned}$$

At which points x will one observe a displacement at time t? What can you conclude about the speed of propagation?

11.3.7. Same questions for the problem

$$\begin{aligned} &u_{xx} = \frac{1}{c^2}\, u_{tt}, \quad -\infty < x < \infty, \ \ t > 0; \\ &u(x,0) = 0, \ \ u_t(x,0) = \delta(x), \quad -\infty < x < \infty. \end{aligned}$$

11.3.8. Same questions for the heat flow (or diffusion) problem

$$\begin{aligned} &u_{xx} = u_t, \quad -\infty < x < \infty, \ \ t > 0; \\ &u(x,0) = \delta(x), \quad -\infty < x < \infty. \end{aligned}$$

Here 'displacement' should be understood as change in temperature or concentration.

11.4. Multiple Fourier integrals

Readers who have to get used to notations with many indices may wish to start with concrete Example 11.5.2 below.

In the following x denotes a point or vector in $\mathbb{R}^n$: $x = (x_1, x_2, \cdots, x_n)$, and similarly $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$, with standard inner product

$$\xi \cdot x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n.$$

For functions f in $L^1(\mathbb{R}^n)$ one naturally defines the *Fourier transform* and the reflected Fourier transform by the formulas

$$g(\xi) = \widehat{f}(\xi) = (\mathcal{F}f)(\xi) = \int_{\mathbb{R}^n} f(x)e^{-i\xi\cdot x}dx, \tag{11.4.1}$$

$$h(\xi) = (\mathcal{F}_R f)(\xi) = (\mathcal{F}f)_R(\xi) = \int_{\mathbb{R}^n} f(x)e^{i\xi\cdot x}dx \tag{11.4.2}$$

for all ξ in $\mathbb{R}^n$. By Fubini's theorem, the multiple integral for $g(\xi)$ may be written as a repeated integral:

$$\begin{aligned} g(\xi) &= \int_{\mathbb{R}^n} f(x_1, x_2, \cdots, x_n)e^{-i\xi_1 x_1}e^{-i\xi_2 x_2}\cdots e^{-i\xi_n x_n}dx_1 dx_2\cdots dx_n \\ &= \int_{\mathbb{R}} e^{-i\xi_1 x_1}dx_1 \int_{\mathbb{R}} e^{-i\xi_2 x_2}dx_2 \cdots \int_{\mathbb{R}} f(x_1, x_2, \cdots, x_n)e^{-i\xi_n x_n}dx_n. \end{aligned} \tag{11.4.3}$$

Symbolically,

$$g = \mathcal{F}^x f = \mathcal{F}^{x_1}\mathcal{F}^{x_2}\cdots\mathcal{F}^{x_n} f, \tag{11.4.4}$$

where $\mathcal{F}^{x_j}$ represents 1-dimensional Fourier transformation relative to x_j.

In the special case $f(x) = f_1(x_1)f_2(x_2)\cdots f_n(x_n)$ with f_j in $L^1(\mathbb{R})$, it immediately follows that $\widehat{f}(\xi) = \widehat{f_1}(\xi_1)\widehat{f_2}(\xi_2)\cdots\widehat{f_n}(\xi_n)$, where $\widehat{f_j}$ is the 1-dimensional Fourier transform of f_1. Thus by Example 9.1.4, taking $a > 0$,

$$\begin{aligned} \mathcal{F}\left[e^{-a|x|^2}\right] &= \mathcal{F}\left[e^{-ax_1^2}\cdots e^{-ax_n^2}\right] \\ &= \left(\frac{\pi}{a}\right)^{n/2} e^{-\xi_1^2/(4a)}\cdots e^{-\xi_n^2/(4a)} = \left(\frac{\pi}{a}\right)^{n/2} e^{-|\xi|^2/(4a)}. \end{aligned} \tag{11.4.5}$$

Fourier inversion on $\mathbb{R}^n$. For well-behaved functions f one may invert formula (11.4.4) step by step, that is, relative to one variable at a time:

$$\begin{gathered} \mathcal{F}^{x_2}\cdots\mathcal{F}^{x_n} f = \frac{1}{2\pi}\,\mathcal{F}_R^{\xi_1} g, \quad \mathcal{F}^{x_3}\cdots\mathcal{F}^{x_n} f = \frac{1}{(2\pi)^2}\,\mathcal{F}_R^{\xi_2}\mathcal{F}_R^{\xi_1} g, \ \cdots, \\ f = \frac{1}{(2\pi)^n}\,\mathcal{F}_R^{\xi_n}\cdots\mathcal{F}_R^{\xi_2}\mathcal{F}_R^{\xi_1} g = \frac{1}{(2\pi)^n}\,\mathcal{F}_R^{\xi} g. \end{gathered} \tag{11.4.6}$$

This procedure works in particular for functions of class $\mathcal{S}$ in $\mathbb{R}^n$, that is, the $\mathcal{C}^\infty$ functions ϕ on $\mathbb{R}^n$ for which all seminorms

$$M_{\alpha\beta}(\phi) = \sup_{x\in\mathbb{R}^n} \left|x^\alpha D^\beta \phi\right| = \sup_x \left|x_1^{\alpha_1}\cdots x_n^{\alpha_n} D_1^{\beta_1}\cdots D_n^{\beta_n}\phi\right|$$

are finite. Here α and β are multi-indices ≥ 0: $\alpha = (\alpha_1, \cdots, \alpha_n)$ with nonnegative integers α_j, and similarly for β. The expression x^α is the standard abbreviation for the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, while

$$D^\beta \phi = D_1^{\beta_1} \cdots D_n^{\beta_n} \phi = \frac{\partial^{\beta_1 + \cdots + \beta_n}}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}} \phi.$$

Convergence $\phi_\lambda \to \phi$ in $\mathcal{S}$ as $\lambda \to \lambda_0$ shall mean that $M_{\alpha\beta}(\phi - \phi_\lambda) \to 0$ for all multi-indices α and $\beta \geq 0$.

For the space $\mathcal{S}$ one readily verifies the assertions

$$\begin{aligned}
&\mathcal{F} D_j \phi = i\xi_j \widehat{\phi}, \quad \mathcal{F} x_j \phi = i D_j \widehat{\phi}, \\
&\text{if } \ \psi = \mathcal{F}\phi \ \text{ then } \ \phi = \frac{1}{(2\pi)^n} \mathcal{F}_R \psi, \qquad (11.4.7)\\
&\mathcal{F} \ \text{ defines a } 1-1 \text{ continuous linear map of } S \text{ onto itself.}
\end{aligned}$$

The space $\mathcal{S}'$ of the *tempered distributions* on $\mathbb{R}^n$ consists of the continuous linear functionals T on $\mathcal{S}$, with the associated (weak) convergence:

$$T_\lambda \to T \ \text{ in } \ \mathcal{S}' \ \text{ if } \ < T_\lambda, \phi > \to < T, \phi >, \ \forall \phi \in \mathcal{S}.$$

An important subclass $\mathcal{P}$ of $\mathcal{S}'$ is given by the locally integrable functions f on $\mathbb{R}^n$ of at most polynomial growth. More precisely, f is in $\mathcal{P}$ if

$$\text{the quotient } \ \frac{f(x)}{(x_1 + i)^{q_1} \cdots (x_n + i)^{q_n}} \ \text{ is in } \ L^1(\mathbb{R}^n)$$

for some $q = (q_1, \cdots, q_n) \in \mathbb{N}_0^n$.

For tempered distributions T, the product $x^\alpha T$ by a monomial x^α, the derivative $D^\beta T$ of order $\beta = (\beta_1, \cdots, \beta_n)$, the Fourier transform $\widehat{T} = \mathcal{F}T$ and the reflected transform $\mathcal{F}_R T$ are defined as continuous linear functionals on $\mathcal{S}$ by the formulas

$$\begin{aligned}
&< x^\alpha T, \phi > \ = < T, x^\alpha \phi >, \\
&< D^\beta T, \phi > \ = (-1)^{\beta_1 + \cdots + \beta_n} < T, D^\beta \phi >, \qquad (11.4.8)\\
&< \mathcal{F}T, \phi > \ = < T, \mathcal{F}\phi >, \quad < \mathcal{F}_R T, \phi > = < T, \mathcal{F}_R \phi > .
\end{aligned}$$

In this way the operators $x^\alpha \cdot$, D^β, $\mathcal{F}$ and $\mathcal{F}_R$ inherit the nice properties which they have on $\mathcal{S}$; cf. (11.4.7).

For suitably matched functions or distributions S and T there is a convolution $S * T$, and $\mathcal{F}(S * T) = \mathcal{F}S \cdot \mathcal{F}T$. This holds in particular if S and T are L^1 functions on $\mathbb{R}^n$. It also holds for δ and arbitrary T in $\mathcal{S}'$, where the *delta distribution* δ is given by the usual formula, $< \delta, \phi > = \phi(0)$, $\forall \phi$.

We finally observe that for f in $\mathcal{P}$,

$$\widehat{f}(\xi) = (\mathcal{F}f)(\xi) = \mathcal{S}' \lim_{A\to\infty} \int_{|x|<A} f(x)e^{-i\xi\cdot x}dx$$

$$= \mathcal{S}' \lim_{\varepsilon\searrow 0} \int_{\mathbb{R}^n} f(x)e^{-\varepsilon|x|}e^{-i\xi\cdot x}dx. \tag{11.4.9}$$

Exercises. 11.4.1. Prove that for functions ϕ in $\mathcal{S}(\mathbb{R}^n)$, $\mathcal{F}(\partial\phi/\partial x_j) = i\xi_j\mathcal{F}T$. Deduce that for tempered distributions T on $\mathbb{R}^n$, $\mathcal{F}D_jT = i\xi_j\mathcal{F}T$.

11.4.2. Show that the "n-variable Hermite functions"

$$h_\alpha(x) \stackrel{\text{def}}{=} h_{\alpha_1}(x_1)h_{\alpha_2}(x_2)\cdots h_{\alpha_n}(x_n), \quad \alpha \in \mathbb{N}_0$$

form an orthonormal basis for $L^2(\mathbb{R}^n)$. How do these functions behave under n-dimensional Fourier transformation?

11.4.3. Sketch Fourier theory for $L^2(\mathbb{R}^n)$.

11.4.4. Prove that for the delta distribution on $\mathbb{R}^n$, one has $\mathcal{F}\delta = 1$ independently of n.

11.4.5. Show that for the n-dimensional Laplacian $\Delta = \Delta_n = D_1^2+\cdots+D_n^2$ and T in $\mathcal{S}'$,

$$\mathcal{F}[\Delta T](\xi) = -\rho^2\mathcal{F}T, \quad \text{where } \rho^2 = \xi_1^2+\cdots+\xi_n^2.$$

11.4.6. Let $f(x) = F(r)$ be a function in $\mathcal{P}$ on $\mathbb{R}^3$ which depends only on $|x| = r$. Show that

$$\int_{|x|<A} F(r)e^{-i\xi\cdot x}dx = 4\pi\int_0^A F(r)\,r\,\frac{\sin\rho r}{\rho}\,dr, \quad \text{where } \rho = |\xi|.$$

Deduce that the Fourier transform $\widehat{f}(\xi)$ depends only on ρ.

Hint. Fixing $\xi \neq 0$, one may introduce a system of 3-dimensional polar coordinates with polar axis along the vector ξ. Setting $x_1 = r\sin\theta\cos\phi$, $x_2 = r\sin\theta\sin\phi$, $x_3 = r\cos\theta$, the vector ξ becomes $(0,0,\rho)$ and dx becomes $r^2\sin\theta\,drd\theta d\phi$.

11.5. Fundamental solutions of partial differential equations

Fourier transformation is especially useful for the determination of so-called *fundamental solutions* for partial differential operators $p(D)$, where $p(x)$ is a polynomial in $x_1,\cdots,x_n$.

DEFINITION 11.5.1. A fundamental or elementary solution for the differential operator $p(D)$ is a function or distribution E such that $p(D)E = \delta$.

If there is a tempered fundamental solution E for $p(D)$, then by Fourier transformation, $p(i\xi)\widehat{E} = \widehat{\delta} = 1$, so that

$$\widehat{E}(\xi) = \frac{1}{p(i\xi)} = \frac{1}{p(i\xi_1, \cdots, i\xi_n)}. \tag{11.5.1}$$

Thus one will try to determine $E(x)$ from (11.5.1). In terms of a fundamental solution, a solution of the non-homogeneous equation $p(D)u = f$ will (for suitable f) be given by

$$u(x) = (E * f)(x) = \int_{\mathbb{R}^n} E(x-y)f(y)dy. \tag{11.5.2}$$

Indeed, if u is a solution, then by (formal) Fourier transformation, $p(i\xi)\widehat{u} = \widehat{f}$, hence $\widehat{u} = \widehat{E}\widehat{f}$. Finally apply Fourier inversion.

Example 11.5.2. We will discuss the Dirichlet problem for the upper half-space H in $\mathbb{R}^3$. Using the notation (x, y, z) for points in $\mathbb{R}^3$ instead of $x = (x_1, x_2, x_3)$, the boundary value problem takes the form

$$\begin{aligned} &\Delta u = u_{xx} + u_{yy} + u_{zz} = 0 \ \text{ in } \ H = \{(x,y,z) \in \mathbb{R}^3 : \ z > 0\}, \\ &u(x,y,0) = f(x,y), \ \ (x,y) \in \mathbb{R}^2; \ \ u(x,y,z) \ \text{ bounded on } \ H. \end{aligned}$$

The boundedness condition on u is a condition as $z \to \infty$; it would be more or less implied by a condition of finite energy, $\int_{\mathbb{R}^3}(u_x^2 + u_y^2 + u_z^2) < \infty$.

It makes sense to apply 2-dimensional Fourier transformation relative to (x, y), or repeated 1-dimensional Fourier transformation, first with respect to x and then with respect to y. Accordingly we set

$$\begin{aligned} v(\xi, \eta, z) &= \mathcal{F}^{x,y}u = \int_{\mathbb{R}^2} u(x,y,z)e^{-i(\xi x + \eta y)}dxdy \\ &= \int_{\mathbb{R}} e^{-i\eta y}dy \int_{\mathbb{R}} u(x,y,z)e^{-i\xi x}dx = \mathcal{F}^y\mathcal{F}^x u \ \big[= \mathcal{F}^x\mathcal{F}^y u \big]. \end{aligned} \tag{11.5.3}$$

Integration by parts with respect to x shows that u_x is transformed into $i\xi v$, and similarly u_y goes over into $i\eta v$, hence $u_{xx} + u_{yy}$ will become $-(\xi^2 + \eta^2)v$; it is convenient to introduce the notation $(\xi^2 + \eta^2)^{\frac{1}{2}} = \rho \geq 0$. Assuming that v_{zz} is [also] obtained by differentiation under the integral sign, so that u_{zz} goes over into v_{zz}, we obtain the new boundary value problem

$$\begin{aligned} &-\rho^2 v + v_{zz} = 0, \quad (\xi, \eta, z) \in H, \\ &v(\xi, \eta, 0) = \mathcal{F}^{x,y}f = g(\xi, \eta), \ \text{ say.} \end{aligned}$$

In the new differential equation, ξ and η occur only as parameters. The general solution will be

$$v(\xi,\eta,z) = a(\xi,\eta)e^{\rho z} + b(\xi,\eta)e^{-\rho z}.$$

Here $a(\xi,\eta)$ and $b(\xi,\eta)$ are not uniquely determined by the boundary condition $a+b=g$, but since we look for a bounded solution u on H, it is reasonable to demand that v not become exponentially large as $z\to\infty$. Thus we have to take $a=0$, so that

$$v(\xi,\eta,z) = g(\xi,\eta)e^{-\rho z}, \quad \rho = (\xi^2+\eta^2)^{\frac{1}{2}}. \tag{11.5.4}$$

We finally have to invert our 2-dimensional Fourier transform (11.5.3). Doing this in two steps, one finds

$$\begin{aligned}\mathcal{F}^x u = \frac{1}{2\pi}\,\mathcal{F}_R^\eta v, \quad u &= \frac{1}{2\pi}\,\mathcal{F}_R^\xi\,\frac{1}{2\pi}\,\mathcal{F}_R^\eta v\\ &= \frac{1}{4\pi^2}\,\mathcal{F}_R^{\xi,\eta} v = \frac{1}{4\pi^2}\int_{\mathbb{R}^2} v(\xi,\eta,z)e^{i(x\xi+y\eta)}d\xi d\eta.\end{aligned}$$

Now our v in (11.5.4) is a product. Successively applying $\mathcal{F}_R^\eta$ and $\mathcal{F}_R^\xi$, the function u will become a two-fold convolution, a convolution relative to (x,y). Defining

$$\frac{1}{4\pi^2}\,\mathcal{F}_R^{\xi,\eta}\big[e^{-\rho z}\big] = \frac{1}{4\pi^2}\,\mathcal{F}^{\xi,\eta}\big[e^{-\rho z}\big] = P(x,y,z), \tag{11.5.5}$$

one expects the final answer

$$u(x,y,z) = f * P = \int_{\mathbb{R}^2} f(s,t)P(x-s,y-t,z)dsdt, \quad z>0. \tag{11.5.6}$$

Exercises. 11.5.1. Show that a tempered fundamental solution E for the operator $-\Delta$ must have Fourier transform

$$\widehat{E}(\xi) = \frac{1}{\rho^2}, \quad \text{where } \rho = |\xi|.$$

11.5.2. Show that the operator $-\Delta = -\Delta_3$ in $\mathbb{R}^3$ has fundamental solution

$$E(x) = \frac{1}{4\pi r} = \frac{1}{4\pi|x|}.$$

Hint. One may use Exercise 11.4.6 with r and ρ interchanged.

11.5.3. Obtain a solution of Poisson's equation $-\Delta u = f$ in $\mathbb{R}^3$ if f is an integrable function with compact support.

11.5.4. Given a pair of tempered distributions $S = D^p f$, $T = D^q g$ on $\mathbb{R}$, with f and g in $\mathcal{P}$, one may define a tempered distribution $S(x)T(y)$ on $\mathbb{R}^2$ as $D_1^p D_2^q f(x)g(y)$. Show that the delta distribution $\delta(x,y) = \delta_2(x,y)$ on $\mathbb{R}^2$ is equal to the product $\delta_1(x)\delta_1(y)$, where δ_1 is the delta distribution on $\mathbb{R}$.

11.5.5. Obtain a fundamental solution $E(x,t)$ for the 1-dimensional heat or diffusion operator $-D_x^2 + D_t$:

(i) By means of 1-dimensional Fourier transformation applied to the equation

$$-E_{xx} + E_t = \delta_2(x,t) = \delta_1(x)\delta_1(t);$$

(ii) by means of 2-dimensional Fourier transformation.

11.5.6. Use n-dimensional Fourier transformation to obtain a fundamental solution $E(x,t)$ for the n-dimensional heat operator $-\Delta_n + D_t = -(D_1^2 + \cdots + D_n^2) + D_t$.

11.5.7. Setting $|x| = r$, show that on $\mathbb{R}^2$,

$$\frac{2\varepsilon}{(r^2+\varepsilon)^2} \to 2\pi\delta(x) \quad \text{as } \varepsilon \searrow 0.$$

Next verify that

$$E(x) = \lim_{\varepsilon\searrow 0} \frac{1}{4\pi} \log(r^2+\varepsilon) = \frac{1}{2\pi} \log r$$

is a fundamental solution for the Laplacian Δ on $\mathbb{R}^2$.

11.5.8. Setting $z = x + iy$ in $\mathbb{C} \approx \mathbb{R}^2$, verify that

$$E(x,y) = \frac{1}{\pi z}$$

is a fundamental solution for the Cauchy–Riemann operator

$$\frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{1}{i}\frac{\partial}{\partial y}\right).$$

Hint. One may use the fact that

$$\frac{1}{z} = \left(\frac{\partial}{\partial x} + \frac{1}{i}\frac{\partial}{\partial y}\right) \log r.$$

11.6. Functions on $\mathbb{R}^2$ with circular symmetry

Equation (11.5.5) left us with the problem to evaluate the (reflected) Fourier transform of $e^{-z\rho}$, where $z > 0$ is a parameter. This is a special case of the following problem: Compute the Fourier transform of a function

$g(\xi,\eta) = G(\rho)$ with *circular* or rotational symmetry. We will treat this problem for general functions $G(\rho)$ on $\mathbb{R}^2$.

For the computation of the Fourier integral

$$\int_{\mathbb{R}^2} G(\rho)e^{\pm i(x\xi+y\eta)}d\xi d\eta$$

we introduce polar coordinates, setting $\xi = \rho\cos\phi$, $\eta = \rho\sin\phi$, so that $d\xi d\eta$ becomes $\rho d\rho d\phi$; we also set $x = r\cos\theta$, $y = r\sin\theta$. Thus

$$\begin{aligned}\mathcal{F}^{\xi,\eta}[G(\rho)](x,y) &= \mathcal{F}_R^{\xi,\eta}[G(\rho)](x,y) = \int_{\mathbb{R}^2} G(\rho)e^{i(x\xi+y\eta)}d\xi d\eta \\ &= \int_0^\infty G(\rho)\rho d\rho \int_{-\pi}^{\pi} e^{ir\rho\cos(\theta-\phi)}d\phi = \int_0^\infty G(\rho)\rho d\rho \int_{-\pi}^{\pi} e^{ir\rho\cos\phi}d\phi.\end{aligned}$$

The answer depends only on r, not on θ: the transform also has circular symmetry!

The inner integral may be calculated by termwise integration of a series:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{it\cos\phi}d\phi &= \frac{1}{\pi}\int_0^\pi = \frac{1}{\pi}\sum_{p=0}^\infty \int_0^\pi \frac{(it\cos\phi)^p}{p!}\,d\phi \\ &= \sum_{k=0}^\infty \frac{2k-1}{2k}\,\frac{2k-3}{2k-2}\cdots\frac{3}{4}\,\frac{1}{2}\,\frac{(it)^{2k}}{(2k)!} \\ &= \sum_{k=0}^\infty (-1)^k\,\frac{t^{2k}}{2^2 4^2\cdots(2k)^2} = J_0(t).\end{aligned}$$

Here we have recognized an old friend from Examples 8.1.6, the Bessel function $J_0(t)$ of order zero. [For $I_p = \int_0^\pi \cos^p\phi\,d\phi$ one may use the recurrence relation $I_p = -\{(p-1)/p\}I_{p-2}$, which is obtained through integration by parts.]

The answer for our repeated integral above is thus given by

THEOREM 11.6.1. *For functions $G(\rho)$, where $\rho = (\xi^2+\eta^2)^{\frac{1}{2}}$, one has*

$$\mathcal{F}^{\xi,\eta}[G(\rho)](x,y) = 2\pi\int_0^\infty G(\rho)\rho J_0(r\rho)d\rho, \quad r = (x^2+y^2)^{\frac{1}{2}}. \tag{11.6.1}$$

We return now to our special case in (11.5.5). Using (11.6.1), one finds

$$P(x,y,z) = \frac{1}{4\pi^2}\,\mathcal{F}^{\xi,\eta}[e^{-z\rho}] = \frac{1}{2\pi}\int_0^\infty e^{-z\rho}\rho J_0(r\rho)d\rho$$
$$= \frac{1}{2\pi r^2}\int_0^\infty e^{-(z/r)t} t J_0(t)dt.$$

To evaluate the integral we make use of the Laplace transform of the Bessel function $J_0(t)$,

$$\mathcal{L}[J_0](s) = \int_0^\infty J_0(t)e^{-st}dt = \text{p.v.}\,(s^2+1)^{-\frac{1}{2}}, \quad \operatorname{Re} s > 0; \tag{11.6.2}$$

cf. Exercise 11.2.8. Differentiation with respect to s next gives

$$\int_0^\infty tJ_0(t)e^{-st}dt = \frac{s}{(s^2+1)^{\frac{3}{2}}}.$$

Substituting $s = z/r$ and $r = (x^2+y^2)^{\frac{1}{2}}$, one finally obtains

$$P(x,y,z) = \frac{1}{2\pi}\,\frac{z}{(x^2+y^2+z^2)^{\frac{3}{2}}}. \tag{11.6.3}$$

The solution of the Dirichlet problem for the upper half-space H in Example 11.5.2 can now be obtained from formula (11.5.6):

$$u(x,y,z) = \int_{\mathbb{R}^2} f(s,t)\,\frac{1}{2\pi}\,\frac{z}{\{(x-s)^2+(y-t)^2+z^2\}^{\frac{3}{2}}}\,dsdt$$
$$= \int_{\mathbb{R}^2} f(x-s,y-t)\,\frac{1}{2\pi}\,\frac{z}{(s^2+t^2+z^2)^{\frac{3}{2}}}\,dsdt. \tag{11.6.4}$$

Verification. Formula (11.6.4) expresses $u(x,y,z)$ as the *Poisson integral* of f for the upper half-space H in $\mathbb{R}^3$. Taking f locally integrable and bounded, the first integral may be used to verify that u satisfies Laplace's equation for $z > 0$. For bounded continuous f, the second integral will show that $u(x,y,z) \to f(x_0,y_0)$ as $(x,y,z) \to (x_0,y_0,0)$ from H.

Exercises. 11.6.1. Verify the final statement above.

11.7. General Fourier problem with spherical symmetry

Let Q be a $1-1$ linear transformation of $\mathbb{R}^n$; we also write Q for the representing $n\times n$ [invertible real] matrix. If $Qx = y$ then $x = Q^{-1}y$, and

$$\xi\cdot x = \xi\cdot Q^{-1}y = (Q^{-1})^{\text{tr}}\xi\cdot y = R\xi\cdot y,$$

where we have written R for the transpose of the matrix Q^{-1}. The Jacobi determinant of the transformation $x = Q^{-1}y$ is $\det Q^{-1} = \det R$. For f in $L^1(\mathbb{R}^n)$, the composition $f \circ Q$ is also in L^1, and by the transformation rule for integrals,

$$\begin{aligned} \mathcal{F}(f \circ Q)(\xi) &= \int_{\mathbb{R}^n} f(Qx) e^{-i\xi\cdot x} dx \\ &= \int_{\mathbb{R}^n} f(y) e^{-iR\xi\cdot y} |\det R| dy = |\det R| \, (\mathcal{F}f)(R\xi). \end{aligned} \tag{11.7.1}$$

This rule will hold for *tempered distributions* as well. Indeed, such distributions T are limits of well-behaved functions, and for consistency with the case of integrals, one has to define

$$\langle T(Qx), \phi(x) \rangle \; = \; \left\langle T(y), \phi(Q^{-1}y) |\det R| \right\rangle . \tag{11.7.2}$$

For tempered distributions rule (11.7.1) then follows from the definition of $\mathcal{F}T$; see (11.4.8).

Suppose now that the function f in L^1 is *spherically symmetric*, that is, $f(x)$ depends only on the length $|x| = r$. An equivalent statement is that $f(Px) \equiv f(x)$ for all orthogonal transformations P. [Recall the definition $P^{-1} = P^{\mathrm{tr}}$, so that $\det P = \pm 1$ and $R = P$.] Then (11.7.1) shows that

$$(\mathcal{F}f)(\xi) = \mathcal{F}(f \circ P)(\xi) = (\mathcal{F}f)(P\xi), \tag{11.7.3}$$

so that $\mathcal{F}f$ also has spherical symmetry. For distributions T one will *define* spherical symmetry by the condition $T(Px) = T(x)$, so that the conclusion is the same:

PROPOSITION 11.7.1. *For a spherically symmetric distribution T, the Fourier transform $\mathcal{F}T$ is also spherically symmetric.*

Before we prove an evaluation theorem, we need some auxiliary results.

PROPOSITION 11.7.2. *For spherically symmetric functions $h(x) = H(r)$ in $\mathbb{R}^k$ of at most polynomial growth* [h of class $\mathcal{P}$], *one has*

$$\int_{|x|<A} h(x) dx = \sigma_k \int_0^A H(r) r^{k-1} dr. \tag{11.7.4}$$

Here σ_k denotes the area of the unit sphere $S_1 = \{|x| = 1\}$ in $\mathbb{R}^k$,

$$\sigma_k = 2\pi^{k/2}/\Gamma(k/2) \qquad \left[\text{note that } \Gamma(1/2) = \pi^{1/2}\right]. \tag{11.7.5}$$

The proof is a straightforward application of Fubini's theorem. In the case of spherically symmetric $h(x)$, the volume element dx in $\mathbb{R}^k$ may be replaced by $d\sigma(S_r)\cdot dr$, where $d\sigma(S_r)$ denotes the area element of the sphere $S_r = S(0,r)$ in $\mathbb{R}^k$. By similarity, $d\sigma(S_r) = r^{k-1}d\sigma(S_1)$. Finally, the total area $\sigma_k = \sigma(S_1)$ in $\mathbb{R}^k$ is equal to $2\pi^{k/2}/\Gamma(k/2)$; cf. Exercise 11.7.2.

DEFINITION 11.7.3. The Bessel functions $J_\nu(t)$ $(\nu > -1)$ are given by the power series

$$J_\nu(t) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{\nu+2k}k!\,\Gamma(\nu+k+1)}\, t^{\nu+2k}. \tag{11.7.6}$$

The cases $\nu = \pm 1/2$ are special:

$$J_{-1/2}(t) = \left(\frac{2}{\pi t}\right)^{1/2} \cos t, \quad J_{1/2}(t) = \left(\frac{2}{\pi t}\right)^{1/2} \sin t. \tag{11.7.7}$$

PROPOSITION 11.7.4. *For $J_\nu(t)$ one has the following integral representations when $\nu > -1/2$:*

$$\begin{aligned} J_\nu(t) &= \frac{(t/2)^\nu}{\Gamma(1/2)\Gamma(\nu+1/2)} \int_0^\pi e^{\pm it\cos\theta} \sin^{2\nu}\theta\, d\theta \\ &= \frac{(t/2)^\nu}{\Gamma(1/2)\Gamma(\nu+1/2)} \int_{-1}^1 (1-s^2)^{\nu-1/2} e^{\pm its} ds. \end{aligned}$$

Also, there are constants β_ν such that for $t \to \infty$,

$$J_\nu(t) = \left(\frac{2}{\pi t}\right)^{1/2} \cos(t-\beta_\nu) + \mathcal{O}(t^{-3/2}).$$

The second integral formula may be derived with the aid of Laplace transformation; cf. Exercise 12.4.14. One may *verify* the formulas by using Euler's *Beta function*,

$$\begin{aligned} B(p,q) &\stackrel{\text{def}}{=} \int_0^1 x^{p-1}(1-x)^{q-1}dx \qquad (p,\, q > 0) \\ &= 2\int_0^{\pi/2} \cos^{2p-1}\theta\, \sin^{2q-1}\theta\, d\theta = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \end{aligned} \tag{11.7.8}$$

For the asymptotic result, cf. Exercise 8.2.5.

THEOREM 11.7.5. *Let* $f(x) = F(r)$ (with $r = |x|$) *be a spherically symmetric function of class* $\mathcal{P}$ *on* $\mathbb{R}^n$. *Then*

$$\begin{aligned} g(\xi) &= \widehat{f}(\xi) = \widehat{f}_R(\xi) = G(\rho) \\ &= \mathcal{S}' \lim_{A\to\infty} (2\pi)^{n/2}\rho^{1-n/2} \int_0^A F(r) r^{n/2} J_{(n/2)-1}(\rho r)dr \\ &= \mathcal{S}' \lim_{\varepsilon\searrow 0} (2\pi)^{n/2}\rho^{1-n/2} \int_0^\infty e^{-\varepsilon r} F(r) r^{n/2} J_{(n/2)-1}(\rho r)dr, \quad \rho = |\xi|. \end{aligned}$$

PROOF. We will derive the first formula for $\widehat{f}(\xi)$. By (11.4.9) it will suffice to compute the transform of the truncated function f_A,

$$\widehat{f}_A(\xi) = \int_{|x|<A} F(r) e^{-i\xi\cdot x} dx.$$

By (11.7.3) $\widehat{f}_A(\xi)$ depends only on $|\xi| = \rho$, hence (thinking of $n \geq 2$),

$$\widehat{f}_A(\xi_1, \xi_2, \cdots, \xi_n) = \widehat{f}_A(\rho, 0, \cdots, 0) = \int_{|x|<A} F(r) e^{-i\rho x_1} dx.$$

We now split $x = (x_1, x_2, \cdots, x_n)$ as (x_1, x'), where $x' = (x_2, \cdots, x_n)$, and apply Fubini's theorem to the last integral to obtain

$$\widehat{f}_A(\xi) = \int_{|x_1|<A} e^{-i\rho x_1} dx_1 \int_{|x'|<\sqrt{A^2 - x_1^2}} F\left(\sqrt{x_1^2 + |x'|^2}\right) dx'.$$

In the final inner integral x_1 is constant, hence there the integrand depends only on $|x'| = r'$. To that integral we apply Proposition 11.7.2, with x replaced by x', r by r', A by $A' = \sqrt{A^2 - x_1^2}$, and $H(r') = F\big(\sqrt{x_1^2 + (r')^2}\big)$, $k = n - 1$. One thus finds

$$\widehat{f}_A(\xi) = \int_{|x_1|<A} e^{-i\rho x_1} dx_1 \cdot \sigma_{n-1} \int_0^{A'} F\left(\sqrt{x_1^2 + (r')^2}\right) (r')^{n-2} dr'.$$

This repeated integral may also be written as a double integral over the semidisc in the (x_1, r') plane given by $x_1^2 + (r')^2 < A^2$, $r' > 0$. For the evaluation of that double integral we introduce polar coordinates $x_1 = r\cos\theta$, $r' = r\sin\theta$, $0 < r < A$, $0 < \theta < \pi$:

$$\widehat{f}_A(\xi) = \sigma_{n-1} \int_0^A \int_0^\pi e^{-i\rho r\cos\theta} F(r)(r\sin\theta)^{n-2} r\, dr\, d\theta.$$

Here the integration with respect to θ can be carried out with the aid of Proposition 11.7.4. Taking $\nu = (n/2)-1$ (with $n \geq 2$) and $t = \rho r$, one finds

$$\int_0^\pi e^{-i\rho r\cos\theta}(\sin\theta)^{n-2}d\theta = \frac{\Gamma(1/2)\Gamma\{(n-1)/2\}}{(\rho r/2)^{(n/2)-1}}\, J_{(n/2)-1}(\rho r).$$

As a result,

$$\begin{aligned}\widehat{f}_A(\xi) = \sigma_{n-1}&2^{(n/2)-1}\Gamma(1/2)\Gamma\{(n-1)/2\}\rho^{1-n/2}\\ &\times\int_0^A F(r)r^{n/2}\, J_{(n/2)-1}(\rho r)dr.\end{aligned}$$

Using formula (11.7.5) for σ_{n-1}, this gives the desired result for $n \geq 2$. For $n = 1$ the result may be obtained by a simple direct computation. $\square$

Exercises. 11.7.1. Let δ denote the delta distribution on $\mathbb{R}^n$. Prove that

(i) $\delta(x)$ is spherically symmetric;

(ii) $\delta(\lambda x) = (1/\lambda^n)\,\delta(x), \quad \lambda > 0.$

Hint. The expression $< T(\lambda x), \phi(x) >$ is of course defined as if it is an ordinary integral $\int T(\lambda x)\phi(x)dx$ over $\mathbb{R}^n$.

11.7.2. Prove by induction that the volume of the ball $B_k(0,r)$ in $\mathbb{R}^k$ is equal to $\pi^{k/2}r^k/\Gamma\{(k/2)+1\}$. Deduce that the area of the sphere $S_r = S_k(0,r)$ in $\mathbb{R}^k$ is equal to $2\pi^{k/2}r^{k-1}/\Gamma(k/2)$.

Hint. Setting $x = (x_1, x_2, \cdots, x_k) = (x_1, x')$, one has

$$\begin{aligned}\operatorname{vol} B_k(0,r) &= \int_{|x|<r} dx = \int_{|x_1|<r} dx_1 \int_{|x'|<r'=\sqrt{r^2-x_1^2}} dx'\\ &= \int_{-r}^{r} \operatorname{vol} B_{k-1}(0,r')\, dx_1.\end{aligned}$$

11.7.3. (i) Verify the formulas for $J_{\pm 1/2}(t)$ in (11.7.7).

(ii) Prove the 'recurrence relation'

$$\frac{1}{z}\frac{d}{dz}\left\{z^{-\nu}J_\nu(z)\right\} = -z^{-\nu-1}J_{\nu+1}(z).$$

(iii) Compute $J_{3/2}(t)$.

11.7.4. Give direct proofs for the cases $n = 1$ and $n = 2$ of Theorem 11.7.5.

11.7.5. (i) Show that the function $f(x) = r^\alpha$ with $r = |x|$ on $\mathbb{R}^n$ belongs to the class $\mathcal{P}$ if and only if $\operatorname{Re}\alpha > -n$.

(ii) Prove that for $-n < \alpha < -(n+1)/2$, and even for $-n < \alpha < -(n-1)/2$,

$$\mathcal{F} r^{\alpha} = C_{\alpha,n} \rho^{-\alpha-n}, \quad \rho = |\xi|. \tag{11.7.9}$$

(iii) Show with the aid of a well-chosen test function that

$$C_{\alpha,n} = 2^{\alpha+n} \pi^{n/2} \frac{\Gamma\{(\alpha+n)/2\}}{\Gamma(-\alpha/2)}.$$

11.7.6. (i) Show that the function r^{α} on $\mathbb{R}^n$ depends analytically on α for $\operatorname{Re}\alpha > -n$ in the sense that $h_{\phi}(\alpha) = < r^{\alpha}, \phi >$ is an analytic function of α for every test function ϕ.

(ii) Use the relation

$$r^{\alpha} = \frac{1}{(\alpha+2)(\alpha+1)} \Delta\, r^{\alpha+2}$$

to define r^{α} recursively as a tempered distribution on $\mathbb{R}^n$ which depends analytically on α throughout the domain $\mathbb{C} \setminus \{-n, -n-2, -n-4, \cdots\}$. [The points $-2, -4, \cdots$ require special attention.]

(iii) Prove that $\mathcal{F} r^{\alpha}$ also depends analytically on α, and use this fact to extend formula (11.7.9) to all values of $\alpha \neq -n, -n-2, -n-4, \cdots$ and $\neq 0, 2, 4, \cdots$.

11.7.7. Determine $\mathcal{F}_R(1/\rho^2)$ in $\mathbb{R}^n$, $\rho = |\xi|$, when $n \geq 3$. Use the answer to obtain the fundamental solution $E(x)$ for (minus) the Laplace operator in $\mathbb{R}^n$ that tends to zero as $r = |x| \to \infty$. Put the answer into the final form

$$E(x) = \frac{1}{(n-2)\sigma_n} \frac{1}{r^{n-2}} \qquad (\sigma_n = \text{area of } S_1 \text{ in } \mathbb{R}^n).$$

CHAPTER 12

Other integral transforms

There are so-called half-line integral transformations (related to Fourier transformation) that can be applied to a large class of functions defined on $\mathbb{R}^+ = (0, \infty)$. The most important of these is Laplace transformation, which is especially useful for the treatment of initial value problems. It maps functions on $\mathbb{R}^+$ onto analytic functions in a right half-plane, to which one can apply methods of Complex Analysis.

In the following, applications of integral transforms to ordinary and partial differential equations will play an important role. In the n-dimensional case, the most interesting application of our integral transforms involves the wave equation. We will see in Section 12.6 that *communication* governed by that equation works poorly in even dimensions, and *works* really well *only* in $\mathbb{R}^3$!

We will also discuss Fourier cosine and sine transforms, and in the next chapter, two-sided Laplace transformation.

12.1. Laplace transforms

Functions f on $\mathbb{R}^+ = (0, \infty)$ that are integrable over finite intervals $(0, A)$ and of at most exponential growth towards infinity have a Laplace transform [also called one-sided Laplace transform]

$$g(s) = (\mathcal{L}f)(s) \stackrel{\text{def}}{=} \int_0^\infty f(t)e^{-st}dt, \quad s = \sigma + i\tau. \tag{12.1.1}$$

EXAMPLES 12.1.1. For $f(t) = e^{at}$, $a \in \mathbb{C}$ and $\operatorname{Re} s > \operatorname{Re} a$,

$$(\mathcal{L}f)(s) = \int_0^\infty e^{at}e^{-st}dt = \left[\frac{e^{(a-s)t}}{a-s}\right]_{t=0}^{\infty} = \frac{1}{s-a}.$$

Indeed, for $\operatorname{Re}(s-a) = \delta > 0$ one has $\left|e^{(a-s)t}\right| = e^{-\delta t}$.

From this one may derive that

$$\begin{aligned}\mathcal{L}[\sin bt](s) &= \mathcal{L}\left[\frac{1}{2i}\left(e^{ibt} - e^{-ibt}\right)\right](s) \\ &= \frac{1}{2i}\left(\frac{1}{s-ib} - \frac{1}{s+ib}\right) = \frac{b}{s^2+b^2}.\end{aligned}$$

If b is real this holds for all $s \in \mathbb{C}$ with $\operatorname{Re} s > 0$.

The (one-sided) Laplace transform is a continuous analog of a power series in $z = e^{-s}$,

$$\sum_0^\infty a_n z^n = \sum_0^\infty a_n e^{-ns}. \tag{12.1.2}$$

The typical domain of convergence for such a series is a circular disc in the complex z-plane:

$$\left\{|z| = \left|e^{-s}\right| = e^{-\sigma} < R\right\}.$$

In terms of s, the domain of convergence becomes a right half-plane, given by

$$\sigma = \operatorname{Re} s > -\log R.$$

[If $R = +\infty$, the half-plane becomes the whole plane.] In this domain, the sum of the series (12.1.2) is analytic.

Similarly, the (one-sided) Laplace transform $g(s) = (\mathcal{L}f)(s)$ will be analytic in its half-plane of convergence.

THEOREM 12.1.2. *Suppose that the Laplace integral* (12.1.1) *converges* [as a Lebesgue integral, hence converges absolutely] *at the point $s = a = \alpha + i\beta$. Then it converges for every $s \in \mathbb{C}$ with $\sigma = \operatorname{Re} s \geq \alpha$. The function $g(s) = (\mathcal{L}f)(s)$ is continuous and bounded on the closed half-plane $\{\sigma \geq \alpha\}$, and it tends to zero as $\sigma = \operatorname{Re} s \to +\infty$.The transform $g(s)$ is differentiable in the complex sense – hence analytic – throughout the open half-plane $\{\sigma > \alpha\}$. There one has*

$$g'(s) = -\int_0^\infty t f(t) e^{-st} dt. \tag{12.1.3}$$

PROOF. (i) The integral (12.1.1) will converge for all s in the closed half-plane $\{\sigma = \operatorname{Re} s \geq \alpha\}$. Indeed, for such s, the integrand is the product of an integrable and a bounded continuous function,

$$f(t)e^{-st} = f(t)e^{-at} \cdot e^{(a-s)t}, \quad \left|f(t)e^{-st}\right| = \left|f(t)e^{-at}\right| e^{(\alpha-\sigma)t}. \tag{12.1.4}$$

A direct estimate shows that $\int_0^A f(t)e^{-st}dt$ is continuous in s on $\mathbb{C}$. As $A \to \infty$,

$$\int_0^A f(t)e^{-st}dt \to \int_0^\infty f(t)e^{-st}$$

uniformly in s for $\operatorname{Re} s \geq \alpha$. Hence $g(s)$ is continuous there. The boundedness of $g(s)$ for $\sigma \geq \alpha$ follows immediately from (12.1.4). That $g(s) \to 0$ as $\sigma \to \infty$ may be derived by dominated convergence, or directly from the inequality

$$\begin{aligned} |g(s)| = \left| \int_0^\infty f(t)e^{-st}dt \right| &\leq \int_0^\delta \left| f(t)e^{-at} \right| dt \\ &+ \int_\delta^\infty \left| f(t)e^{-at} \right| dt \cdot e^{(\alpha-\sigma)\delta}, \quad \sigma \geq \alpha. \end{aligned}$$

The right-hand side can be made small by fixing a small $\delta > 0$ and then taking σ large.

(ii) We will prove that the complex derivative $g'(s)$ exists throughout the open half-plane $\{\sigma > \alpha\}$. Fix $s \in \mathbb{C}$ with $\sigma = \operatorname{Re} s > \alpha$. For $t \geq 0$ and complex $h \neq 0$, one has

$$\begin{aligned} \left| \frac{e^{-ht} - 1}{h} + t \right| &= \left| \frac{1}{h} \left\{ \frac{(-ht)^2}{2!} + \frac{(-ht)^3}{3!} + \frac{(-ht)^4}{4!} + \cdots \right\} \right| \\ &\leq |h| \, \frac{1}{2} \, t^2 \left(1 + \frac{|h|t}{3} + \frac{(|h|t)^2}{3 \cdot 4} + \cdots \right) \\ &\leq |h| \, \frac{1}{2} \, t^2 \, e^{|h|t} \leq |h| \, \frac{e^{\delta t}}{\delta^2} \, e^{|h|t}, \quad \forall \, \delta > 0. \end{aligned}$$

Hence for fixed $\delta < \sigma - \alpha$ and $|h| \leq \sigma - \alpha - \delta$, so that $\delta + |h| - \sigma \leq -\alpha = -\operatorname{Re} a$,

$$\begin{aligned} &\left| \frac{g(s+h) - g(s)}{h} + \int_0^\infty t f(t) e^{-st} dt \right| \\ &= \left| \int_0^\infty \left(\frac{e^{-ht} - 1}{h} + t \right) f(t) e^{-st} dt \right| \\ &\leq |h| \, \frac{1}{\delta^2} \int_0^\infty |f(t)| e^{(\delta + |h| - \sigma)t} dt \leq |h| \, \frac{1}{\delta^2} \int_0^\infty \left| f(t) e^{-at} \right| dt. \end{aligned}$$

The final expression tends to zero as $h \to 0$. Conclusion: g is differentiable at s in the complex sense, with derivative $g'(s)$ as in (12.1.3). By Complex

Analysis, g then is analytic: it has local representations by convergent power series. It is attractive to prove this directly; cf. Exercise 12.1.3. $\square$

EXAMPLE 12.1.3. Let $\operatorname{Re} a > -1$, $f(t) = t^a = \text{p.v.}\, t^a$, $t > 0$. The product $t^a e^{-st}$ is in $L^1(\mathbb{R}^+)$ for all complex s with $\operatorname{Re} s > 0$. For *real* $s > 0$,

$$\begin{aligned} g(s) &= \int_0^\infty t^a e^{-st} dt = \int_0^\infty \left(\frac{v}{s}\right)^a e^{-v} \frac{dv}{s} \\ &= \Gamma(a+1) s^{-a-1} = \Gamma(a+1)\, \text{p.v.}\, s^{-a-1}. \end{aligned} \tag{12.1.5}$$

By Theorem 12.1.2, the left-hand side has an analytic extension $g = \mathcal{L}f$ to the right half-plane $\{\sigma = \operatorname{Re} s > 0\}$. The same is true for the final member, so that we have two analytic functions for $\operatorname{Re} s > 0$ that agree on the positive real axis. Hence by the Uniqueness Theorem for analytic functions, the final member gives the Laplace transform of t^a for all s with $\sigma > 0$.

For the applications, the most important property of Laplace transformation involves its action on derivatives:

PROPOSITION 12.1.4. *Let f be equal to an indefinite integral on $[0, \infty)$ and suppose that $f'(t)e^{-at}$ is in $L^1(\mathbb{R}^+)$ for some constant a. Then*

$$(\mathcal{L}f')(s) = s(\mathcal{L}f)(s) - f(0) \quad \textit{for } \operatorname{Re} s > \max\{\operatorname{Re} a, 0\}. \tag{12.1.6}$$

PROOF. Setting $\alpha = \max\{\operatorname{Re} a, 0\}$, the product $f'(v)e^{-\alpha v}$ will be in $L^1(\mathbb{R}^+)$. Thus

$$f(t) = f(0) + \int_0^t f'(v) dv = f(0) + \int_0^t f'(v) e^{-\alpha v} \cdot e^{\alpha v} dv,$$

$$|f(t)| \le |f(0)| + e^{\alpha t} \int_0^\infty |f'(v)| e^{-\alpha v} dv \le C_\alpha e^{\alpha t}, \quad \forall\, t \in \mathbb{R}^+.$$

Taking $\operatorname{Re} s = \sigma > \alpha$, integration by parts now gives the desired result:

$$\begin{aligned} \int_0^\infty f'(t) e^{-st} dt &= \Big[f(t) e^{-st}\Big]_{t=0}^\infty - \int_0^\infty f(t)(-s) e^{-st} dt \\ &= -f(0) + s \int_0^\infty f(t) e^{-st} dt. \end{aligned}$$

Indeed, one has $|f(t)e^{-st}| \le C_\alpha e^{(\alpha - \sigma)t}$, so that the last integral converges. The final bound also shows that the integrated term reduces to $-f(0)$. $\square$

Corollary 12.1.5. *For f as in Proposition 12.1.4,*

$$\lim_{s\to+\infty} s(\mathcal{L}f)(s) \quad \textit{exists and} \quad = f(0). \tag{12.1.7}$$

Indeed, $(\mathcal{L}f')(s) \to 0$ as $\sigma = \operatorname{Re} s \to \infty$: apply Theorem 12.1.2 to f'.

Exercises. 12.1.1. Compute $\mathcal{L}[\cos bt]$:

(i) directly;

(ii) from the formula $\cos bt = (d/dt)(\sin bt)/b$.

12.1.2. Compute $\mathcal{L}[e^{at}\cos bt]$ and $\mathcal{L}[e^{at}\sin bt]$.

12.1.3. Let $f(t)e^{-\alpha t}$ (with $\alpha \in \mathbb{R}$) be in $L^1(\mathbb{R}^+)$. Choosing any $s_0 = \sigma_0 + i\tau_0$ with $\sigma_0 > \alpha$, prove by termwise integration of a suitable expansion that for $|s - s_0| < \sigma_0 - \alpha$, and even for $|s - s_0| \le \sigma_0 - \alpha$,

$$g(s) = \int_0^\infty f(t)e^{-st}dt = \sum_0^\infty c_n(s - s_0)^n,$$

where

$$c_n = \frac{1}{n!}\int_0^\infty (-t)^n f(t)e^{-s_0 t}dt, \quad n = 0, 1, 2, \cdots.$$

12.1.4. Use the power series for $J_0(t)$ and termwise integration to show that for $\operatorname{Re} s = \sigma > 1$,

$$(\mathcal{L}J_0)(s) = \text{p.v.}\,(s^2+1)^{-\frac{1}{2}}.$$

Extend the result to all s with real part $\sigma > 0$.

12.2. Rules for Laplace transforms

We consider functions f of at most exponential growth on $\mathbb{R}^+$ in the sense that $f(t)e^{-\alpha t}$ is in $L^1(\mathbb{R}^+)$ for some real constant α. Taking $s = \sigma + i\tau$, we set $(\mathcal{L}f)(s) = g(s)$ for $\sigma \ge \alpha$.

Discussion of the Table below. Rules (i)–(iii) follow immediately from the Definition of the Laplace transform and rule (v) follows from Theorem 12.1.2. A sufficient condition for rule (iv) was given in Proposition 12.1.4. For the *half-line convolution* in rule (vi) we have the following

Proposition 12.2.1. (i) *Let the functions $f_j(t)$ be locally integrable on* $\mathbb{R}^+$ [integrable over finite subintervals]. *Then the half-line convolution*

$$(f_1 * f_2)(t) = \int_0^t f_1(t-v)f_2(v)dv \tag{12.2.1}$$

exists almost everywhere and is locally integrable on $\mathbb{R}^+$.

	Original $f(t)$	Laplace trf $g(s)$	Remark
(i)	$f(\lambda t)$	$\dfrac{1}{\lambda}\, g\Big(\dfrac{s}{\lambda}\Big)$	$\lambda > 0$, $\sigma \geq \lambda\alpha$
(ii)	$f(t-\lambda)U(t-\lambda)$	$e^{-\lambda s}g(s)$	$\lambda \geq 0$
(iii)	$e^{\lambda t}f(t)$	$g(s-\lambda)$	$\lambda \in \mathbb{C}$, $\sigma \geq \alpha + \operatorname{Re}\lambda$
(iv)	$Df(t) = f'(t)$	$sg(s) - f(0)$	Prop 12.1.4
(v)	$tf(t)$	$-Dg(s) = -g'(s)$	Thm 12.1.2
(vi)	$(f_1 * f_2)(t)$	$g_1(s)g_2(s)$	$(f_1 * f_2)(t) = \int_0^t \cdots$

(ii) *Suppose now that* $f_j(t)e^{-\alpha_j t}$ *with real* α_j *is in* $L^1(\mathbb{R}^+)$, $j = 1, 2$. *Then the function* $(f_1 * f_2)(t)e^{-st}$ *will be integrable over* $\mathbb{R}^+$ *for* $\operatorname{Re} s = \sigma \geq \alpha = \max\{\alpha_1, \alpha_2\}$, *and*

$$\mathcal{L}(f_1 * f_2) = \mathcal{L}f_1 \cdot \mathcal{L}f_2 \quad \textit{for } \sigma \geq \alpha.$$

Proof. (i) Extending f_1 and f_2 to $\mathbb{R}$ by setting $f_j = 0$ on $\mathbb{R}^-$, we have

$$f_1(t-v)f_2(v) = 0 \quad \text{for } v < 0 \text{ and for } v > t,$$

hence the ordinary convolution $\int_{\mathbb{R}} f_1(t-v)f_2(v)dv$ reduces to 0 for $t < 0$ and to the half-line convolution (12.2.1) for $t \geq 0$.

In order to prove a.e. existence and integrability of the half-line convolution on $0 \leq t \leq A$, we temporarily redefine f_1, f_2 as equal to 0 for $t > A$. The modified f_1 and f_2 will be in $L^1(\mathbb{R})$ and hence the result for $0 \leq t \leq A$ follows from Proposition 9.4.3 for ordinary convolution.

(ii) Keeping $f_1 = f_2 = 0$ for $t < 0$, the hypothesis implies integrability of $f_j(t)e^{-st}$ over $\mathbb{R}$ when $\operatorname{Re} s \geq \alpha$, $j = 1, 2$. For such s, the function

$$\begin{aligned}(f_1 * f_2)(t)e^{-st} &= \int_{\mathbb{R}} f_1(t-v)e^{-s(t-v)} \cdot f_2(v)e^{-sv}dv \\ &= f_1(t)e^{-st} * f_2(t)e^{-st}\end{aligned}$$

will be integrable over $\mathbb{R}$ by Proposition 9.4.3. By the same proposition, the integral of the left-hand side over $\mathbb{R}$ will be equal to the product of the integrals of $f_1(t)e^{-st}$ and $f_2(t)e^{-st}$. $\square$

Example 12.2.2. The Laplace transform of the *Bessel function* J_0 can be obtained from the characterization of $J_0(t)$ as the solution of the initial value problem

$$ty'' + y' + ty = 0, \quad y(0) = 1, \;\; y'(0) = 0$$

that was discussed in Examples 8.1.3 and 8.1.6. The power series for $J_0(t)$ readily shows that $|J_0(t)| \le e^t$ for $t \ge 0$. Using the initial conditions, rules (iv) and (v) in the Table give

$$\mathcal{L}y' = s\mathcal{L}y - y(0) = s\mathcal{L}y - 1, \;\; \mathcal{L}y'' = s\mathcal{L}y' - y'(0) = s^2\mathcal{L}y - s,$$
$$\mathcal{L}[ty] = -(\mathcal{L}y)', \;\; \mathcal{L}[ty''] = -(\mathcal{L}y'')' = -s^2(\mathcal{L}y)' - 2s\mathcal{L}y + 1.$$

Thus, transforming our differential equation and taking $\operatorname{Re} s = \sigma > 1$, simple calculations will give

$$(s^2+1)(\mathcal{L}y)' + s\mathcal{L}y = 0, \quad \text{so that } \; \mathcal{L}y = c(s^2+1)^{-\frac{1}{2}}.$$

Choosing the principal value of the fractional power, we must have

$$s\mathcal{L}y = s \cdot cs^{-1}(1+1/s^2)^{-\frac{1}{2}} \to y(0) = 1 \quad \text{as} \;\; s \to +\infty;$$

see Corollary 12.1.5. Hence $c = 1$ and

$$(\mathcal{L}J_0)(s) = \text{p.v.}\,(s^2+1)^{-\frac{1}{2}}. \tag{12.2.2}$$

Our proof gives this for $\sigma > 1$, but by the boundedness of $J_0(t)$ and analytic continuation, the result will hold for all s with real part $\sigma > 0$. Cf. also Exercise 12.1.4.

Exercises. 12.2.1. Starting with the formula $\mathcal{L}1 = 1/s$, use the rules in the Table to compute $\mathcal{L}[e^{at}]$ and $\mathcal{L}[t^n e^{at}]$.

12.2.2. Compute $\mathcal{L}[\cos bt]$ from the initial value problem

$$y'' + b^2 y = 0, \quad y(0) = 1, \;\; y'(0) = 0.$$

Next use the Table to compute $\mathcal{L}[e^{at}\cos bt]$.

12.2.3. Give a direct proof of rule (vi) for the Laplace transform of a half-line convolution by inverting order of integration.

12.3. Inversion of the Laplace transformation

Here we will use the close connection between Laplace transformation and Fourier transformation. It will then be convenient to think of f as a function on $\mathbb{R}$ which vanishes on the negative real axis; as a reminder we sometimes write $f(t)U(t)$, where $U(t)$ is the unit step function, $1_+(t)$ [Examples 10.5.2].

THEOREM 12.3.1. *For functions f of at most exponential growth on $\mathbb{R}^+$, more precisely, $f(t)e^{-\alpha t}$ in $L^1(\mathbb{R}^+)$ for some real constant α, one has*

$$g(s) = (\mathcal{L}f)(s) = (\mathcal{L}f)(\sigma + i\tau) = \mathcal{F}\left[f(t)e^{-\sigma t}U(t)\right](\tau) \tag{12.3.1}$$

for $\sigma \geq \alpha$, $\tau \in \mathbb{R}$. Conversely one has the so-called complex inversion formula*,*

$$f(t)U(t) = \lim_{A\to\infty} \frac{1}{2\pi i} \int_{\sigma - iA}^{\sigma + iA} g(s)e^{ts}ds, \quad \sigma \geq \alpha. \tag{12.3.2}$$

The limit in (12.3.2) will exist pointwise at the points t where $f(t)U(t)$ is differentiable or satisfies a Hölder–Lipschitz condition. If $g(s)$ is integrable over the vertical line $\{\operatorname{Re} s = \sigma\}$, the limit is equal to an ordinary integral from $\sigma - i\infty$ to $\sigma + i\infty$. The limit relation will always hold in the sense of general distributions; see below. The corresponding limit relation for $f(t)e^{-\sigma t}U(t)$ holds in the sense of tempered distributions.

COROLLARY 12.3.2. *Laplace transformation is one to one on the class of functions f of at most exponential growth: if $\mathcal{L}f = 0$, then $f = 0$ in the sense that $f(t) = 0$ almost everywhere on $\mathbb{R}^+$. In particular, $f(t)$ must then vanish at every point of continuity.*

PROOF OF THEOREM 12.3.1. For (12.3.1) one need only observe that for $\sigma \geq \alpha$,

$$g(\sigma + i\tau) = \int_0^\infty f(t)e^{-\sigma t}e^{-i\tau t}dt = \int_{\mathbb{R}} f(t)e^{-\sigma t}U(t) \cdot e^{-i\tau t}dt.$$

Applying Fourier inversion to this formula, cf. Theorems 9.2.2 and 10.1.7, one finds that for $\sigma \geq \alpha$,

$$f(t)e^{-\sigma t}U(t) = \frac{1}{2\pi}\mathcal{F}_R[g(\sigma + i\tau](t) = \frac{1}{2\pi}\lim_{A\to\infty}\int_{-A}^{A} g(\sigma + i\tau)e^{it\tau}d\tau. \tag{12.3.3}$$

Here the limit is an ordinary limit at points t where the left-hand side is well-behaved; the limit relation always holds in the sense of $\mathcal{S}'$.

Multiplying both sides of (12.3.3) by $e^{\sigma t}$ and replacing τ by $\sigma + i\tau = s$ as variable of integration, one obtains the complex inversion formula

$$\begin{aligned} f(t)U(t) &= \frac{1}{2\pi}\lim_{A\to\infty}\int_{-A}^{A} g(\sigma + i\tau)e^{t(\sigma + i\tau)}d\tau \\ &= \lim_{A\to\infty}\frac{1}{2\pi i}\int_{\sigma - iA}^{\sigma + iA} g(s)e^{ts}ds. \end{aligned} \tag{12.3.4}$$

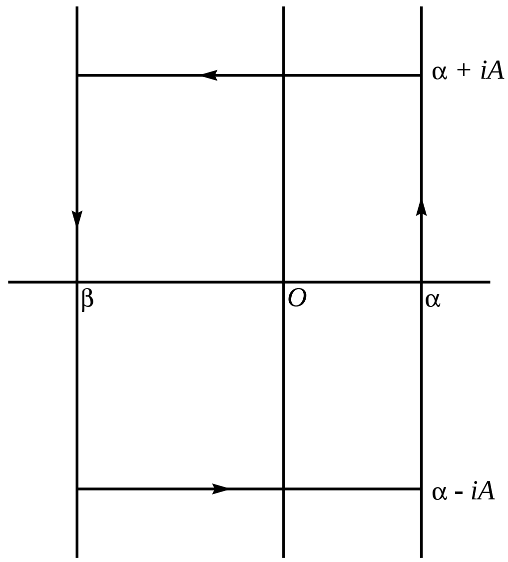

FIGURE 12.1

Here the limit relation holds in the sense of general distributions, that is, it holds relative to the test class of $\mathcal{C}^\infty$ functions ψ with compact support. [For such functions ψ, the products $e^{\sigma t}\psi$ will be in $\mathcal{S}$.] Cf. Chapter 13. $\square$

EXAMPLE 12.3.3. Let $g(s) = 1/(s^2+1)$, $\operatorname{Re} s > 0$. Then

$$f(t)U(t) = \left(\mathcal{L}^{-1}g\right)(t) = \frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty} \frac{1}{s^2+1}\, e^{ts} ds \qquad (\alpha > 0).$$

For $t > 0$ we move the vertical line of integration $\sigma = \text{const}$ far to the left, because $|e^{ts}|$ will become small there; cf. Figure 12.1. The Residue Theorem now gives

$$\begin{aligned} f(t) = \sum &\left(\text{residues of } \frac{1}{s^2+1}\, e^{ts} \text{ at } s = \pm i\right) \\ &+ \frac{1}{2\pi i}\int_{\beta-i\infty}^{\beta+i\infty} \frac{1}{s^2+1}\, e^{ts} ds \qquad (\beta < 0). \end{aligned} \tag{12.3.5}$$

Indeed, the integrals along horizontal segments $s = \sigma \pm iA$, $\alpha \geq \sigma \geq \beta$ will go to zero when $A \to \infty$. Letting β go to $-\infty$, the final integral tends to

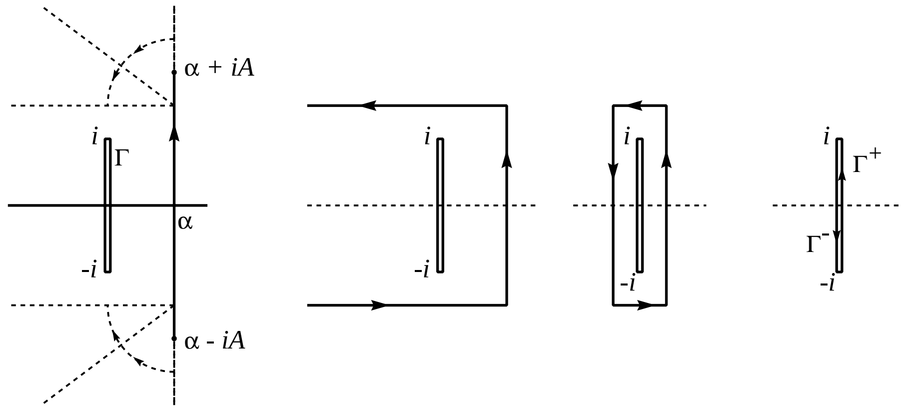

FIGURE 12.2

zero. [Being constant for $\beta < 0$, it is actually equal to zero.] Thus

$$f(t) = \frac{1}{2i}\left(e^{it} - e^{-it}\right) = \sin t \quad (t > 0).$$

EXAMPLE 12.3.4. Let $g(s) = \text{p.v.}\,(s^2+1)^{-\frac{1}{2}}$, $\operatorname{Re} s > 0$. We know from Example 12.2.2 that $g = \mathcal{L}[J_0U]$. Thus for $\alpha > 0$,

$$J_0(t)U(t) = \left(\mathcal{L}^{-1}g\right)(t) = \lim_{A\to\infty} \frac{1}{2\pi i}\int_{\alpha-iA}^{\alpha+iA} (s^2+1)^{-\frac{1}{2}}\, e^{ts} ds.$$

Since the function $(s^2+1)^{-\frac{1}{2}}$ is multi-valued, one has to be careful about moving paths of integration. Introducing a cut in the s-plane along the segment $\Gamma = [-i, i]$, one can define an analytic branch $g_1(s) = (s^2+1)^{-\frac{1}{2}}$ outside the cut; we consider the branch that behaves like $1/s$ at infinity. That branch can be considered as an analytic continuation of our principal value $g(s)$ on the half-plane $\operatorname{Re} s > 0$. It has continuous extensions to the two edges of the cut, except for the points $s = \pm 1$.

Keeping $t > 0$, we may successively deform our path of integration as indicated in Figure 12.2, finally contracting the path onto the edges of the cut Γ. Since $g_1(s)$ is positive on $\mathbb{R}^+$ and real on Γ, it will by continuity be positive on the right-hand edge Γ^+ of the cut. On the left-hand edge Γ_- of the cut, $g_1(s)$ will have the opposite values. Thus, setting $s = iv$ on the cut, we find that $g_1(s) = (1-v^2)^{-\frac{1}{2}}$ on Γ^+ and $g_1(s) = -(1-v^2)^{-\frac{1}{2}}$ on Γ^-.

As a result,

$$J_0(t) = \frac{1}{2\pi i}\int_{-i;\, s\in\Gamma^+}^{i} g_1(s)e^{ts}ds + \frac{1}{2\pi i}\int_{i;\, s\in\Gamma^-}^{-i} g_1(s)e^{ts}ds$$

$$= \frac{2}{2\pi}\int_{-1}^{1}(1-v^2)^{-\frac{1}{2}}e^{itv}dv = \frac{2}{\pi}\int_0^1 \frac{1}{\sqrt{1-v^2}}\cos tv\, dv. \tag{12.3.6}$$

This representation will be valid also for $t < 0$ [since J_0 is even] and for $t = 0$.

Observe that by (12.3.6), $|J_0(t)| \le J_0(0) = 1$ for $t \in \mathbb{R}$, while by the Riemann–Lebesgue lemma, $J_0(t) \to 0$ as $t \to \infty$.

12.4. Other methods of inversion

In practice, it may be more convenient to use other methods of inversion than Theorem 12.3.1. We mention several.

(i) *Use of tables of Laplace transforms.* The given function $g(s)$ may occur in a table, or if it does not occur itself, the rules in Section 12.2 may help out. For example, the question may be to determine $f = fU = \mathcal{L}^{-1}g$ when

$$g(s) = \frac{e^{-s}}{s+1} \qquad (\operatorname{Re} s > 0).$$

The list of transforms will surely contain the pair

$$f_0(t) = 1, \quad g_0(s) = \frac{1}{s} \qquad (\operatorname{Re} s > 0).$$

Applying the rules, one will obtain

$$\frac{1}{s+1} = \mathcal{L}\left[e^{-t}\right] = \mathcal{L}\left[e^{-t}U(t)\right], \quad e^s\frac{1}{s+1} = \mathcal{L}\left[e^{-(t-1)}U(t-1)\right].$$

(ii) *Decomposition into partial fractions.* Suppose

$$g(s) = \frac{P(s)}{Q(s)} \qquad (\operatorname{Re} s > \alpha),$$

where P and Q are polynomials. We may assume that $\deg P < \deg Q$ [so that $g(s) \to 0$ as $\operatorname{Re} s \to \infty$]. Factoring $Q(s) = C(s-a)^m(s-b)^n\cdots$ with distinct $a, b, \cdots$, one has

$$g(s) = \frac{A_0}{(s-a)^m} + \frac{A_1}{(s-a)^{m-1}} + \cdots + \frac{A_{m-1}}{s-a}$$
$$+ \frac{B_0}{(s-b)^n} + \cdots + \frac{B_{n-1}}{s-b} + \cdots.$$

Here A_k is the coefficient of $(s-a)^k$ in the power series for $(s-a)^m g(s)$ around the point a:

$$A_k = \frac{1}{k!} D^k\big\{(s-a)^m g(s)\big\}\Big|_{s=a}, \quad k = 0, 1, \cdots, m-1,$$

etc. One finally uses the standard formula

$$\frac{1}{(s-a)^p} = \mathcal{L}\left[\frac{t^{p-1}}{(p-1)!}\, e^{at} U(t)\right]. \tag{12.4.1}$$

Thus for example,

$$\begin{aligned}\frac{4}{(s^2+1)^2} &= \frac{-1}{(s-i)^2} + \frac{-i}{s-i} + \frac{-1}{(s+i)^2} + \frac{i}{s+i} \\ &= \mathcal{L}\big[(-te^{it} - ie^{it} - te^{-it} + ie^{-it})U(t)\big] \\ &= \mathcal{L}\big[(-2t\cos t + 2\sin t)U(t)\big] \qquad (\operatorname{Re} s > 0).\end{aligned}$$

(iii) *Termwise inverse transformation* when $g(s)$ is analytic at infinity [that is, $g(1/s$ analytic at $s = 0$]. Suppose that

$$g(s) = \sum_{n=0}^{\infty} \frac{a_n}{s^{n+1}} \quad \text{for } (\operatorname{Re} s > \alpha \geq 0.$$

Then

$$f(t)U(t) = \big(\mathcal{L}^{-1} g\big)(t) = \sum_0^\infty a_n \mathcal{L}^{-1}\left[\frac{1}{s^{n+1}}\right](t) = \sum_0^\infty a_n \frac{t^n}{n!}\, U(t).$$

Verification. The series for g will be (absolutely) convergent for $|s| > \alpha$, hence $|a_n| \leq C_\varepsilon(\alpha+\varepsilon)^n$ for every $\varepsilon > 0$. Thus

$$\sum_0^\infty \left| a_n \frac{t^n}{n!}\, e^{-st}\right| \leq C_\varepsilon \sum_0^\infty \frac{\{(\alpha+\varepsilon)t\}^n}{n!}\, e^{-\sigma t} = C_\varepsilon e^{(\alpha+\varepsilon-\sigma)t}.$$

Hence for $\operatorname{Re} s > \alpha+\varepsilon$, the series $\sum_0^\infty a_n(t^n/n!)e^{-st}$ may be integrated term by term over $\mathbb{R}^+$. The result will be the original formula for $g(s)$.

For example, for $\operatorname{Re} s > 1$,

$$\begin{aligned}&\text{p.v.}\,(s^2+1)^{-\frac12} = \frac{1}{s}\,\text{p.v.}\left(1 + \frac{1}{s^2}\right)^{-\frac12} \\ &= \frac{1}{s}\left\{1 - \frac{1}{2}\frac{1}{s^2} + \frac{(-\frac12)(-\frac32)}{2!}\frac{1}{s^4} + \frac{(-\frac12)(-\frac32)(-\frac52)}{3!}\frac{1}{s^6} + \cdots\right\}.\end{aligned}$$

Hence

$$\begin{aligned}&\mathcal{L}^{-1}\left\{\text{p.v.}\,(s^2+1)^{-\frac{1}{2}}\right\}\\&=\left\{1-\frac{1}{2}\frac{t^2}{2!}+\frac{1\cdot 3}{2!\,2^2}\frac{t^4}{4!}-\frac{1\cdot 3\cdot 5}{3!\,2^3}\frac{t^6}{6!}+\cdots\right\}U(t)\\&=\left\{1-\frac{t^2}{2^2}+\frac{t^4}{2^24^2}-\frac{t^6}{2^24^26^2}+\cdots\right\}U(t)=J_0(t)U(t);\end{aligned}$$

cf. Examples 8.1.6.

(iv) *Inverse transformation applied to a product.* Suppose $g = g_1g_2$ where $g_j = \mathcal{L}f_j = \mathcal{L}[f_jU]$. Then

$$fU = \mathcal{L}^{-1}g = \mathcal{L}^{-1}(g_1g_2) = f_1U * f_2U = (f_1 * f_2)U,$$

the half-line convolution. For example: Solve the initial value problem

$$y'' + y = f(t), \quad t > 0; \quad y(0) = y'(0) = 0.$$

Setting $f = 0$ for $t < 0$ and assuming that f is at most of exponential growth on $\mathbb{R}^+$, one will have $y = 0$ for $t < 0$ and by the rules in Section 12.2,

$$s^2\mathcal{L}y + \mathcal{L}y = \mathcal{L}f = \mathcal{L}[fU], \quad \text{so that} \quad \mathcal{L}y = \mathcal{L}f \cdot \frac{1}{s^2+1},$$

provided $\operatorname{Re} s$ is sufficiently large. It follows that

$$y(t) = f(t) * \sin t = \int_0^t f(v)\sin(t-v)\,dv \quad \text{for} \quad t \geq 0.$$

The solution makes sense for any locally integrable function f on $\mathbb{R}^+$.

Exercises. 12.4.1. Compute the inverse Laplace transforms of

$$\frac{1}{s^2}, \quad \frac{1}{s^2-1}, \quad \frac{1}{s}, \quad \frac{s}{s^2+1} \qquad (\operatorname{Re} s > \alpha):$$

(i) with the aid of the complex inversion formula;
(ii) with the aid of partial fractions.

12.4.2. Use Laplace transformation to solve the initial value problems

$$y'' + y = 0, \quad 0 < t < \infty; \quad y(0) = 0, \quad y'(0) = 1;$$
$$y'' - a^2y = 0, \quad 0 < t < \infty; \quad y(0) = 1, \quad y'(0) = 0.$$

12.4.3. Same question for the initial value problem given by the system

$$4y' - z' - 5z = 0, \quad 4y + z' + z = 4, \quad 0 < t < \infty;$$
$$y(0) = 1, \quad z(0) = 2.$$

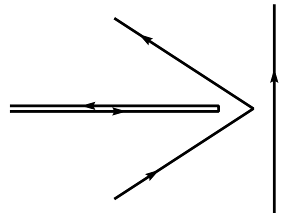

FIGURE 12.3

[Answer: $y = e^{-t}\sin 2t + 1, z = 2e^{-t}\cos 2t.$]

12.4.4. Determine $\mathcal{L}^{-1}[g(s)/s]$ if $g = \mathcal{L}f$ ($\mathrm{Re}\, s > \alpha \geq 0$).

12.4.5. Use Laplace transformation to solve the convolution equation

$$y(t) = e^{-t} + 2\int_0^t y(v)\cos(t-v)\,dv, \qquad t \geq 0.$$

12.4.6. Prove that

$$\int_0^t J_0(v)J_0(t-v)dv = \sin t, \qquad t \geq 0.$$

12.4.7. Solve Abel's integral equation

$$\int_0^t y(v)(t-v)^\beta dv = f(t), \quad t \geq 0 \quad (-1 < \beta < 0).$$

Hint. Solve first for $\int_0^t y(v)dv$.

12.4.8. Let $g(s)$ be an analytic function in a half-plane $\mathrm{Re}\, s > \alpha$ such that $|g(s)| \leq B/|s|^{1+\varepsilon}$, $\varepsilon > 0$, for $\mathrm{Re}\, s > \alpha' = \max\{\alpha, 0\}$. Prove that g is the Laplace transform of a continuous function $f(t)$ on $[0,\infty)$ with $f(0) = 0$.

12.4.9. Let $g(s)$ be an analytic function in a half-plane $\mathrm{Re}\, s > \alpha$ such that

$$\left|g(s) - \frac{c}{s}\right| \leq \frac{B}{|s|^{1+\varepsilon}} \;\; (\varepsilon > 0) \;\text{ for }\; \mathrm{Re}\, s > \alpha' \geq \max\{\alpha, 0\}.$$

Prove that g is the Laplace transform of a continuous function $f(t)$ on $[0,\infty)$ with $f(0) = c$.

12.4.10. Determine $\mathcal{L}^{-1}g$ when $g(s) = (1/\sqrt{s})e^{-x\sqrt{s}}$ ($\mathrm{Re}\, s > 0$), where $\sqrt{s}$ denotes the principal value of $s^{\frac{1}{2}}$ and x is a positive parameter. Use the answer $\left(1/\sqrt{\pi t}\right) e^{-x^2/(4t)}U(t)$ to compute $\mathcal{L}^{-1}\left[e^{-x\sqrt{s}}\right]$.

Hint. Make a cut in the s-plane along $\mathbb{R}^-$; cf. Figure 12.3. When $t > 0$, the path of integration for $\mathcal{L}^{-1}g(t)$ may be moved to the edges of the cut. Now set $s = w^2$ and finally set $w = iv$.

12.4.11. (*Heat conduction in semi-infinite medium*) Solve the boundary value problem

$$u_{xx} = u_t, \;\; x > 0, \; t > 0; \;\; u(x,0) = 0, \; x > 0; \;\; u(x,t) \;\text{ bounded;}$$
$$u(0,t) = f(t), \; t > 0.$$

Hint. Introduce the Laplace transform

$$v(x,s) = \mathcal{L}^t[u(x,t)](s) = \int_0^\infty u(x,t)e^{-st}dt.$$

12.4.12. Solve the boundary value problem

$$u_{xx} = \frac{1}{c^2}\,u_{tt}, \;\; x > 0, \; t > 0; \;\; u(x,0) = u_t(x,0) = 0, \; x > 0;$$
$$u(0,t) = f(t) \;\; (t > 0), \;\; \text{where supp } f = [0,1]$$
$$[\text{for example, } f(t) = \sin\omega t \text{ for } 0 \le t \le 1, \; f(t) = 0 \text{ for all other } t];$$
$$|u(x,t)| \le M \;\text{ for all }\; x, \, t.$$

Over which time interval does one receive a signal at the point x_0? [For which values of t is $u(x_0,t) \neq 0$?]

In a *variation on the Laplace method* for ordinary [or partial] differential equations, one sets $y(t)$ [or $u(x,t)$] equal to $\int_\Gamma g(s)e^{ts}ds$, where Γ is a path in the complex s-plane, with end-points a and b, say, that is to be chosen later.

12.4.13. Apply this form of the Laplace method to Bessel's equation of order zero.

Hint. The differential equation leads to the following condition on g:

$$\int_\Gamma (s^2+1)g(s)d_s e^{ts} + \int_\Gamma sg(s)e^{ts}ds = 0, \;\; \forall\, t \in \mathbb{R}^+ \text{ or } \mathbb{R}.$$

Integration by parts transforms the condition to

$$\int_\Gamma \left[-\{(s^2+1)g(s)\}' + sg(s)\right] e^{ts}ds = 0, \;\; \forall\, t,$$

provided the integrated term $[(s^2+1)\,g(s)e^{ts}]_a^b$ is equal to zero.

12.4.14. Apply the same method to Bessel's equation of order ν [Exercise 8.1.5], after it has been reduced to the form

$$tz'' + (2\nu+1)z' + tz = 0$$

by the substitution $y(t) = t^\nu z(t)$. [The solution of the z-equation for which $z(0) = 1/\{2^\nu\Gamma(\nu+1)\}$ is $J_\nu(t)/t^\nu$; cf. Proposition 11.7.4.]

12.5. Fourier cosine and sine transformation

For problems involving a half-line it is sometimes convenient to use integral analogs of cosine and sine series instead of Laplace transformation.

DEFINITION 12.5.1. For integrable functions on $\mathbb{R}^+ = (0, \infty)$, the (Fourier) cosine transform $g = Cf$ and the (Fourier) sine transform $h = Sf$ are given by the formulas

$$
\begin{aligned}
g(\xi) = (Cf)(\xi) &\stackrel{\text{def}}{=} \int_0^\infty f(x) \cos \xi x \, dx, \quad \xi \in \mathbb{R}^+ \text{ or } \xi \in \mathbb{R}, \\
h(\xi) = (Sf)(\xi) &\stackrel{\text{def}}{=} \int_0^\infty f(x) \sin \xi x \, dx, \quad \xi \in \mathbb{R}^+ \text{ or } \xi \in \mathbb{R}.
\end{aligned}
$$

As a function on $\mathbb{R}$, g is even and h is odd.

Cosine and sine transform are closely related to Fourier transforms. Indeed, let f_e be the even, f_o the odd extension of f to $\mathbb{R}$. Then

$$
\begin{aligned}
g = Cf &= \frac{1}{2} \int_{\mathbb{R}} f_e(x) \cos \xi x \, dx \\
&= \frac{1}{2} \int_{\mathbb{R}} f_e(x) e^{-i\xi x} dx = \frac{1}{2} \mathcal{F} f_e = \frac{1}{2} \mathcal{F}_R f_e,
\end{aligned} \tag{12.5.1}
$$

$$
\begin{aligned}
h = Sf &= \frac{1}{2} \int_{\mathbb{R}} f_o(x) \sin \xi x \, dx \\
&= \frac{i}{2} \int_{\mathbb{R}} f_o(x) e^{-i\xi x} dx = \frac{i}{2} \mathcal{F} f_o = \frac{-i}{2} \mathcal{F}_R f_o.
\end{aligned} \tag{12.5.2}
$$

Hence by Fourier inversion, and appropriate interpretation of the formulas, cf. Theorems 9.2.2 and 10.1.7,

$$
\begin{aligned}
f_e &= \frac{1}{2\pi} \mathcal{F}_R \, 2g = \frac{1}{\pi} \mathcal{F} g = \frac{2}{\pi} Cg, \\
f_o &= \frac{1}{2\pi} \mathcal{F}_R \, \frac{2}{i} \, h = \frac{i}{\pi} \mathcal{F} h = \frac{2}{\pi} Sh.
\end{aligned}
$$

Restriction to $\mathbb{R}^+$ thus gives

THEOREM 12.5.2. (Inversion theorem)*:*

$$
\textit{if } g = CF \textit{ then } f = \frac{2}{\pi} Cg, \quad \textit{if } h = Sf \textit{ then } f = \frac{2}{\pi} Sh.
$$

These formulas have to be interpreted in the proper way. For example, if f is in $L^1(\mathbb{R}^+)$ and differentiable at the point $x > 0$, then

$$f(x) = \frac{2}{\pi}\,(Cg)(x) = \lim_{A\to\infty} \frac{2}{\pi} \int_0^A g(\xi) \cos x\xi \, d\xi.$$

For arbitrary even or odd tempered distributions (on $\mathbb{R}$) one may define the cosine and sine transform in terms of the Fourier transform as indicated in (12.5.1), (12.5.2). For locally integrable functions f on $(0,\infty)$ of at most polynomial growth one thus has

$$(Cf)(\xi) = \mathcal{S}' \lim_{A\to\infty} \int_0^A f(x) \cos \xi x \, dx, \quad \xi \in \mathbb{R}, \ \text{etc.} \tag{12.5.3}$$

Examples 12.5.3. Earlier examples of Fourier transforms readily give most of the following cosine and sine transforms (where $a > 0$):

$f(x)$	$Cf(\xi)$	$Sf(\xi)$
$e^{-a\lvert x\rvert}$	$\dfrac{a}{\xi^2 + a^2}$	$\dfrac{\xi}{\xi^2 + a^2}$
e^{-ax^2}	$\dfrac{1}{2}\sqrt{\dfrac{\pi}{a}}\, e^{-\xi^2/(4a)}$	
$\begin{cases} 1 & \text{for } x < a \\ 0 & \text{for } x > a \end{cases}$	$\dfrac{\sin a\xi}{\xi}$	$\dfrac{1 - \cos a\xi}{\xi}$
1	$\pi\delta(\xi)$	$\text{pv}\,\dfrac{1}{\xi}$
$J_0(x)$	$\begin{cases} \dfrac{1}{\sqrt{1-\xi^2}} & \text{for } \xi < 1 \\ 0 & \text{for } \xi > 1 \end{cases}$	$\begin{cases} 0 & \text{for } \xi < 1 \\ \dfrac{1}{\sqrt{\xi^2 - 1}} & \text{for } \xi > 1 \end{cases}$

For the transforms of the Bessel function $J_0(x)$, see Exercises 11.2.8, 11.2.9. As an alternative one may start with the Laplace transform of $J_0(t)U(t)$; cf. Example 12.2.2.

Rules for cosine and sine transformation. The following rules hold under appropriate conditions, and then the proofs are straightforward; λ denotes a positive constant.

	$f(x)$	$Cf(\xi) = g(\xi)$	$Sf(\xi) = h(\xi)$
(i)	$f(\lambda x)$	$\dfrac{1}{\lambda}\, g\left(\dfrac{\xi}{\lambda}\right)$	$\dfrac{1}{\lambda}\, h\left(\dfrac{\xi}{\lambda}\right)$
(ii)	Df	$\xi Sf - f(0)$	$-\xi Cf$
(iii)	xf	DSf	$-DCf$
(iv)	$D^2 f$	$-\xi^2 Cf - f'(0+)$	$-\xi^2 Sf + f(0)\xi$
(v)	$x^2 f$	$-D^2 Cf$	$-D^2 Sf$

Rule (ii) holds in the classical sense whenever f is an indefinite integral on $\mathbb{R}^+$ with f and f' in $L^1(\mathbb{R}^+)$. It holds in extended sense if f is equal to an indefinite integral on $\mathbb{R}^+$ with f' of at most polynomial growth [so that f is polynomially bounded]; cf. formula (12.5.3).

Observe that only transforms of even order derivatives are expressed in terms of the same transform.

Exercises. 12.5.1. Compute $C\left[e^{-ax^2}\right]$ and $S\left[xe^{-ax^2}\right]$, paying special attention to the case $a = 1/2$.

12.5.2. Prove the rules for SDf, Sxf and $SD^2 f$ under appropriate conditions on f.

12.5.3. Show that for f in $L^2(\mathbb{R}^+)$,

$$\int_{\mathbb{R}^+} |Cf|^2 = \int_{\mathbb{R}^+} |Sf|^2 = \frac{1}{2}\pi \int_{\mathbb{R}^+} |f|^2.$$

12.5.4. Prove that $C^2 = S^2 = (\pi/2) \cdot \text{identity}$ on $L^2(\mathbb{R}^+)$. What are the eigenvalues of C and S? Indicate corresponding eigenfunctions.

12.5.5. Determine the transform $S\left[x^{-\frac{1}{2}}\right]$ after observing that it must have the form $c\xi^{-\frac{1}{2}}$ with $c > 0$. Similarly for $C\left[x^{-\frac{1}{2}}\right]$.

12.5.6. Determine $S\left[x^{-1}\right]$, where x^{-1} is interpreted as the odd distribution $\mathrm{pv}\,(1/x)$. Also determine $C[(\sin \lambda x)/x]$ where $\lambda > 0$.

12.5.7. Obtain the solution of the following boundary value problem in the form of a sine transform:

$$\begin{aligned}
&u_{xx} + u_{yy} = 0, \quad 0 < x < 1, \;\; y > 0;\\
&u(0, y) = 0, \quad u(1, y) = f(y), \;\; y > 0;\\
&u(x, 0) = 0, \quad 0 < x < 1; \;\; u(x, y) \;\text{ bounded.}
\end{aligned}$$

12.5.8. (*Heat conduction in semi-infinite medium*; cf. Exercise 12.4.11) Use sine transformation to solve the boundary value problem

$$\begin{aligned} &u_{xx} = u_t, \quad x > 0, \ \ t > 0;\\ &u(x,0) = 0, \quad x > 0; \quad u(x,t) \ \text{bounded};\\ &u(0,t) = f(t), \quad t > 0. \end{aligned}$$

[The solution may be written in the final form

$$\begin{aligned} u(x,t) &= \frac{1}{2\sqrt{\pi}} \int_0^t f(t-\tau) x \tau^{-3/2} e^{-x^2/(4\tau)} d\tau \\ &= \frac{2}{\sqrt{\pi}} \int_{x/(2\sqrt{t})}^{\infty} f\left(t - \frac{x^2}{4w^2}\right) e^{-w^2} dw. \,] \end{aligned}$$

What happens to the temperature $u(x,t)$ as $t \to \infty$ in the special case $u(0,t) = f(t) = 1$?

12.5.9. How would one solve the following boundary value problem:

$$\begin{aligned} &u_{xx} + u_{yy} = 0, \quad 0 < x < 1, \ \ y > 0;\\ &u(0,y) = u(1,y) = 0, \ \ y > 0;\\ &u(x,0) = f(x), \quad 0 < x < 1; \quad u(x,y) \ \text{bounded}. \end{aligned}$$

Determine the solution explicitly in the special case $u(x,0) = f(x) = 1$. Show that $u(x,y)$ tends to zero exponentially as $y \to \infty$.

12.6. The wave equation in $\mathbb{R}^n$

The emphasis will be on the cases $n = 1, 2, 3$, and for those it is not really necessary to use the general Inversion Theorem 11.7.5 for spherically symmetric functions. However, the case of arbitrary n is interesting because it brings out the difference between odd and even dimensions. We therefore begin by determining the fundamental solution $E = E(x,t)$ for the wave operator in arbitrary $\mathbb{R}^n$. It satisfies the equation

$$\begin{aligned} \Box\, E = \Box_n\, E &\overset{\text{def}}{=} \left(-\Delta_n^x + \frac{1}{c^2}\, D_t^2\right) E(x,t) = \delta(x,t) \\ &= \delta_n(x)\delta_1(t) \quad \text{on } \ \mathbb{R}^n \times \mathbb{R}. \end{aligned} \tag{12.6.1}$$

[The wave operator $\Box$, pronounced 'box', is also called the d'Alembertian (after d'Alembert); cf. [**4**].] The physical question is as follows. For displacements or disturbances governed by the wave equation, one wishes to determine the displacement $E(x,t)$ at the point x and time t, due to an "impulsive force" or "thrust" 1 at the point $x=0$ and time $t=0$.

Carrying out Fourier transformation relative to x: $\mathcal{F}^x E(x,t) = \widehat{E}(\xi,t)$, one obtains the equation

$$\left(\frac{1}{c^2}\,D_t^2 + \rho^2\right)\widehat{E}(\xi,t) = \delta_1(t), \quad \rho = |\xi|. \tag{12.6.2}$$

It is reasonable to look for a solution which vanishes for $t<0$. [When $t<0$, "nothing has happened yet".] Problem (12.6.2) may then be solved by Laplace transformation, provided one thinks of $\delta_1(t)$ as a limit of functions on $\mathbb{R}^+$. Let us take

$$\delta_1(t) = \lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\,\chi_{(0,\varepsilon)}(t),$$

where χ_J denotes the characteristic function of the interval J. Thus

$$(\mathcal{L}\delta_1)(t) = \lim\int_0^\varepsilon \frac{1}{\varepsilon}\,e^{-st}dt = \lim\frac{1-e^{-\varepsilon s}}{\varepsilon s} = 1.$$

One now finds $\mathcal{L}^t\widehat{E} = c^2/(s^2+c^2\rho^2)$, so that

$$\widehat{E}(\xi,t) = c\,\frac{\sin c\rho t}{\rho}\,U(t), \quad \rho = |\xi|. \tag{12.6.3}$$

Note that the answer for $\widehat{E}$ is independent of the dimension!

We finally apply Fourier inversion, making use of Theorem 11.7.5 with the roles of x and ξ interchanged. The result is

THEOREM 12.6.1. *The wave operator in $\mathbb{R}^n$ has the following fundamental solution which vanishes for $t<0$:*

$$\begin{aligned} E(x,t) &= (2\pi)^{-n}c\,\mathcal{F}_R^\xi\left[\frac{\sin ct\rho}{\rho}\right](x)\,U(t) \\ &= (2\pi)^{-n/2}cU(t)\;\mathcal{S}'\lim_{A\to\infty} r^{1-n/2}\int_0^A(\sin ct\rho)\rho^{(n/2)-1}J_{(n/2)-1}(r\rho)d\rho, \end{aligned} \tag{12.6.4}$$

where $\rho=|\xi|$ and $r=|x|$.

We will look closely at the cases $n=1,2,3$.

The case $n=1$. Here we may replace ρ by ξ because the resulting function is even in ξ:

$$\widehat{E}(\xi,t) = c\,\frac{\sin ct\xi}{\xi}\,U(t).$$

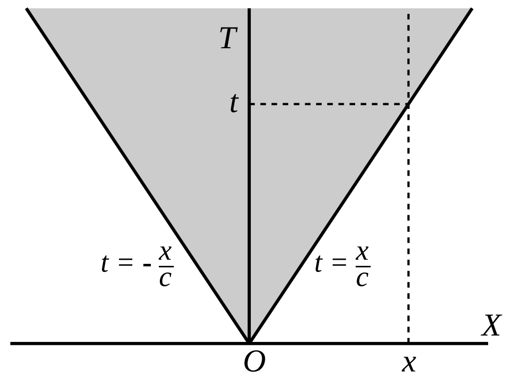

FIGURE 12.4

Inversion will give, cf. Exercise 9.1.2,

$$E(x,t) = \begin{cases} 0 & \text{for } t < |x|/c, \\ c/2 & \text{for } t > |x|/c. \end{cases}$$

Thus $E(x,t)$ is constant, equal to $c/2$, throughout the "forward light cone" $\{|x| < ct,\, t > 0\}$ with vertex at the point $(0,0)$; cf. Figure 12.4. At a given point $x \neq 0$, a disturbance arrives at time $t = |x|/c$; the displacement remains constant forever after. A succession of impulsive forces at the origin leads to a superposition of displacements at the point x.

What will be observed at the point $x > 0$ if the signal at the origin is a vibration or "tone" of short duration with circular frequency ω? For example, we might consider the signal

$$\Phi(x,t) = \delta(x)(\sin \omega t)\chi_{(0,\varepsilon)}(t) = \text{“}\delta(x) \int_0^\varepsilon (\sin \omega\tau)\delta_1(t-\tau)d\tau\text{”}. \tag{12.6.5}$$

The equation $\Box\, u = \Phi(x,t)$, with the condition $u = 0$ for $t < 0$, will [for $n = 1$] have the solution

$$\begin{aligned} u(x,t) &= \int_0^\varepsilon (\sin \omega\tau)E(x, t-\tau)d\tau \\ &= \begin{cases} 0 & \text{for } t < x/c, \\ \int_0^{t-x/c} (c/2)\sin \omega\tau\, d\tau & \text{for } (x/c) < t < (x/c) + \varepsilon, \\ (c/2)\,\dfrac{1-\cos \omega\varepsilon}{\omega}, \text{ a constant,} & \text{for } t > (x/c) + \varepsilon. \end{cases} \end{aligned}$$

Here the answer on the time-interval $(x/c) < t < (x/c) + \varepsilon$ works out to

$$(c/2)\,\frac{1 - \cos\omega(t - x/c)}{\omega}.$$

One will be able to recognize the frequency ω in this 'middle' time-interval, provided its length ε is a good deal larger than the period $2\pi/\omega$ of the vibration.

The case $n = 2$. We now find, either from Theorems 12.6.1 and 11.7.5, or by referring to Theorem 11.6.1, that

$$\begin{aligned} E(x,t) &= \frac{1}{2\pi}\,c\,U(t)\;\mathcal{S}'\lim_{A\to\infty}\int_0^A (\sin ct\rho)J_0(r\rho)d\rho \\ &= \frac{1}{2\pi}\,c\,U(t)\;\mathcal{S}'\lim_{\varepsilon\searrow 0}\int_0^\infty e^{-\varepsilon\rho}(\sin ct\rho)J_0(r\rho)d\rho. \end{aligned}$$

Thus we need the sine transform of J_0. It may be looked up under Examples 12.5.3, or one may use the second formula above in conjunction with the Laplace transform of J_0. The result is

$$E(x,t) = \begin{cases} 0 & \text{for } t < r/c, \\ \dfrac{1}{2\pi}\,\dfrac{c}{(c^2t^2 - r^2)^{1/2}} & \text{for } t > r/c. \end{cases}$$

The support of $E(x,t)$ is again the (closed) solid forward light cone; cf. Figure 12.5. At a fixed point x different from the origin, a disturbance arrives at time $t = r/c$. Afterwards, the displacement tends to zero, but this happens rather slowly. A time-limited signal emanating from the origin is not received as such at the point x.

The case $n = 3$. Either from Theorem 12.6.1 together with the form of $J_\nu(t)$ for $\nu = 1/2$ [formula (11.7.7)], or by using Exercise 11.4.6 with x and ξ interchanged, one obtains

$$\begin{aligned} E(x,t) &= \frac{c\,U(t)}{2\pi^2}\;\mathcal{S}'\lim_{A\to\infty}\frac{1}{r}\int_0^A (\sin ct\rho)\sin r\rho\;d\rho \\ &= \frac{c\,U(t)}{4\pi^2}\;\mathcal{S}'\lim_{A\to\infty}\frac{1}{r}\left\{\frac{\sin A(ct - r)}{ct - r} - \frac{\sin A(ct + r)}{ct + r}\right\} \\ &= \frac{c\,U(t)}{4\pi r}\big\{\delta_1(ct - r) - \delta_1(ct + r)\big\}; \end{aligned}$$

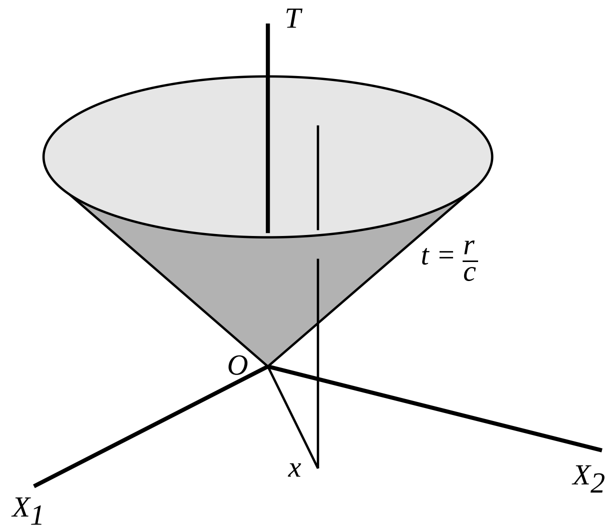

FIGURE 12.5

cf. Examples 10.4.6 (ii). Now for $t > 0$ one has $\delta_1(ct + r) = 0$. Also using the fact that $\delta_1(\lambda y) = (1/\lambda)\delta_1(y)$ for $\lambda > 0$, we obtain the simple answer

$$E(x,t) = \frac{1}{4\pi r}\,\delta_1\left(t - \frac{r}{c}\right). \tag{12.6.6}$$

This time the support of $E(x,t)$ is just the *boundary* of the forward light cone, the set $\{(x,t) \in \mathbb{R}^3 \times \mathbb{R} : r = ct,\, t \geq 0\}$. At a given point x different from the origin, a sharply time-limited signal is received at the instant $t = r/c$. This signal has precisely the same shape as the original one at the origin at time $t = 0$! A time-limited signal of the form (12.6.5), emanating from the origin, will be received at the point x as

$$u(x,t) = \text{“}\int_0^{\varepsilon} (\sin\omega\tau)E(x, t-\tau)d\tau\text{”} = \frac{1}{4\pi r}\,\sin\omega(t - r/c),$$

$(r/c) < t < (r/c) + \varepsilon$. Observe that *the signal is received without distortion!* There is only attenuation because of the distance: the amplitude at the point x is inversely proportional to the distance r from the origin.

REMARKS 12.6.2. In dimensions $n \geq 4$, the disturbance $E(x,t) = E_n(x,t)$ also reaches the point x at time $t = r/c$. When $n = 4, 6, \cdots$, the resulting displacement $E(x,t)$ tends to zero relatively slowly as $t \to \infty$. When $n = 5, 7, \cdots$, the displacement $E(x,t)$ is sharply limited in time, but there is a great deal of distortion from the original. Indeed, the fundamental solution will now contain derivatives of $\delta_1(t-r/c)$!

One has the symbolic relation

$$E_{n+2}(x,t) = -\frac{1}{2\pi r}\,\frac{\partial}{\partial r}\,E_n(x,t); \tag{12.6.7}$$

cf. Exercise 12.6.3. Here $E(x,t)$ is considered as a function $\tilde{E}(r,t)$ of $r=|x|$.

Exercises. 12.6.1. Compute

$$\begin{aligned}&\mathcal{S}'\lim_{\varepsilon\searrow 0}\int_0^\infty e^{-\varepsilon\rho}(\sin ct\rho)J_0(r\rho)d\rho\\ &= \operatorname{Im}\mathcal{S}'\lim\int_0^\infty e^{-(\varepsilon-ict)\rho}J_0(r\rho)d\rho.\end{aligned}$$

12.6.2. Compute

$$\mathcal{S}'\lim_{\varepsilon\searrow 0}\frac{1}{r}\int_0^\infty e^{-\varepsilon\rho}(\sin ct\rho)\sin r\rho\,d\rho.$$

12.6.3. Use the recurrence relation

$$\frac{1}{z}\,\frac{d}{dz}\left\{z^{-\nu}J_\nu(z)\right\} = -z^{-\nu-1}J_{\nu+1}(z)$$

of Exercise 11.7.3 to verify the important recursion formula (12.6.7) for the fundamental solution of the wave equation in different dimensions.

CHAPTER 13

General distributions and Laplace transforms

Tempered distributions on $\mathbb{R}^n$ correspond to functions of at most polynomial growth. However, in practice one also encounters functions of much more rapid growth at infinity. In order to embed such functions in a system of distributions, it is necessary to restrict the test functions ϕ to $\mathcal{C}^\infty$ functions which vanish outside some bounded set $K = K_\phi$.

For a subclass of the corresponding general distributions one can introduce two-sided Laplace transformation.

13.1. General distributions on $\mathbb{R}$ and $\mathbb{R}^n$

We begin by defining suitable test functions.

Definitions 13.1.1. The test class $\mathcal{C}_0^\infty$ on $\mathbb{R}^n$ consists of the $\mathcal{C}^\infty$ functions ϕ that have compact support.

This class is made into the Schwartz *space* $\mathcal{D}$ of *test functions* by the following definition of convergence for sequences. One says that

$$\phi_j \to \phi \ \text{ in } \ \mathcal{D} \quad \text{if}$$

(i) the supports of the functions ϕ_j and ϕ belong to a fixed compact set K, and
(ii)

$$D^\alpha \phi_j = \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}\,\phi_j \to D^\alpha \phi = \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}\,\phi$$

uniformly on K (and hence uniformly on $\mathbb{R}^n$) for every multi-index $\alpha = (\alpha_1, \cdots, \alpha_n) \geq 0$.

There is a corresponding definition of convergence $T_\lambda \to T$ for distributions T_λ depending on a real parameter λ tending to λ_0.

Examples 13.1.2. An important test function on $\mathbb{R}^n$ is

$$\theta(x) \stackrel{\text{def}}{=} \begin{cases} \dfrac{1}{c}\exp\left(-\dfrac{1}{1-|x|^2}\right) & \text{for } x \in \mathbb{R}^n \text{ with } |x| < 1, \\ 0 & \text{for } |x| \geq 1, \end{cases} \tag{13.1.1}$$

where

$$c = \int_{B(0,1)} \exp\left(-\frac{1}{1-|y|^2}\right) dy.$$

Cf. Examples 4.1.9 for the one-dimensional case [with different notations].

A very useful family of related test functions is given by

$$\theta_\varepsilon(x) = \frac{1}{\varepsilon^n}\,\theta\left(\frac{x}{\varepsilon}\right), \quad x \in \mathbb{R}^n, \;\; \varepsilon > 0. \tag{13.1.2}$$

For these functions

$$\int_{\mathbb{R}^n} \theta_\varepsilon(x)dx = \int_{\mathbb{R}^n} \theta(y)dy = 1, \quad \operatorname{supp}\theta_\varepsilon(x) = \overline{B}(0,\varepsilon).$$

For arbitrary compact sets $K \in \mathbb{R}^n$ there are test functions ω_ε that are equal to 1 on K and equal to 0 outside $K_{2\varepsilon}$, the 2ε-neighborhood of K. One may obtain such a function by setting

$$\begin{aligned} \omega_\varepsilon(x) = \big\{\chi(K_\varepsilon) * \theta_\varepsilon\big\}(x) &= \int_{\mathbb{R}^n} \chi_{K_\varepsilon}(y)\theta_\varepsilon(x-y)dy \\ &= \int_{K_\varepsilon} \theta_\varepsilon(x-y)dy. \end{aligned} \tag{13.1.3}$$

Definitions 13.1.3. A *distribution* T on $\mathbb{R}^n$ is a continuous linear functional on the test space $\mathcal{D}$: whenever $\phi_\lambda \to \phi$ in $\mathcal{D}$, it is required that the numbers $< T, \phi_\lambda >$ tend to $< T, \phi >$.

The class of distributions is made into the distribution *space* $\mathcal{D}'$ by the following definition of (weak) convergence when $\lambda \to \lambda_0$:

$$T_\lambda \to T \quad \text{if} \quad < T_\lambda, \phi > \to < T, \phi >, \;\; \forall\, \phi \in \mathcal{D}. \tag{13.1.4}$$

One could also define the space $\mathcal{D}'$ by *completion* of the space of locally integrable functions, provided with the definition of convergence relative to test functions corresponding to (13.1.4); cf. [**68**].

Every locally integrable function f on $\mathbb{R}^n$ defines a distribution T_f by the formula

$$< T_f, \phi > = \int_{\mathbb{R}^n} f\phi, \quad \forall\, \phi \in \mathcal{D}.$$

These special distributions are in $1-1$ correspondence with the defining functions f:

$$T_f = 0 \;\text{ if and only if }\; f = 0,$$

provided we identify functions that are equal almost everywhere. One identifies T_f with f and writes $< T_f, \phi > = < f, \phi >$.

The delta distribution on $\mathbb{R}^n$ has its usual definition $< \delta, \phi > = \phi(0)$, $\forall \phi \in \mathcal{D}$. The tempered distributions on $\mathbb{R}^n$ are distributions in the present sense; cf. Exercise 13.1.2.

One says that $T_1 = T_2$ on an open set $\Omega \subset \mathbb{R}^n$ if $< T_1, \phi > = < T_2, \phi >$ for all test functions ϕ with support in Ω. The support of T is defined as usual; cf. Definition 10.4.3. All distributions T may be multiplied by arbitrary $\mathcal{C}^\infty$ functions ω:

$$< \omega T, \phi > = < T\omega, \phi > \stackrel{\text{def}}{=} < T, \omega\phi >, \quad \forall \phi \in \mathcal{D}.$$

Definition 13.1.4. (*Derivatives*) For a distribution T on $\mathbb{R}^n$, the (partial) derivative $D_k T = \partial T / \partial x_k$ is the distribution on $\mathbb{R}^n$ given by

$$< D_k T, \phi > = -\left\langle T, \frac{\partial \phi}{\partial x_k} \right\rangle, \quad \forall \phi \in \mathcal{D}.$$

In the case $n = 1$ one simply writes DT. In $\mathbb{R}^n$ one uses the notation

$$D^\alpha T = D_1^{\alpha_1} \cdots D_n^{\alpha_n} T = \frac{\partial^{\alpha_1 + \cdots + \alpha_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} T,$$

where α stands for the multi-index $(\alpha_1, \cdots, \alpha_n) \geq 0$. Thus

$$< D^\alpha T, \phi > = (-1)^{\alpha_1 + \cdots + \alpha_n} < T, D^\alpha \phi > .$$

Since the order of differentiation is immaterial for test functions, the same is true for distributional derivatives D^α. Furthermore, one readily proves

Proposition 13.1.5. *Distributional differentiation is continuous: if distributions T_λ converge to T in $\mathcal{D}'$, then $D^\alpha T_\lambda$ converges to $D^\alpha T$ in $\mathcal{D}'$.*

Examples 13.1.6. For $n = 1$ one has $\delta = DU$, where $U = 1_+$ is the unit step function. If g is an indefinite integral on $\mathbb{R}$ then $Dg = g'$; cf. Section 10.5. If a function f on $\mathbb{R}$ is equal to an indefinite integral both on $(-\infty, 0)$ and on $(0, \infty)$, while f' is integrable over all finite intervals $(-A, 0)$ and $(0, A)$, then

$$Df = f' + s\delta,$$

where s is the jump $f(0+) - f(0-)$. [Indeed, $f - sU$ will be equal to an indefinite integral g on $\mathbb{R}$.]

The Laplacian ΔT of a distribution T on $\mathbb{R}^n$ is given by the formula

$$< \Delta T, \phi > = < T, \Delta \phi >, \quad \forall \phi \in \mathcal{D}.$$

EXAMPLE 13.1.7. (*Distributional Laplacian of* $u(x) = 1/r = 1/|x|$ *in* $\mathbb{R}^3$) The function $u(x) = 1/r$ is locally integrable on $\mathbb{R}^3$ and it satisfies Laplace's equation in the classical sense on $\mathbb{R}^3 \setminus \{0\}$, but Δu does not vanish throughout $\mathbb{R}^3$!

Since $1/r = \lim \{1/(r+\varepsilon)\}$ in $\mathcal{D}'$ as $\varepsilon \searrow 0$, one will have $\Delta(1/r) = \lim \Delta\{1/(r+\varepsilon)\}$. However, it is easier to compute $\Delta(1/r)$ from the relation

$$\frac{1}{r} = \lim u_\varepsilon(x) \quad \text{with} \quad u_\varepsilon(x) = (r^2+\varepsilon)^{-\frac{1}{2}}.$$

From the form of the Laplacian in polar coordinates one finds

$$\begin{aligned}\Delta u_\varepsilon(x) &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial u_\varepsilon}{\partial r}\right) = \frac{1}{r^2}\frac{d}{dr}\left\{-r^3(r^2+\varepsilon)^{-\frac{3}{2}}\right\}\\ &= -3\varepsilon\left(r^2+\varepsilon\right)^{-\frac{5}{2}}.\end{aligned}$$

Thus for a test function ϕ, taking R so large that $\operatorname{supp}\phi$ belongs to the ball $B(0,R)$,

$$\begin{aligned}< \Delta u_\varepsilon, \phi > &= \int_{B(0,R)} \Delta u_\varepsilon(x)\phi(x)dx = \int_{B(0,R)} \Delta u_\varepsilon(x)\phi(0)dx\\ &\quad - 3\varepsilon \int_{B(0,R)} \{\phi(x)-\phi(0)\}(r^2+\varepsilon)^{-\frac{5}{2}}\,dx. \qquad (13.1.5)\end{aligned}$$

We first compute the integral of $\Delta u_\varepsilon(x)$ itself:

$$\begin{aligned}\int_{B(0,R)} \Delta u_\varepsilon(x)dx &= \int_0^R \frac{1}{r^2}\frac{d}{dr}\left\{-r^3(r^2+\varepsilon)^{-\frac{3}{2}}\right\}\cdot 4\pi r^2 dr\\ &= -4\pi\left[r^3(r^2+\varepsilon)^{-\frac{3}{2}}\right]_0^R = -4\pi + \mathcal{O}(\varepsilon).\end{aligned}$$

For the test function ϕ one has $\phi(x)-\phi(0) = \mathcal{O}(r)$. With this inequality, the final term in (13.1.5) can be estimated as $\mathcal{O}(\varepsilon)$. Combining results, one finds that

$$\left\langle \Delta\frac{1}{r}, \phi\right\rangle = \lim_{\varepsilon\searrow 0} < \Delta u_\varepsilon, \phi > = -4\pi\phi(0) = < -4\pi\delta, \phi >$$

for all ϕ. Hence

$$\Delta\frac{1}{r} = \Delta\frac{1}{|x|} = -4\pi\delta(x) \quad \text{in } \mathbb{R}^3. \qquad (13.1.6)$$

We end with a fundamental result on the structure of general distributions.

THEOREM 13.1.8. *When restricted to a bounded open set Ω, a distribution T on $\mathbb{R}^n$ is equal to a distributional derivative D^α of some order $\alpha = (\alpha_1, \cdots, \alpha_n)$ of a locally integrable function.*

We sketch a proof for $n = 1$. Let ω be a test function on $\mathbb{R}$ which is equal to 1 on $(-A, A)$. Then the distribution ωT has compact support, and hence may be considered as a tempered distribution. Indeed, the formula

$$< \omega T, \psi > = < T, \omega\psi >, \quad \psi \in \mathcal{S},$$

defines ωT as a continuous linear functional on $\mathcal{S}$. Thus by the Structure Theorem 10.6.2 for tempered distributions, one has $\omega T = D^s f = D^s f_A$ with $f_A \in \mathcal{P}$. It follows that

$$T = \omega T = D^s f_A \quad \text{on} \;\; (-A, A).$$

Globally a distribution on $\mathbb{R}^n$ need not be a derivative of a locally integrable function; cf. Exercise 13.1.6.

Exercises. 13.1.1. Verify that formula (13.1.3) defines a test function ω_ε on $\mathbb{R}^n$ which is equal to 1 on K and equal to 0 outside $K_{2\varepsilon}$.

13.1.2. Show that convergence $\phi_\lambda \to \phi$ in $\mathcal{D}$ implies convergence $\phi_\lambda \to \phi$ in $\mathcal{S}$ [Section 10.4]. Deduce that every tempered distribution on $\mathbb{R}^n$ is equal to a distribution in $\mathcal{D}'(\mathbb{R}^n)$.

13.1.3. Prove that distributional differentiation is continuous.

13.1.4. Use the approach of Example 13.1.7 to derive that in $\mathbb{R}^n$ (with $n \neq 2$)

$$\Delta \frac{1}{|x|^{n-2}} = -(n-2)\sigma_n \delta(x), \quad \text{where} \;\; \sigma_n = \text{area of}\, S_1 \text{ in } \mathbb{R}^n.$$

13.1.5. Show that in $\mathbb{R}^2$, one has $\delta(x) = \delta(x_1, x_2) = D_1 D_2\{U(x_1)U(x_2)\}$.

13.1.6. Verify that the series

$$\delta(x) + D\delta(x-1) + \cdots + D^k\delta(x-k) + \cdots$$

converges to a distribution T on $\mathbb{R}$. Prove also that T cannot be represented in the form $D^m F$ on $\mathbb{R}$, with $m \geq 0$ and F locally integrable.

13.2. Two-sided Laplace transformation

Here we restrict ourselves to the case of one independent variable. For suitable functions f on $\mathbb{R}$ one may define

$$(13.2.1) \qquad g(s) = (\mathcal{L}f)(s) = (\mathcal{L}_{II}f)(s) \overset{\text{def}}{=} \int_{\mathbb{R}} f(t)e^{-st}dt, \quad s = \sigma + i\tau.$$

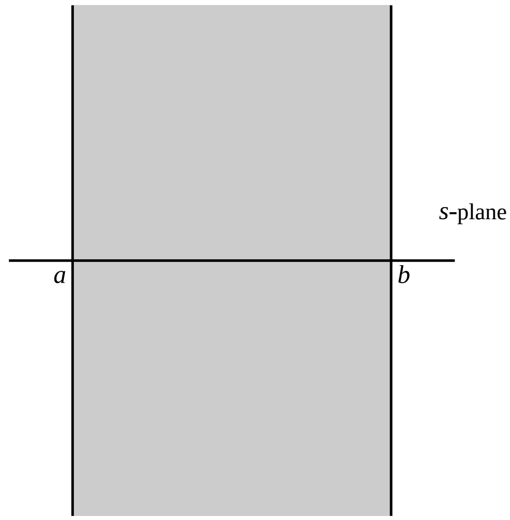

FIGURE 13.1

The two-sided transform is a continuous analog of a Laurent series

$$\sum_{n=-\infty}^{\infty} a_n z^n = \sum_{-\infty}^{\infty} a_n e^{-ns}.$$

The typical domain of convergence for such a series is an annulus

$$\rho < |z| = \left|e^{-s}\right| = e^{-\sigma} < R.$$

In terms of s this becomes a vertical strip [Figure 13.1]:

$$-\log R = a < \sigma = \operatorname{Re} s < b = -\log \rho.$$

The sum of the series is analytic throughout the strip.

PROPOSITION 13.2.1. *Let f be a function on $\mathbb{R}$ such that the product $f(t)e^{-\sigma t}$ is integrable over $\mathbb{R}$ for $a < \sigma < b$.* [An equivalent condition would be that $f(t)e^{-\sigma t}U(t)$ is in $L^1(\mathbb{R}^+)$ for all $\sigma > a$, while $f(t)e^{-\sigma t}U(-t)$ is in $L^1(\mathbb{R}^-)$ for all $\sigma < b$.] *Then the* (two-sided) *Laplace transform $g(s) = (\mathcal{L}f)(s)$ exists for all complex s in the strip $\{a < \sigma = \operatorname{Re} s < b\}$. The transform is an analytic function which is bounded on every 'interior strip' $\alpha \le \sigma \le \beta$* [that is, $a < \alpha < \beta < b$], *and one has*

$$g'(s) = -\int_{\mathbb{R}} t f(t) e^{-st} dt. \tag{13.2.2}$$

The proof follows readily from Theorem 12.1.2 since

$$g(s) = \left(\int_0^\infty + \int_{-\infty}^0 \right) f(t)e^{-st}dt = \int_0^\infty f(t)e^{-st}dt + \int_0^\infty f(-t)e^{st}dt. \tag{13.2.3}$$

The first integral on the right represents a bounded analytic function on every right half-plane $\operatorname{Re} s \geq \alpha > a$, and the final integral, a bounded analytic function on every left half-plane $\operatorname{Re} s \leq \beta < b$.

The two-sided Laplace transform is closely related to a Fourier transform:

$$g(\sigma + i\tau) = \int_{\mathbb{R}} f(t)e^{-\sigma t}e^{-i\tau t}dt = \mathcal{F}\big[f(t)e^{-\sigma t}\big](\tau). \tag{13.2.4}$$

Definition 13.2.2. (*Extended Laplace transformation*) Let T be a distribution on $\mathbb{R}$ such that the product $T(t)e^{-st}$ is a tempered distribution on $\mathbb{R}$ for $a < \sigma = \operatorname{Re} s < b$. Then the (two-sided) Laplace transform $\mathcal{L}f$ is given by

$$g(s) = (\mathcal{L}T)(s) = (\mathcal{L}T)(\sigma + i\tau) \stackrel{\text{def}}{=} \mathcal{F}\big[T(t)e^{-\sigma t}\big](\tau), \quad a < \sigma < b. \tag{13.2.5}$$

Thinking of an integral, one sometimes writes symbolically

$$g(s) = (\mathcal{L}T)(s) = < T(t), e^{-st} > . \tag{13.2.6}$$

For distributions T as in Definition 13.2.2, it is indeed possible to extend the class of test functions in such a way that it includes the functions e^{-st} for $a < \operatorname{Re} s < b$.

Proposition 13.2.3. *Under the conditions of Definition 13.2.2, the Laplace transform $g(s) = (\mathcal{L}T)(s)$ is analytic in the strip $a < \sigma = \operatorname{Re} s < b$. On 'interior strips', the transform is bounded by* (the absolute value of) *a polynomial in s. One has the complex inversion formula*

$$T(t) = \lim_{A \to \infty} \frac{1}{2\pi i} \int_{\sigma - iA}^{\sigma + iA} g(s)e^{ts}ds, \quad \forall\, \sigma \in (a, b). \tag{13.2.7}$$

Here the limit is in the sense of general distributions.

For a proof of the first two parts one would like to go back to one-sided transforms, but a decomposition as in (13.2.3) is not always possible. Indeed, the product $T(t)U(t)$ need not be well-defined. However, in terms of a $\mathcal{C}^\infty$ function $\omega(t)$ that is equal to 0 for $t \leq -1$ and equal to 1 for $t \geq +1$, one may write

$$g(s) = \mathcal{L}(\omega T)(s) + \mathcal{L}\{(1 - \omega)T\}(s).$$

	Original $f(t)$	Laplace trf $g(s)$	Remark
(i)	$f(\lambda t)$	$\dfrac{1}{\lvert\lambda\rvert}\, g\Big(\dfrac{s}{\lambda}\Big)$	$\lambda \in \mathbb{R} \setminus \{0\}$, $\sigma/\lambda \in (a, b)$
(ii)	$f(t+\lambda)$	$e^{\lambda s} g(s)$	$\lambda \in \mathbb{R}$
(iii)	$e^{\lambda t} f(t)$	$g(s-\lambda)$	$\lambda \in \mathbb{C}$, $\sigma - \operatorname{Re}\lambda \in (a, b)$
(iv)	$Df(t)$	$sg(s)$	
(v)	$tf(t)$	$-g'(s)$	
(vi)	$(f_1 * f_2)(t)$	$g_1(s) g_2(s)$	validity limited

This decomposition essentially gives $g(s)$ as the sum of two one-sided Laplace transforms. The proof of the first two parts of the theorem may now be completed with the aid of Exercise 13.2.6.

Fourier inversion of (13.2.5) as in Chapter 11 [see Theorems 11.1.2 and 11.2.4] finally shows that for $a < \sigma < b$,

$$T(t)e^{-\sigma t} = \frac{1}{2\pi}\, \mathcal{F}_R^{\tau}[g(\sigma + i\tau)](t) = \lim_{A\to\infty} \frac{1}{2\pi} \int_{-A}^{A} g(\sigma + i\tau) e^{it\tau} d\tau.$$

Here the limit is to be taken in the sense of tempered distributions. Multiplication by $e^{\sigma t}$ gives (13.2.7) in the sense of general distributions on $\mathbb{R}$.

EXAMPLES 13.2.4. By (13.2.5) or (13.2.6), $\mathcal{L}\delta = 1$. The function

$$g(s) = \frac{1}{s^2+1}, \quad \operatorname{Re} s > 0$$

is the Laplace transform of $(\sin t)U(t)$; cf. Examples 12.1.1 and 12.3.3. However, the function

$$\tilde{g}(s) = \frac{1}{s^2+1}, \quad \operatorname{Re} s < 0$$

is the Laplace transform of $-(\sin t)U(-t)$!

Rules for the (two-sided) Laplace transformation; see the table. Here it has been assumed that f is a function or distribution on $\mathbb{R}$ such that $f(t)e^{-\sigma t}$ is an integrable function on $\mathbb{R}$ for $a < \sigma < b$, or at least a tempered distribution.

Discussion. For well-behaved functions, rules (i)-(v) follow directly from the transformation rules for integrals. For distributions, the formal equation (13.2.6) is very suggestive. For example,

$$\mathcal{L}DT = < DT, e^{-st} > = - < T, D_t e^{-st} > = s < T, e^{-st} > = s\mathcal{L}T.$$

For genuine proofs one may appeal to (13.2.5) and the rules for Fourier transformation.

For the convolution rule (vi) we start with a classical case. Suppose that f_1 and f_2 are locally integrable functions on $\mathbb{R}$ with support on the half-line $\{t \geq c\}$, and such that the products $f_j(t)e^{-\sigma t}$ are integrable over $\mathbb{R}$ for $\sigma = \operatorname{Re} s > a$. Then by Fubini's theorem, the product $(f_1 * f_2)(t)e^{-\sigma t}$ will also be integrable over $\mathbb{R}$ for $\sigma > a$; cf. Proposition 12.2.1. Furthermore,

$$\begin{aligned}\mathcal{L}(f_1 * f_2)(s) &= \int_{\mathbb{R}} \left\{ \int_{\mathbb{R}} f_1(v) f_2(t-v) dv \right\} e^{-st} dt \\ &= \int_{\mathbb{R}} f_1(v) e^{-sv} dv \int_{\mathbb{R}} f_2(t-v) e^{-s(t-v)} dt = \mathcal{L}f_1(s)\mathcal{L}f_2(s),\end{aligned}$$

provided $\sigma > a$.

The rule is also valid for distributions f_j on $\mathbb{R}$ with support on the half-line $\{t \geq c\}$, and such that the products $f_j(t)e^{-\sigma t}$ are tempered for $\sigma > a$.

Exercises. 13.2.1. Prove that Laplace transformation according to Definition 13.2.2 is one to one.

13.2.2. Compute $\mathcal{L}[U(t)]$ and $\mathcal{L}[-U(-t)]$, and compare the answers. Explain why the result does not contradict the previous exercise.

13.2.3. Verify the results stated in Examples 13.2.4.

13.2.4. Prove rules (iv) and (v) in the table of Laplace transforms.

13.2.5. Deduce rule (iv) in the table of Section 12.2 for the one-sided Laplace transform of a derivative from rule (iv) in the new table.

Hint. Assuming that $f(t)$ can be written as an indefinite integral $c + \int_0^t f'(v)dv$ on $\mathbb{R}$, one has

$$D\{f(t)U(t)\} = f'(t)U(t) + f(0)\delta(t) \qquad \text{(cf. Examples 13.1.6).}$$

13.2.6. Let T be a distribution on $\mathbb{R}$ with support in $[c, +\infty]$ such that $T(t)e^{-\alpha t}$ is tempered. Prove:

(i) There exist a function $F \in \mathcal{P}$ with support in $[c, \infty]$ and an integer $m \geq 0$ such that $T = e^{\alpha t} D^m F$.

Hint. Part (iii) of Structure Theorem 10.6.2 gives a representation of the form $Te^{-\alpha t} = D^m f$. In the present case, $D^m f = 0$ on $(-\infty, c)$, hence on that interval, f is equal to a polynomial P of degree $< m$.

(ii) $(\mathcal{L}T)(s)$ is analytic for $\operatorname{Re} s > \alpha$, and on every half-plane $\{\operatorname{Re} s > \alpha'\}$ with $\alpha' > \alpha$, $|\mathcal{L}T(s)|$ is bounded by the absolute value of a polynomial in s.

Hint. By part (i) $\mathcal{L}T(s) = (s-\alpha)^m G(s-\alpha)$, where $G = \mathcal{L}F$.

13.2.7. Discuss the *Mellin transformation* [named after the Finnish mathematician Hjalmar Mellin (1854–1933; [**85**]), cf. [**86**]]:

$$g(s) \stackrel{\text{def}}{=} \int_0^\infty f(x) x^{s-1} dx, \tag{13.2.8}$$

under the hypothesis that $f(x)x^{\sigma-1}$ is a function that is integrable over $(0, \infty)$ for every $\sigma \in (a, b)$. Formulate a complex inversion formula.

Bibliography

[1] Abel, N.H., Internet, http://en.wikipedia.org/wiki/Niels_Henrik_Abel

[2] Abelian and Tauberian theorems, Internet, http://en.wikipedia.org/wiki/Abelian_and_tauberian_theorems

[3] d'Alembert, J.L.R., Internet, http://nl.wikipedia.org/wiki/Jean_Le_Rond_d27Alembert

[4] d'Alembert operator, Internet, en.wikipedia.org/wiki/D'Alembert_operator

[5] Banach, S., Internet, en.wikipedia.org/wiki/Stefan_Banach

[6] Bernoulli, D., Internet, http://en.wikipedia.org/wiki/Daniel_Bernoulli

[7] Bessel, F., Internet, http://en.wikipedia.org/wiki/Friedrich_Bessel

[8] du Bois-Reymond, P.D.G., Internet, http://en.wikipedia.org/wiki/Paul_David_Gustav_du_Bois-Reymond

[9] Borel, É., Internet, http://en.wikipedia.org/wiki/C389mile_Borel

[10] Borel measure, Internet, http://en.wikipedia.org/wiki/Borel_measure

[11] Carleson, L., Internet, http://en.wikipedia.org/wiki/Lennart_Carleson

[12] Cauchy, A.-L., Internet, en.wikipedia.org/wiki/Augustin-Louis_Cauchy

[13] Cauchy integral formula, Internet, en.wikipedia.org/wiki/Cauchy's_integral_formula

[14] Cauchy integral theorem, Internet, en.wikipedia.org/wiki/Cauchy's_integral_theorem

[15] Cauchy–Schwarz inequality, Internet, en.wikipedia.org/wiki/CauchySchwarz_inequality

[16] Cesàro, E., Internet, en.wikipedia.org/wiki/Ernesto_Cesro

[17] Cesàro summability, Internet, http://planetmath.org/encyclopedia/CesaroSummability.html

[18] Chebyshev, P., Internet, en.wikipedia.org/wiki/Pafnuty_Chebyshev

[19] Chebyshev polynomials, Internet, http://en.wikipedia.org/wiki/Chebyshev_polynomials

[20] Churchill, R.V. and Brown, J.W., Fourier series and boundary value problems. Fourth edition. McGraw-Hill, New York, 1987.

[21] Differential equations, Internet, en.wikipedia.org/wiki/Ordinary_differential_equation

[22] Dirac, P., Internet, http://en.wikipedia.org/wiki/Paul_Dirac

[23] Dirac, P.A.M., The principles of quantum mechanics. Oxford, at the Clarendon Press, 1930. Third ed. 1947.

[24] Dirac delta function, Internet, http://en.wikipedia.org/wiki/Dirac_delta_function
[25] Dirichlet, J.P.G.L., Internet, http://en.wikipedia.org/wiki/Johann_Peter_Gustav_Lejeune_Dirichlet
[26] Dirichlet kernel, Internet, http://en.wikipedia.org/wiki/Dirichlet_kernel
[27] Duistermaat, J.J. and Kolk, J.A.C., Distributions. Theory and Applications. Birkhäuser (Springer), New York etc, 2010.
[28] Dym, H. and McKean, H.P., Fourier series and integrals. Academic Press, New York, 1972.
[29] Euler, L., Internet, http://en.wikipedia.org/wiki/Leonhard_Euler
[30] Fejér, L., Internet, http://en.wikipedia.org/wiki/LipC3B3t_FejC3A9r
[31] Fischer, E.S., Internet, en.wikipedia.org/wiki/Ernst_Sigismund_Fischer
[32] Folland, G.B., Fourier analysis and its applications. Wadsworth & Brooks/Cole, Pacific Grove, CA, 1992.
[33] Fourier, J., Internet, http://nl.wikipedia.org/wiki/Joseph_Fourier
[34] Fubini, G., Internet, en.wikipedia.org/wiki/Guido_Fubini
[35] Gauss, C.F., Internet, http://nl.wikipedia.org/wiki/Carl_Friedrich_Gauss
[36] Gauss quadrature, Internet, http://en.wikipedia.org/wiki/Gaussian_quadrature
[37] Gelfand, I.M. and coauthors, Generalized functions, vol 1-5. (Translated from the Russian) Acad.Press, New York, 1964–1966.
[38] Gibbs, J.W., Internet, en.wikipedia.org/wiki/Josiah_Willard_Gibbs
[39] Gibbs phenomenon, Internet, http://en.wikipedia.org/wiki/Gibbs_phenomenon
[40] Gram, J.P., Internet, http://en.wikipedia.org/wiki/Jrgen_Pedersen_Gram
[41] Gram-Schmidt process, Internet, http://en.wikipedia.org/wiki/Gram-Schmidt_process
[42] Hadamard, J., Internet, http://en.wikipedia.org/wiki/Jacques_Hadamard
[43] Harmonic oscillator, Internet, en.wikipedia.org/wiki/Quantum_harmonic_oscillator
[44] Hardy, G.H., Internwet, en.wikipedia.org/wiki/G._H._Hardy
[45] Hardy, G.H. and Rogosinski, W.W., Fourier series. Cambridge Tracts no. 38, Cambridge Univ. Press, 1944. (Second ed. 1950.)
[46] Hermite, C., Internet, http://en.wikipedia.org/wiki/Charles_Hermite
[47] Hermite polynomials, http://en.wikipedia.org/wiki/Hermite_polynomials
[48] Hilbert, D., Internet, en.wikipedia.org/wiki/David_Hilbert
[49] Hilbert space, Internet, en.wikipedia.org/wiki/Hilbert_space
[50] Hilbert–Schmidt operator, Internet, http://en.wikipedia.org/wiki/Hilbert--Schmidt_operator
[51] Hölder, O., Internet, en.wikipedia.org/wiki/Otto_Hlder
[52] Hörmander, L., The analysis of linear partial differential operators, vol I-IV [I: Distribution theory and Fourier analysis.] Springer, Berlin, 1983–85.
[53] Hull, T.E., Internet, www.linkedin.com/pub/tom-hull/10/A68/355
[54] Ince, E.L., Ordinary Differential Equations. Dover Publications, 1958.
[55] Infeld, L., Internet, en.wikipedia.org/wiki/Leopold_Infeld

[56] Infeld, L. and Hull, T.E., The factorization method. Rev. Modern Physics **23** (1951), 21–68.

[57] Isoperimetric inequality, Internet, http://en.wikipedia.org/wiki/Isoperimetric_inequality

[58] Jacobi, C.G.J., Internet, en.wikipedia.org/wiki/Carl_Gustav_Jacob_Jacobi

[59] Jordan, P., Internet, http://en.wikipedia.org/wiki/Pascual_Jordan

[60] Jordan, P. and von Neumann, J., On inner products in metric spaces. Ann. of Math. **36** (1935), 719–723.

[61] Kellogg, O.D., Foundations of potential theory. Grundl. math. Wiss. vol 31, Springer, Berlin, 1967. (Reprint of 1929 edition; also reprinted by Dover Publications, 1953.)

[62] Kelvin, 1st Baron –: William Thomson, Internet, en.wikipedia.org/wiki/William_Thomson,_1st_Baron_Kelvin

[63] Kelvin transform, Internet, en.wikipedia.org/wiki/Kelvin_transform

[64] Kloosterman, H.D., Internet, en.wikipedia.org/wiki/Hendrik_Kloosterman

[65] Kolmogorov, A., Internet, http://en.wikipedia.org/wiki/Andrey_Kolmogorov

[66] Korevaar, J., Distributions defined by fundamental sequences, I-V. Indag. Math. **17** (1955), 369–389, 483–503, 663–674.

[67] Korevaar, J., Pansions (formal Hermite expansions) and the theory of Fourier transforms. Trans. Amer. Math. Soc. **91** (1959), 53–101.

[68] Korevaar, J., Mathematical methods, vol 1, Linear algebra, normed spaces, distributions, integration. Academic Press, New York, 1968. (Reprinted by Dover Publ., Mineola, N.Y., 2008.)

[69] Korevaar, J., Tauberian theory. A century of developments. Grundl. Math. Wiss., Springer, 2004.

[70] Körner, T.W., Fourier analysis. Second edition. Cambridge Univ. Press, 1989.

[71] Laguerre, E., Internet, http://en.wikipedia.org/wiki/Edmond_Laguerre

[72] Laguerre polynomials, Internet, http://en.wikipedia.org/wiki/Laguerre_polynomials

[73] Laplace, P.-S., Internet, http://en.wikipedia.org/wiki/Pierre-Simon_Laplace

[74] Laplace's equation, Internet, http://en.wikipedia.org/wiki/Laplace27s_equation

[75] Laplace operator, Internet, http://en.wikipedia.org/wiki/Laplace_operator

[76] Lebesgue, H., Internet, \http://en.wikipedia.org/wiki/Henri_Lebesgue

[77] Lebesgue integration, Internet, http://en.wikipedia.org/wiki/Lebesgue_integration

[78] Legendre, A.-M., Internet, http://en.wikipedia.org/wiki/Adrien-Marie_Legendre

[79] Legendre polynomials, Internet, http://en.wikipedia.org/wiki/Legendre_polynomials

[80] Levi, B., Internet, http://en.wikipedia.org/wiki/Beppo_Levi

[81] Lighthill, M.J., Introduction to Fourier analysis and generalised functions. Cambridge Univ. Press, 1960.

[82] Liouville, J., Internet, http://en.wikipedia.org/wiki/Joseph_Liouville

[83] Lipschitz, R.O.S., Internet, http://en.wikipedia.org/wiki/Rudolf_Lipschitz
[84] Littlewood, J.E., Internet, en.wikipedia.org/wiki/John_Edensor_Littlewood
[85] Mellin, H., Internet, http://en.wikipedia.org/wiki/Hjalmar_Mellin
[86] Mellin transform, Internet, en.wikipedia.org/wiki/Mellin_transform
[87] Neumann, J. von, Internet, en.wikipedia.org/wiki/John_von_Neumann
[88] Paley, R.E.A.C. and Wiener, N., Fourier transforms in the complex dpmain. Amer. Math. Soc. Publ. vol 19. Amer. Math. Soc. Providence, RI, 1934.
[89] Parseval, M.-A., Internet, http://en.wikipedia.org/wiki/Marc-Antoine_Parseval
[90] Parseval's theorem, Internet, http://en.wikipedia.org/wiki/Parseval's_theorem
[91] Plancherel, M., Internet, http://en.wikipedia.org/wiki/Michel_Plancherel
[92] Plancherel's theorem, Internet, http://en.wikipedia.org/wiki/Plancherel_theorem
[93] Poisson, S.D., Internet, http://en.wikipedia.org/wiki/SimC3A9on_Denis_Poisson
[94] Poisson kernel, Internet, http://en.wikipedia.org/wiki/Poisson_kernel
[95] Poisson's sum formula, Internet, en.wikipedia.org/wiki/Poisson_summation_formula
[96] Prime number theorem, Internet, http//:en.wikipedia.org/wiki/Prime_number_theorem
[97] Quantum harmonic oscillator, Internet, http://en.wikipedia.org/wiki/Quantum_harmonic_oscillator.
[98] Riemann, B., Internet, en.wikipedia.org/wiki/Bernhard_Riemann
[99] Riemann integral, Internet, http://en.wikipedia.org/wiki/Riemann_integral
[100] Riesz, F., Internet, http://encyclopedia.thefreedictionary.com/Frederic+Riesz
[101] Riesz representation theorem, Internet, http://en.wikipedia.org/wiki/Riesz_representation_theorem
[102] Riesz-Fischer theorem, Internet, en.wikipedia.org/wiki/RieszFischer_theorem
[103] Rodrigues, B.O., Internet, http://en.wikipedia.org/wiki/Olinde_Rodrigues
[104] Rodrigues' formula, Internet, http://en.wikipedia.org/wiki/Rodrigues_formula
[105] Schauder, J.P., Internet, en.wikipedia.org/wiki/Juliusz_Schauder
[106] Schauder basis, Internet, en.wikipedia.org/wiki/Schauder_basis
[107] Schläfli. L., Internet, htpp://en.wikipedia.org/wiki/Ludwig_Schlfli
[108] Schmidt, E., Internet, http://en.wikipedia.org/wiki/Erhard_Schmidt
[109] Schwartz, L., Internet, http://en.wikipedia.org/wiki/Laurent_Schwartz
[110] Schwartz, L., Théorie des distributions, I, II. Hermann, Paris, 1950. New expanded edition in one volume, 1966.
[111] Schwartz, L., Mathematics for the physical sciences. Hermann, Paris; Addison-Wesley, Reading, Mass., 1966.
[112] Schwarz, K.H.A., Internet, en.wikipedia.org/wiki/Hermann_Schwarz
[113] Spherical harmonics, Internet, en.wikipedia.org/wiki/Spherical_harmonics

[114] Stein, E.M. and Weiss, G., Introduction to Fourier analysis on Euclidean spaces. Princeton Univ. Press, 1971.

[115] Sturm, J.C.F., Internet, `http://en.wikipedia.org/wiki/Jacques_Charles_Francois_Sturm`

[116] Sturm–Liouville theory, Internet, `http://en.wikipedia.org/wiki/Sturm--Liouville_theory`

[117] Szegő, G., Orthogonal polynomials, Colloquium Publications vol. 23 (fourth ed). American Mathematical Society, Providence, R.I., 1975.

[118] Tauber, A., Internet, `en.wikipedia.org/wiki/Alfred_Tauber`

[119] Taylor, A., Introduction to functional analysis. Wiley, New York, second edition 1980.

[120] Titchmarsh, E.C., Introduction to the theory of Fourier integrals. Oxford Univ. press, 1937. Third ed., Chelsea Publ. Co, New York, 1986.

[121] de la Vallée Poussin, C.-J., Internet, `http://nl.wikipedia.org/wiki/Charles-Jean_de_La_Vallee_Poussin`

[122] Volterra, V., Internet, `en.wikipedia.org/wiki/Vito_Volterra`

[123] Weierstrass, K., Internet, `en.wikipedia.org/wiki/Karl_Weierstrass`

[124] Wiener, N., The Fourier integral and certain of its applications. Cambridge Univ. Press, Cambridge, 1933. (Reprinted 1988.)

[125] Wolff, J., Fouriersche Reihen. Noordhoff, Groningen, 1931.

[126] Zhang, Gong-zhing, Theory of distributions of $\mathcal{S}$ type and pansions. Chinese Math.-Acta **4** (1963), 211–221. (Translated from the Chinese.)

[127] Zorn, M.A., Internet, `http://en.wikipedia.org/wiki/Max_August_Zorn`

[128] Zorn's Lemma, Internet, `http://en.wikipedia.org/wiki/Zorn's_lemma`

[129] Zygmund, A., Trigonometric series, vol I,II. Third ed. Cambridge Univ. Press, 2002.

Index