Maxwell's Theory of Electrodynamics

From lectures given by
Yaroslav Kurylev

Edited, annotated and expanded by
Adam Townsend & Giancarlo Grasso

19 May 2012

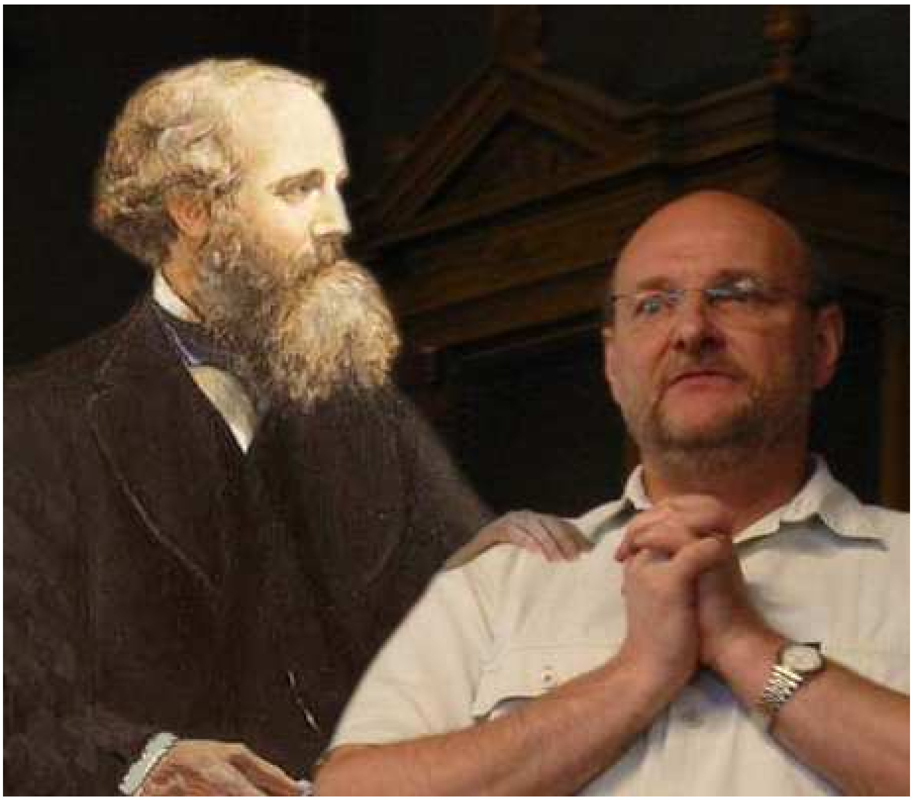

Figure 1: Our Dear Leader and His Spirit Guide

Contents

Editors' note

This set of course notes was released initially in April 2011 for the University College London mathematics course MATH3308, *Maxwell's Theory of Electrodynamics*, a ten-week course taught by Professor Yaroslav Kurylev between January and March 2011. The editors were members of this class. These notes were published to `http://adamtownsend.com` and if you wish to distribute these notes, *please direct people to the website* in order to download the latest version: we expect to be making small updates or error corrections sporadically. If you spot any errors, you can leave a comment on the website or send an email to the authors at `adam@adamtownsend.com`. We hope you enjoy the notes and find them useful!

AKT GWG
London, April 2011

Acknowledgments

The editors would like to thank the following for their help in the production of this set of course notes:

- **Lisa** for catapults and providing essential Russian to English translation;
- **Kav** for questions;
- **The Permutation Boys** (Jonathan, James and Charles) for permuting;
- **Ewan** for patiently being Ian;
- **Prof. Kurylev** for coffee and many, many invaluable life lessons.

We would also like to thank those who have offered corrections since the original publication.

Chapter 1

Introduction to the Theory of Distributions

1.1 Introduction

James Clerk Maxwell (1831–1879), pictured on page 2, was a Scottish mathematician and physicist who is attributed with formulating classical electromagnetic theory, uniting all previously unrelated observations, experiments and equations of electricity, magnetism and even optics into one consistent theory. Maxwell's equations demonstrated that electricity, magnetism and even light are all manifestations of the same phenomenon, namely the *electromagnetic field*. By the end of the course we'll have built up the mathematical foundation that we need in order to understand this fully (more so than Maxwell himself!), and we'll have derived and used Maxwell's famous equations. Although they bear his name, like much of mathematics, his equations (in earlier forms) come from the work of other mathematicians, but they were brought together for the first time in his paper, *On Physical Lines of Force*, published between 1861 and 1862*.

The first chapter of this course is concerned with putting together a mathematically sound foundation which we can build Maxwell's theory of electrodynamics upon. Maxwell's theory introduces functions that don't have derivatives in the classical sense, so we use a relatively new addition to analysis in order to understand this. This foundation is the *Theory of Distributions*. Distributions (not related to probability distributions) are the far-reaching generalisations of the notion of a *function* and their theory has become an indispensable tool in modern analysis, particularly in the theory of partial differential equations. The theory as a whole was put forth by the French† mathematician Laurent Schwartz in the late 1940s, however, its most important ideas (of so-called 'weak solutions') were introduced back in 1935 by the Soviet mathematician Sergei Sobolev. This course will cover only the very beginnings of this theory.

*Paraphrased from Wikipedia

†Honest!

1.2 Definitions and notation

Let me start by reminding you of some definitions and notation:

Definition 1.1 A *compact set* K in a domain Ω, denoted $K \Subset \Omega$ is a *closed, bounded* set.

Notation 1.2 $C^\infty(\Omega)$ is the set of *continuous* functions which can be *differentiated* infinitely many times in a domain Ω.

Notation 1.3 $C_0(\Omega)$ is the set of *continuous* functions which are *equal to 0* outside some compact $K \Subset \Omega$. That is to say, if $f \in C_0$, $\forall x \notin K, f(x) = 0$. This is known as the set of continuous functions with *compact support*

Notation 1.4 $C_0^\infty(\Omega)$ is the set of *continuous*, *infinitely differentiable* functions with *compact support*

Notation 1.5 We use $B_r(x_0)$ to denote the closed ball of radius r centred at a point x_0 in our domain Ω: that is to say,

$$B_r(x_0) = \{x \in \Omega : |x - x_0| \leq r\}$$

where $|\cdot|$ is the norm of our domain. For this course we'll always say that our domains are subsets of $\mathbb{R}^n$, so we don't need to worry about special types of norm: $|\mathbf{v}|$ is the standard (Euclidean) length of the vector $\mathbf{v}$, which we work out in the usual Pythagorean way.*

Note: we'll use the convention of writing vectors in boldface (e.g. $\mathbf{v}$) from chapter 2 onwards—the first chapter deals heavily in analysis and since we are being general in which domain we're working in ($\mathbb{R}$, $\mathbb{R}^2$ or $\mathbb{R}^3$), the distinction between one-dimensional scalars and higher-dimensional vectors is unnecessary; most of our results hold in n dimensions, and excessive boldface would clutter the notes. As such we'll use standard italic (e.g. v) unless clarity demands it of us.

1.2.1 The du Bois-Reymond lemma and test functions

We start by proving a very useful lemma, first proved by the German mathematician Paul du Bois-Reymond.

Lemma 1.6 Let $\Omega \subset \mathbb{R}^n$ be an open domain (bounded or unbounded). Let $f(x) \in C(\Omega)$. Assume that for any $\phi \in C_0^\infty(\Omega)$,

$$\int_\Omega f(x)\phi(x)dx = 0.$$

*You may find that other authors use the notations $B(x_0, r)$, $\overline{B}(x_0, r)$ or $B_c(x_0, r)$ (where the bar and c denote 'closed').

Then $f \equiv 0$.

Proof: We'll approach this using contradiction. Say we have $x_0 \in \Omega$ such that $f(x_0) > 0$.

$f \in C(\Omega)$, i.e.

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. } |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon$$

Let $\varepsilon = \frac{1}{2} f(x_0)$, i.e. there exists a ball of radius δ, $B_\delta(x_0)$ such that $f(x) > \frac{1}{2} f(x_0)$ for all $x \in B_\delta(x_0)$.

Choose ϕ from C_0^∞ such that:

1. $\phi(x) \geq 0$ for all x
2. $\phi(x) = 1$ for $x \in B_{\frac{\delta}{2}}(x_0)$
3. $\phi(x) = 0$ for $x \notin B_\delta(x_0)$.

Then look at

$$\int\limits_{\Omega} f(x)\phi(x)dx = \int\limits_{B_\delta(x_0)} f(x)\phi(x)dx$$

since ϕ is zero outside of this ball. But

$$\int\limits_{B_\delta(x_0)} f(x)\phi(x)dx \geq \int\limits_{B_{\frac{\delta}{2}}(x_0)} f(x)\phi(x)dx$$

since all the $f(x)\phi(x)$ terms are positive. Then

$$\int\limits_{B_{\frac{\delta}{2}}(x_0)} f(x)\phi(x)dx = \int\limits_{B_{\frac{\delta}{2}}(x_0)} f(x)dx$$

since $\phi(x) = 1$ by condition 2 above. Then

$$\int\limits_{B_{\frac{\delta}{2}}(x_0)} f(x)dx \geq \frac{1}{2} f(x_0) \cdot \mathrm{Vol}(B_{\frac{\delta}{2}}(x_0)) > 0$$

by our choice of ε above.

But our initial assumption was that

$$\int\limits_{\Omega} f(x)\phi(x)dx = 0$$

which is a **contradiction**. $\square$

The crucial result underlining the proof of this lemma is the following:

Proposition 1.7 Let K_1 and K_2 be two compacts in $\mathbb{R}^n$ such that $K_1 \Subset K_2$. Then there exists $\phi \in C_0^\infty(\mathbb{R}^n)$ such that

1. $0 \le \phi(x) \le 1$ on $\mathbb{R}^n$;
2. $\phi(x) = 1$ on K_1;
3. $\phi(x) = 0$ on $\mathbb{R}^n \setminus K_2$.

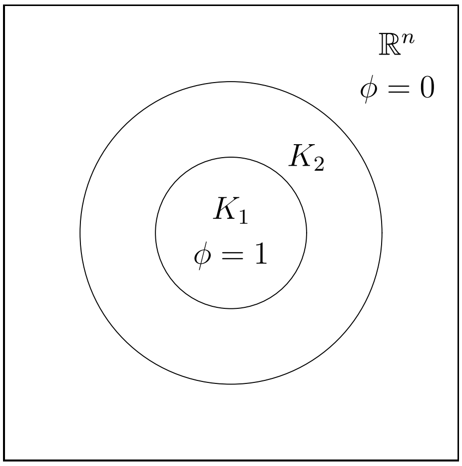

Figure 1.1: Proposition 1.7 tells us this exists

We can take from the lemma a simple corollary:

Corollary 1.8 Let $f, g \in C(\Omega)$. Then if $\forall \phi \in C_0^\infty(\Omega)$,

$$\int_\Omega f(x)\phi(x)dx = \int_\Omega g(x)\phi(x)dx$$

then $f = g$.

Proof: Subtracting one side from the other gives

$$\int_\Omega f(x)\phi(x)dx - \int_\Omega g(x)\phi(x)dx = 0$$

$$\int_\Omega (f(x) - g(x))\phi(x)dx = 0$$

Then $f(x) - g(x) = 0$ by Lemma 1.6, i.e. $f(x) = g(x)$. $\square$

To understand the importance of this result, remember that normally to test whether two functions f and g are equal, we have to compare their values at all points $x \in \Omega$. By the Du Bois-Reymond lemma, however, this is equivalent to the fact that integrals of f and g with all $\phi \in C_0^\infty(\Omega)$ are equal. Thus, functions are *uniquely characterised* by their integral with $C_0^\infty(\Omega)$ functions. It is, of course, crucial that we look at the values of integrals of f with *all* $C_0^\infty(\Omega)$ functions, i.e. to *test* f with all $C_0^\infty(\Omega)$ functions which are, therefore, called *test functions*.

Definition 1.9 $C_0^\infty(\Omega)$ is called the *space of test functions* and is often denoted $\mathcal{D}(\Omega)$

1.2.2 Topological vector spaces and multi-indices

Definition 1.10 $\mathcal{D}(\Omega)$ is a *topological vector space*, i.e.

1. It is a linear vector space, i.e. if $f, g \in \mathcal{D}(\Omega)$ and $\lambda, \mu \in \mathbb{R}$,

$$h(x) = (\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x) \quad \in \mathcal{D}(\Omega) \quad x \in \Omega$$

   and $h(x) = 0$ for $x \notin K_f \cup K_g$, i.e. x not in the compact support of either.

2. There is a notion of convergence, i.e. there is a meaning to $f_k \xrightarrow[k\to\infty]{} f$. This is defined shortly.

For the next few definitions we need to introduce a further few convenient definitions:

Definition 1.11 A *multi-index* $\boldsymbol{\alpha}$ is an n-dimensional vector, where $n = \dim(\Omega)$, with non-negative integer components:

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n), \quad \alpha_i = 0, 1, \ldots \quad i = 1, \ldots, n$$

Multi-indices are convenient when we want to write high-order partial derivatives, namely

$$\partial^{\boldsymbol{\alpha}} \phi(x) = \frac{\partial^{|\boldsymbol{\alpha}|}\phi}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \quad j = 1, \ldots, n, \quad |\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_n.$$

1.2.3 Null set and support

Definition 1.12 The *null set* of a function f, $\mathcal{N}(f)$, is the maximal open set in Ω where $f \equiv 0$. In other words, $x_0 \in \mathcal{N}(f)$ if it has a neighbourhood $x_0 \in U \subset \Omega$ with $f(x) = 0$ when $x \in U$ and if any other open set A holds this property, then $A \subset \mathcal{N}(f)$

Definition 1.13 The *support* of a function f, $\operatorname{supp}(f)$ is the complement of $\mathcal{N}(f)$ in Ω, i.e.

$$\operatorname{supp}(f) = \Omega \setminus \mathcal{N}(f)$$

Note that by this definition, $\operatorname{supp}(f)$ is always closed in Ω.

Example 1.14 Let $\Omega = \mathbb{R}$ and

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$$

Then $\mathcal{N}(f) = (-\infty, 0)$ and $\operatorname{supp}(f) = [0, \infty)$.

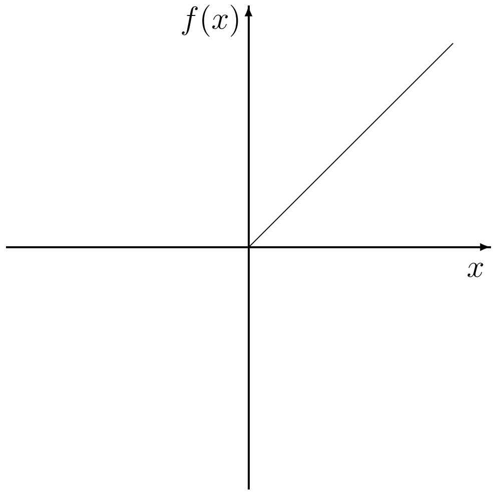

Figure 1.2: Look, a picture of the function!

1.2.4 Convergence

Definition 1.15 A function ϕ_k *converges* to ϕ in $\mathcal{D}(\Omega)$ if the following two conditions are satisfied:

1. For any multi-index $\boldsymbol{\alpha}$ and any $x \in \Omega$,

$$\partial^{\boldsymbol{\alpha}} \phi_p(x) \to \partial^{\boldsymbol{\alpha}} \phi(x), \quad \text{as } p \to \infty$$

   i.e. all derivatives converge

2. There is a compact $K \Subset \Omega$ such that

$$\operatorname{supp}(\phi_p), \operatorname{supp}(\phi) \subset K, \quad p = 1, 2, \ldots$$

Problem 1.16 For any $\phi \in \mathcal{D}(\Omega)$ and any multi-index $\boldsymbol{\beta}$, show that

$$\partial^{\boldsymbol{\beta}} \phi \in \mathcal{D}(\Omega)$$

Proof:

$$\mathcal{N}(\partial^{\boldsymbol{\beta}} \phi) \supset \mathcal{N}(\phi) \implies \operatorname{supp}(\partial^{\boldsymbol{\beta}} \phi) \subset \operatorname{supp}(\phi) \Subset \Omega$$

□

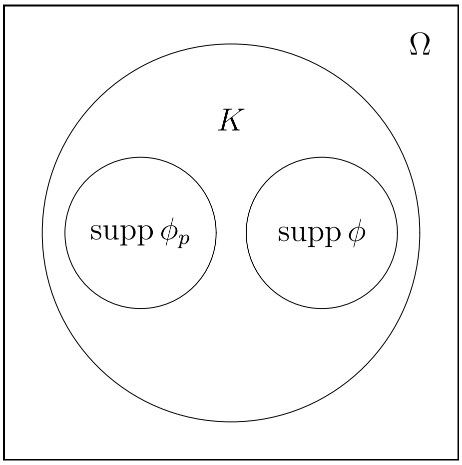

Figure 1.3: Definition 1.15(2) in picture form

Problem 1.17 If $\zeta \in C^\infty(\Omega)$, then for any $\phi \in \mathcal{D}(\Omega)$, show

$$\psi = \zeta\phi \in \mathcal{D}(\Omega)$$

Proof:

$$\operatorname{supp}(\psi) \subset \operatorname{supp}(\phi) \Subset \Omega$$

□

Problem 1.18 Let $\phi_p \to \phi$ in $\mathcal{D}(\Omega)$ and let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ be some fixed multi-index. Show that

$$\partial^{\boldsymbol{\beta}}\phi_p \to \partial^{\boldsymbol{\beta}}\phi \quad \text{in } \mathcal{D}(\Omega).$$

Solution This just follows from condition 1 of the definition of ϕ_p converging to ϕ. ✓

Problem 1.19 $*$ Let $\zeta \in C^\infty(\Omega)$, i.e. ζ is an infinitely differentiable function in Ω. Show that, for $\phi_p \to \phi$ in $\mathcal{D}(\Omega)$,

$$\zeta\phi_p \to \zeta\phi \quad \text{in } \mathcal{D}(\Omega).$$

Solution Want to show that if $\psi_p = \zeta\phi_p$, $\psi = \zeta\phi$, that

1. $\forall \boldsymbol{\beta}$, $\partial^{\boldsymbol{\beta}}\psi_p \to \partial^{\boldsymbol{\beta}}\psi$
2. $\exists K$ s.t. $\operatorname{supp}(\psi_p), \operatorname{supp}(\psi) \subset K$

So let's do it.

1.
$$\begin{aligned}\partial^{\boldsymbol{\beta}}\psi_p &= \partial^{\boldsymbol{\beta}}(\zeta\phi_p)\\ &= \sum_{\gamma:\gamma\leq\beta}\binom{\beta}{\gamma}\left(\partial^{\beta-\gamma}\zeta\right)\left(\partial^{\gamma}\phi_p\right)\end{aligned}$$

which is the sum of C^∞ terms. So as $p\to\infty$, since $\phi_p\to\phi$ in $\mathcal{D}(\Omega)$,

$$\begin{aligned}&\to \sum_{\gamma:\gamma\leq\beta}\binom{\beta}{\gamma}\left(\partial^{\beta-\gamma}\zeta\right)\left(\partial^{\gamma}\phi\right)\\ &= \partial^{\boldsymbol{\beta}}(\zeta\phi).\end{aligned}$$

2. $\mathcal{N}(\psi)\supset\mathcal{N}(\phi) \implies \text{supp}(\psi)\subset\text{supp}(\phi)\Subset\Omega$ since $\phi_p\to\phi$ in $\mathcal{D}(\Omega)$.
$\mathcal{N}(\psi_p)\supset\mathcal{N}(\phi_p) \implies \text{supp}(\psi_p)\subset\text{supp}(\phi_p)\Subset\Omega$.
And so $\exists K\Subset\Omega$ such that $\text{supp}(\psi),\text{supp}(\psi_p)\in K$.

✓

Problem 1.20 If $\phi_p\to\phi$, $\psi_p\to\psi$ in $\mathcal{D}(\Omega)$ and $\lambda,\mu\in\mathbb{R}$, show that

$$\lambda\phi_p+\mu\psi_p\to\lambda\phi+\mu\psi \quad \text{in } \mathcal{D}(\Omega)$$

Problem 1.21 If $\phi_p\to\phi$, $\psi_p\to\psi$ in $\mathcal{D}(\Omega)$ and $\lambda_p\to\lambda$, $\mu_p\to\mu$ in $\mathbb{R}$, show that

$$\lambda_p\phi_p+\mu_p\psi_p\to\lambda\phi+\mu\psi \quad \text{in } \mathcal{D}(\Omega)$$

1.3 Functionals

Now we move on to *functionals*:

Definition 1.22 A function $F:V\to\mathbb{R}$, mapping from a vector space onto the reals, is called a *functional*.

Definition 1.23 A *continuous, linear functional* F has the properties,

1. Linearity:
$$F(\lambda\mathbf{v}+\mu\mathbf{w})=\lambda F(\mathbf{v})+\mu F(\mathbf{w});\quad \mathbf{v},\mathbf{w}\in V;\quad \lambda,\mu\in\mathbb{R} \tag{1.1}$$

2. Continuity:
$$F(\mathbf{v}_p)\to F(\mathbf{v}) \quad \text{as } \mathbf{v}_p\to\mathbf{v} \tag{1.2}$$

Problem 1.24 1. Let F be a linear functional on a topological vector space V. Show that $F(\mathbf{0}) = 0$, where $\mathbf{0}$ is the zero vector in V.

2. Assume, in addition, that F satisfies

$$F(\mathbf{v}_p) \to 0, \quad \text{as } \mathbf{v}_p \to \mathbf{0}$$

Show that then F is continuous.

1.3.1 Dual space

Definition 1.25 The space of *continuous linear functionals* on V is called the *dual space* to V and is denoted[a] by V'.

The dual space is itself a topological vector space if, for $\lambda, \mu \in \mathbb{R}$ and $F, G \in V'$ we define $\lambda F + \mu G$ by

$$(\lambda F + \mu G)(\mathbf{v}) = \lambda F(\mathbf{v}) + \mu G(\mathbf{v}), \quad \mathbf{v} \in V \tag{1.3}$$

and $F_p \to F$ if, for any $\mathbf{v} \in V$,

$$F_p(\mathbf{v}) \to F(\mathbf{v}), \quad \text{as } p \to \infty. \tag{1.4}$$

Problem 1.26 Show that $H = \lambda F + \mu G$ defined by (1.3) is a continuous linear functional on V, i.e. H satisfies (1.1), (1.2).

Proof: We need to check that this object is linear and continuous:

1. Linearity:

$$(\lambda F + \mu G)(\alpha \mathbf{v} + \beta \mathbf{w}) = \lambda F(\alpha \mathbf{v} + \beta \mathbf{w}) + \mu G(\alpha \mathbf{v} + \beta \mathbf{w})$$

   and each F, G are linear functionals, therefore

$$\begin{aligned} &= \lambda(\alpha F(\mathbf{v}) + \beta F(\mathbf{w})) + \mu(\alpha G(\mathbf{v}) + \beta G(\mathbf{w})) \\ &= \alpha[\lambda F(\mathbf{v}) + \mu G(\mathbf{v})] + \beta[\lambda F(\mathbf{w}) + \mu G(\mathbf{w})] \\ &= \alpha(\lambda F + \mu G)(\mathbf{v}) + \beta(\lambda F + \mu G)(\mathbf{w}) \end{aligned}$$

2. Continuity:

$$\begin{aligned} (\lambda F + \mu G)(\mathbf{v}_p) =& \lambda F(\mathbf{v}_p) + \mu G(\mathbf{v}_p) \\ \xrightarrow[p\to\infty]{} & \lambda F(\mathbf{v}) + \mu G(\mathbf{v}) \\ =& (\lambda F + \mu G)(\mathbf{v}) \end{aligned}$$

[†]Other, if not most, authors denote the dual space to V by V^*

□

Problem 1.27 $*$ Show that if $F_p \to F$ and $G_p \to G$, where $F_p, F, G_p, G \in V'$, then

$$\lambda F_p + \mu G_p \to \lambda F + \mu G.$$

Solution Again, just like above,

$$(\lambda F_p + \mu G_p)(\mathbf{v}) = \lambda F_p(\mathbf{v}) + \mu G_p(\mathbf{v})$$

and, by the information given in the question,

$$\begin{aligned} &\xrightarrow[p\to\infty]{} \lambda F(\mathbf{v}) + \mu G(\mathbf{v}) \\ &= (\lambda F + \mu G)(\mathbf{v}) \end{aligned}$$

✓

Problem 1.28 If $F_p \to F$, $G_p \to G$, $\lambda_p \to \lambda$, $\mu_p \to \mu$, show that

$$\lambda_p F_p + \mu_p G_p \to \lambda F + \mu G$$

Solution We want to show then that

$$(\lambda_p F_p + \mu_p G_p)(\mathbf{v}) \to (\lambda F + \mu G)(\mathbf{v})$$

So here we go, quite succinctly.

$$\begin{aligned} (\lambda_p F_p + \mu_p G_p)(\mathbf{v}) &= \lambda_p F_p(\mathbf{v}) + \mu_p G_p(\mathbf{v}) \\ &\xrightarrow[p\to\infty]{} \lambda F(\mathbf{v}) + \mu G(\mathbf{v}) \\ &= (\lambda F + \mu G)(\mathbf{v}) \end{aligned}$$

✓

Example 1.29 Let $V = \mathbb{R}^2$, the space of 2-dimensional vectors. Recall that $\mathbf{v}_p \to \mathbf{v}$ iff $\|\mathbf{v}_p - \mathbf{v}\| \to 0$. Then V' is actually V itself and we define, for $\mathbf{w} \in V'$ (which is again just a 2-vector) and $\mathbf{v} \in V$,

$$\mathbf{w}(\mathbf{v}) = \mathbf{w} \cdot \mathbf{v} = w_1 v_1 + w_2 v_2$$

i.e. the scalar product of $\mathbf{w}$ and $\mathbf{v}$.

1.4 Distributions

Let us return to the topological vector space $\mathcal{D}(\Omega) = C_0^\infty(\Omega)$. Then it has a dual space of its functionals.

Definition 1.30 The topological vector space $\mathcal{D}'(\Omega)$ of continuous linear functionals on $\mathcal{D}(\Omega)$ is called the space of *distributions* on Ω.

Let me just remind you of a definition and result from real analysis (MATH7102):

Definition 1.31 We say that ϕ_p converges to ϕ ($\phi_p \xrightarrow[p\to\infty]{} \phi$) *uniformly* if:

$$\forall \varepsilon > 0 \; \exists P(\varepsilon) \text{ s.t. } p > P(\varepsilon) \implies |\phi_p(x) - \phi(x)| < \varepsilon \; \forall x \in \Omega$$

Lemma 1.32 Let $\phi_p \in C(\Omega)$ and $\phi_p(x) \to \phi(x)$ for any $x \in \Omega$. Then for any compact $K \Subset \Omega$, $\phi_p \to \phi$ *uniformly.*

What are some examples of distributions?

Proposition 1.33 Take any $f \in C(\Omega)$. For any $\phi \in \mathcal{D}(\Omega)$, consider

$$F_f(\phi) = \int_\Omega f(x)\phi(x)dx.$$

Then F_f is a distribution (a continuous linear functional on $\mathcal{D}(\Omega)$) called *a distribution associated with the function* f. That is to say, the integrals form a linear continuous functional in $\mathcal{D}(\Omega)$.

Proof:

1. Linearity:

$$\begin{aligned} F_f(\lambda\phi + \mu\psi) &= \int_\Omega f(x)\left(\lambda\phi(x) + \mu\psi(x)\right)dx \\ &= \lambda \int_\Omega f(x)\phi(x)dx + \mu \int_\Omega f(x)\psi(x)dx \\ &= \lambda F_f(\phi) + \mu F_f(\psi) \end{aligned} \tag{1.5}$$

   i.e F_f satisfies (1.1).

2. Continuity:

   Let $\phi_p \to \phi$ in $\mathcal{D}(\Omega)$.

   Then we're told in definition 1.15 that for *any* multi-index $\boldsymbol{\alpha}$,

$$\partial^{\boldsymbol{\alpha}}\phi_p(x) \to \partial^{\boldsymbol{\alpha}}\phi(x) \quad x \in \Omega$$

   So take $\boldsymbol{\alpha} = (0, \dots, 0)$, then obviously since $\partial^0\phi = \phi$, we get

$$\begin{gathered} \phi_p(x) \to \phi(x), \quad x \in \Omega \\ \implies f(x)\phi_p(x) \to f(x)\phi(x) \end{gathered}$$

   But we *can't* say that $\int f(x)\phi_p(x)\,dx \to \int f(x)\phi(x)\,dx$ because this simply isn't true.

Counterexample: Take the interval $(0, 1)$ and

$$\phi_p(x) = \begin{cases} p & \text{if } x \in (0, \frac{1}{p}) \\ 0 & \text{if } x \in [\frac{1}{p}, 1) \end{cases}$$

Then $\forall x \in (0, 1)$, $\phi_p(x) \to 0$ but $\int_0^1 \phi_p(x)dx = 1$.

We need the extra condition that $\phi_p \to \phi$ *uniformly*.

Recall that the definition of convergence (1.15) says that there is a compact $K \Subset \Omega$ such that

$$\text{supp}(\phi_p), \text{supp}(\phi) \subset K, \quad p = 1, 2, \dots$$

Then by Lemma 1.32 above, we have all we need and

$$F_f(\phi_p) = \int_K f(x)\phi_p(x)dx \to \int_K f(x)\phi(x)dx = F_f(\phi). \tag{1.6}$$

This completes the proof. □

So in summary, every $f \in C(\Omega)$ defines a *distribution* $F_f \in \mathcal{D}'(\Omega)$, where

$$F_f(\phi) = \int_\Omega f(x)\phi(x)dx$$

Moreover, the Du Bois-Reymond lemma says that to two different functions f_1 and f_2, we associate different F_{f_1} and F_{f_2}.

So we say that C is embedded in $\mathcal{D}'$, $C(\Omega) \subset \mathcal{D}'(\Omega)$, and this embedding is continuous.

Previously we defined uniform convergence of a function, let us state what it means for a function to converge pointwise:

Definition 1.34 A function $f_p \xrightarrow[p\to\infty]{} f$ *pointwise* in $C(\Omega)$ if

$$f_p(x) \to f(x) \quad \forall x \in \Omega$$

Problem 1.35 $*$ If $f_p, f \in C(\Omega)$ and $f_p \to f$ pointwise, show that

$$F_{f_p} \to F_f \quad \text{in } \mathcal{D}'(\Omega)$$

Proof: $F_{f_p} \to F_f$ in $\mathcal{D}'(\Omega)$ means that

$$\begin{aligned}\forall \phi \in \mathcal{D}'(\Omega), \quad F_{f_p}(\phi) \to F_f(\phi) &\iff \int_\Omega f_p(x)\phi(x)dx \to \int_\Omega f(x)\phi(x)dx \\ &\iff \int_K f_p(x)\phi(x)dx \to \int_K f(x)\phi(x)dx\end{aligned}$$

where $K = \text{supp}(\phi)$.

Since $f_p\phi \to f\phi$ *pointwise*, and $f_p\phi$ and $f\phi$ are *continuous*, then $f_p\phi \to f\phi$ *uniformly* in K. □

1.4.1 Distributions of different forms

This leads us to the next question: are there any distributions $F \in \mathcal{D}'(\Omega)$ which are not of the form F_f for some $f \in C(\Omega)$?

Definition 1.36 Let $y \in \Omega$. Define the *Dirac delta functional*,

$$\delta_y(\phi) := \phi(y). \tag{1.7}$$

Problem 1.37 Show that the Dirac delta functional is a distribution.

Proof:

1. Linearity:

$$\begin{aligned} \delta_y(\lambda\phi + \mu\psi) &= (\lambda\phi + \mu\psi)(y) \\ &= \lambda\phi(y) + \mu\psi(y) \\ &= \lambda\delta_y(\phi) + \mu\delta_y(\psi) \end{aligned}$$

2. Continuity: if $\phi_p \to \phi$ in $\mathcal{D}$, then, in particular, $\phi_p(y) \to \phi(y)$, so that

$$\delta_y(\phi_p) = \phi_p(y) \to \phi(y) = \delta_y(\phi)$$

□

Problem 1.38 Show that there is no $f \in C(\Omega)$ such that $\delta_y = F_f$.

Solution Without loss of generality, take $y = 0$. Assume $\delta = F_f$ with $f \in C(\Omega)$ and take a function $\phi_\varepsilon \in \mathcal{D}(\Omega)$ such that

$$\begin{cases} 0 \leq \phi_\varepsilon(x) \leq 1 & \text{in } \Omega \\ \phi_\varepsilon(x) = 1 & \text{in } B_\varepsilon \\ \phi_\varepsilon(x) = 0 & \text{outside } B_{2\varepsilon} \end{cases}$$

$\delta(\phi_\varepsilon) = \phi_\varepsilon(0) = 1$ so

$$\begin{aligned} 1 = \delta(\phi_\varepsilon) = F_f(\phi_\varepsilon) &= \left| \int_\Omega f(x)\phi_\varepsilon(x)\, dx \right| \\ &= \left| \int_{B_{2\varepsilon}} f(x)\phi_\varepsilon(x)\, dx \right| \\ &\le \int_{B_{2\varepsilon}} |f(x)|\, |\phi_\varepsilon(x)|\, dx \\ &\le \int_{B_{2\varepsilon}} |f(x)|\, dx \end{aligned}$$

Let $f \in C(\Omega)$. Since f is continuous, it is bounded in any compact, i.e.

$$|f(x)| \le C \quad \text{in } B_1$$

(where we've picked the radius of the ball to be 1 fairly arbitrarily). Therefore

$$\begin{aligned} 1 \le C &\int_{B_{2\varepsilon}} dx \\ &= C \cdot (\text{vol. of ball in } n\text{-dim. case}) \cdot (2\varepsilon)^n \\ &\to 0 \text{ as } \varepsilon \to 0 \end{aligned}$$

which is a **contradiction**. ✓

Now for more examples of distributions:

Example 1.39 Let $\Omega = \mathbb{R}^3$. Consider the plane, say $x_3 = a$. Let $h(x_1, x_2)$ be a continuous function of two variables, i.e. $h \in C(\mathbb{R}^2)$. Define

$$F_{a,h}(\phi) = \int h(x_1, x_2)\phi(x_1, x_2, a)\, dx_1\, dx_2,$$

i.e. the integral over the plane $x_3 = a$:

$$= \int_{\substack{\text{plane} \\ x_3 = a}} h(x)\phi|_{\mathbb{R}}\, dA$$

Then $F_{a,h} \in \mathcal{D}'(\mathbb{R}^3)$ (we could denote this by $h\delta_a(x_3)$).

Example 1.40 More generally, let $S \subset \Omega \subset \mathbb{R}^3$ be a surface and h a continuous function on S. Define

$$F_{S,h}(\phi) = \int_S h\phi\, dA_S,$$

where dA_S is the area element on S, known as the surface integral from MATH1402.

Example 1.41 Let now, L be a curve in $\Omega \subset \mathbb{R}^2$ or $\Omega \subset \mathbb{R}^3$ and h be some continuous function on L. We define

$$F_{L,h}(\phi) = \int_L h\phi \, d\ell$$

where $d\ell$ is the length element on L.

Problem 1.42 Show that $F_{a,h} \in \mathcal{D}'(\mathbb{R}^3)$, as well as $F_{S,h}, F_{L,f} \in \mathcal{D}'(\Omega)$.

1.5 Support of a Distribution

We know what $\operatorname{supp}(f)$ is when $f \in C(\Omega)$: this is the closure of the set of points where $f(x) \neq 0$. This definition uses the values of the function at all points. Can we find $\operatorname{supp}(f)$ looking only at

$$F_f(\phi) = \int_\Omega f(x)\phi(x)dx?$$

To this end, observe that if $\operatorname{supp}(\phi) \subset \mathcal{N}(f)$, then

$$F_f(\phi) = 0.$$

Moreover, by the Du Bois-Reymond lemma, if $U \subset \Omega$ is open and, for any $\phi \in \mathcal{D}(\Omega)$ with $\operatorname{supp}(\phi) \subset U$,

$$F_f(\phi) = \int_\Omega f(x)\phi(x)dx = 0,$$

then $f = 0$ in U, i.e. $U \subset \mathcal{N}(f)$. This makes it possible to find $\mathcal{N}(f)$ and, therefore, $\operatorname{supp}(f)$, looking only at F_f.

This also motivates the following definition:

Definition 1.43 Let $F \in \mathcal{D}'(\Omega)$. We say that $F = 0$ in an open $U \subset \Omega$ (we can denote this $F|_U$) if

$$F(\phi) = 0, \quad \phi \in \mathcal{D}(\Omega), \operatorname{supp}(\phi) \subset U. \tag{1.8}$$

This is a very natural definition: that, the distribution F is zero in U if, for whichever function $\phi \in \mathcal{D} = C_0^\infty$ we put into it, the result is zero so long as the support of ϕ (i.e. the non-zero bits) is contained entirely in U.

Definition 1.44 The null-set $\mathcal{N}(F)$ is the largest open set U in Ω such that (1.8) is valid when $\operatorname{supp}(\phi) \subset U$ (i.e. the largest open set where $F = 0$).

Definition 1.45 The support $\text{supp}(F)$ is the closed set

$$\text{supp}(F) = \Omega \setminus \mathcal{N}(F).$$

The consistency of this definition is based on the fact (which I shan't prove) that

$$F = 0 \text{ in } U_1 \text{ and } F = 0 \text{ in } U_2 \implies F = 0 \text{ in } U_1 \cup U_2. \tag{1.9}$$

Problem 1.46 Show that $\text{supp}(\delta_y) = \{y\}$

Proof: If $\text{supp}(\phi) \subset \Omega \setminus \{y\}$, then $\phi(y) = 0$. Therefore,

$$\delta_y(\phi) = \phi(y) = 0.$$

Thus

$$(\Omega \setminus \{y\}) \subset \mathcal{N}(\delta_y).$$

Now let's assume that $y \in \mathcal{N}(\delta_y)$. Then $\mathcal{N}(\delta_y) = \Omega$ since we've just added $\{y\}$ to $(\Omega \setminus \{y\})$, and

$$\mathcal{N}(\delta_y) = \Omega \iff \delta_y(\phi) = 0 \quad \forall \phi \in \mathcal{D}(\Omega).$$

But Proposition 1.7 says that there exists $\phi \in \mathcal{D}(\Omega)$ such that $\phi(y) = 1$. But then $\delta_y(\phi) = \phi(y) = 1$, which is clearly a **contradiction**.

Hence $y \notin \mathcal{N}(\delta_y)$. □

Problem 1.47 Suppose $S \subset \Omega$ and h is a continuous function on S. Show that $\text{supp}(F_{S,h}) \subset S$.

Proof:

$$F_{S,h}(\phi) = \int_S h(x)\phi|_S(x)\, dA$$

Look at the supports:

$$\Omega \setminus S \subset \mathcal{N}(F_{S,h})$$

since F is integrating over an area S (if you take away S you're left with nothing). Hence, taking complements,

$$S \supset \text{supp}(F_{S,h})$$

□

Actually, $\text{supp}(F_{S,h}) = S \cap \text{supp}(h)$.

Problem 1.48 $*$ Show that $\text{supp}(F_{L,h}) \subset L$.

Proof:

$$F_{L,h}(\phi) = \int\limits_L h(x)\phi|_L(x)\, d\ell$$

Look at the supports:

$$\Omega \setminus L \subset \mathcal{N}(F_{L,h})$$

since F is integrating over a line L (if you take away L you're left with nothing). Hence, taking complements,

$$L \supset \mathrm{supp}(F_{L,h})$$

□

1.5.1 Distributions with compact support ($\mathcal{E}'$)

The subclass of distributions with compact support is denoted by

$$\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega).$$

Definition 1.49 $F \in \mathcal{E}'(\Omega)$ if $\mathrm{supp}(F)$ is compact in Ω.

The importance of $\mathcal{E}'(\Omega)$ is that we can unambiguously define $F(\psi)$ for any $\psi \in C^\infty(\Omega)$ when $F \in \mathcal{E}'(\Omega)$.

Indeed, let $\mathrm{supp}(F) \subset K_1 \subset K_2 \Subset \Omega$, where K_1 and K_2 satisfy the conditions of Proposition 1.7.

Let

$$\phi(x) \in \mathcal{D}(\Omega) = \begin{cases} 1 & \text{if } x \in K_1 \\ 0 & \text{if } x \in \Omega \setminus K_2 \end{cases}$$

(the existence of such ϕ is guaranteed by Proposition 1.7).

We now define

$$F(\psi) := F(\phi\psi). \tag{1.10}$$

Note that this definition is independent of the choice of ϕ satisfying

$$\phi(x) = 1 \quad x \in \mathrm{supp}(F).$$

Indeed, if $\widehat{\phi}$ is another function with this property, then

$$F(\phi\psi) - F(\widehat{\phi}\psi) = F((\phi - \widehat{\phi})\psi) = 0,$$

since $\mathrm{supp}(F) \cap \mathrm{supp}((\phi - \widehat{\phi})\psi) = \emptyset$.

Definition 1.50 We say that $F_p \to F$ in $\mathcal{E}'(\Omega)$ if $F_p \to F$ in $\mathcal{D}'(\Omega)$. That is to say, $F_p(\phi) \to F(\phi)$ for any $\phi \in \mathcal{D}(\Omega)$ and there is a compact $K \Subset \Omega$ such that $\mathrm{supp}(F_p), \mathrm{supp}(F) \subset K$.

1.6 Differentiation of Distributions

Take a function $f \in C^\infty(\Omega)$. Then, for any multi-index $\boldsymbol{\beta}$, define

$$g := \partial^{\boldsymbol{\beta}} f = \frac{\partial^{|\boldsymbol{\beta}|} f}{\partial x_1^{\beta_1} \cdots \partial x_n^{\beta_n}}$$

Obviously $g \in C^\infty(\Omega)$ as well. So

$$\begin{aligned}
F_g(\phi) &= \int_\Omega g(x)\phi(x)\, dx \\
&= \int_\Omega \partial^{\boldsymbol{\beta}} f(x)\phi(x)\, dx \\
&= (-1)^{|\boldsymbol{\beta}|} \int_\Omega f(x) \underbrace{\partial^{\boldsymbol{\beta}}\phi(x)}_{\in \mathcal{D}(\Omega)}\, dx && (1.11) \\
&= (-1)^{|\boldsymbol{\beta}|} F_f(\partial^{\boldsymbol{\beta}}\phi) && (1.12)
\end{aligned}$$

where line 1.11 has been achieved by integrating by parts $|\boldsymbol{\beta}|$ times. You can think of this process like repeating this:

$$\underbrace{\left[\phi(x)\partial^{\boldsymbol{\beta}-1} f(x)\right]}_{\substack{\text{This disappears since} \\ f,\phi = 0 \text{ at the boundaries}}} - \int_\Omega \partial\phi(x)\, \partial^{\boldsymbol{\beta}-1} f(x)\, dx$$

Therefore:

Definition 1.51 For $F \in \mathcal{D}'(\Omega)$ and a multi-index $\boldsymbol{\beta}$, then $\partial^{\boldsymbol{\beta}} F \in \mathcal{D}'(\Omega)$ is the distribution of the form

$$\partial^{\boldsymbol{\beta}} F(\phi) = (-1)^{|\boldsymbol{\beta}|} F(\partial^{\boldsymbol{\beta}}\phi)$$

and is called the $\boldsymbol{\beta}$-derivative of the distribution F.

Let us just show that this definition is OK.

First, since ϕ has infinitely many derivatives, then $\partial^{\boldsymbol{\beta}}\phi$ also has infinitely many derivatives. Moreover, since $\phi = 0$ on the open set $\mathcal{N}(\phi)$, then $(-1)^{|\boldsymbol{\beta}|}\partial^{\beta}\phi = 0$ on $\mathcal{N}(\partial^{\boldsymbol{\beta}}\phi)$, which is larger than $\mathcal{N}(\phi)$ and

$$\operatorname{supp}(\partial^{\beta}\phi) \subset \operatorname{supp}(\phi).$$

This implies that $\operatorname{supp}(\partial^{\boldsymbol{\beta}}\phi)$ is a compact in Ω. Therefore, line 1.12 above is well-defined.

Now, to show that (1.12) is a distribution, we should prove the relations of linearity (1.5) and continuity (1.6).

1. Linearity:

$$\begin{aligned}\partial^{\boldsymbol{\beta}} F(\lambda_1\phi_1 + \lambda_2\phi_2) &= (-1)^{|\boldsymbol{\beta}|} F(\partial^{\boldsymbol{\beta}}[\lambda_1\phi_1 + \lambda_2\phi_2]) \\ &= (-1)^{|\boldsymbol{\beta}|} F(\lambda_1\partial^{\boldsymbol{\beta}}\phi_1 + \lambda_2\partial^{\boldsymbol{\beta}}\phi_2) \\ &= (-1)^{|\boldsymbol{\beta}|}\lambda_1 F(\partial^{\boldsymbol{\beta}}\phi_1) + (-1)^{|\boldsymbol{\beta}|}\lambda_2 F(\partial^{\boldsymbol{\beta}}\phi_2) \\ &= \lambda_1\partial^{\boldsymbol{\beta}} F(\phi_1) + \lambda_2\partial^{\boldsymbol{\beta}} F(\phi_2).\end{aligned}$$

2. Continuity: by the definition of convergence, if $\phi_p \to \phi$, then

$$\partial^{\boldsymbol{\beta}}\phi_p \to \partial^{\boldsymbol{\beta}}\phi$$

   Thus,

$$\partial^{\boldsymbol{\beta}} F(\phi_p) = (-1)^{|\beta|} F(\partial^{\boldsymbol{\beta}}\phi_p) \to (-1)^{|\beta|} F(\partial^{\boldsymbol{\beta}}\phi) = \partial^{\boldsymbol{\beta}} F(\phi).$$

Thus, for distributions associated with smooth functions, the derivatives of these distributions are associated with the derivatives of the corresponding functions. This observation shows that our definition 1.51 is a natural extension of the notion of differentiation from functions to distributions.

Problem 1.52 The Heaviside step function, $\Theta(x)$ or $H(x)$, for $x \in \mathbb{R}$, is given by

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}.$$

Show that $H' = \delta$, where H represents the distribution derived from the function $H(x)$.

Proof:

$$\begin{aligned}H'(\phi) &= (-1)H(\phi') \\ &= -\int_0^\infty \phi'(x)\,dx \\ &= \phi(0) \qquad [\text{since } \phi(\infty) = 0 \text{ because of compact support}] \\ &= \delta(\phi)\end{aligned}$$

□

Problem 1.53 The sign (or *signum*) function, $\operatorname{sgn}(x)$, for $x \in \mathbb{R}$ is given by

$$\begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}.$$

Show that $\operatorname{sgn}' = 2\delta$.

Proof:

$$
\begin{aligned}
\operatorname{sgn}'(\phi) &= (-1)\operatorname{sgn}(\phi') \\
&= -\int_{-\infty}^{\infty} \operatorname{sgn}(x)\phi'(x)\,dx \\
&= -\left(\int_{-\infty}^{0} \operatorname{sgn}(x)\phi'(x)\,dx + \int_{0}^{\infty} \operatorname{sgn}(x)\phi'(x)\,dx\right) \\
&= -\left(-\int_{-\infty}^{0} \phi'(x)\,dx + \int_{0}^{\infty} \phi'(x)\,dx\right) \\
&= -\left(-\left[\phi(x)\right]_{-\infty}^{0} + \left[\phi(x)\right]_{0}^{\infty}\right) \\
&= -\left(-\left(\phi(0)-\phi(-\infty)\right) + \left(\phi(\infty)-\phi(0)\right)\right) \\
&= 2\phi(0) \qquad [\text{since } \phi(\infty)=\phi(-\infty)=0 \text{ because of compact support}] \\
&= 2\delta(\phi)
\end{aligned}
$$

□

Problem 1.54 Show that, for any multi-index $\boldsymbol{\alpha}$ and any $y \in \Omega$,

$$\partial^{\boldsymbol{\alpha}}\delta_y(\phi) = (-1)^{|\boldsymbol{\alpha}|}(\partial^{\boldsymbol{\alpha}}\phi)(y).$$

Solution

$$
\begin{aligned}
\partial^{\boldsymbol{\alpha}}\delta_y(\phi) &= (-1)^{|\boldsymbol{\alpha}|}\delta_y(\partial^{\boldsymbol{\alpha}}\phi) \\
&= (-1)^{|\boldsymbol{\alpha}|}(\partial^{\boldsymbol{\alpha}}\phi)(y)
\end{aligned}
$$

✓

Note, in a matter unrelated to the problem above, that if $F \in \mathcal{D}'$ is a distribution and $\psi \in C^{\infty}$, then

$$(\psi F)(\phi) = F(\psi\phi), \quad \psi \in \mathcal{D}(\Omega)$$

1.7 Convolution

An important operation acting on distributions is *convolution*. To define it, we start, as usual, with continuous functions.

Definition 1.55 Let

$$f \in C_0(\mathbb{R}^n)$$
$$g \in C(\mathbb{R}^n)$$

i.e. f has compact support, $\mathrm{supp}(f) \subset K \Subset \mathbb{R}^n$. Then define

$$\begin{aligned} h(x) = (f * g)(x) &:= \int_{\mathbb{R}^n} f(x-y)g(y)\, dy \\ &= \int_{\mathbb{R}^n} f(y)g(x-y)dy \\ &= (g * f)(x), \end{aligned}$$

which is called the *convolution* of f and g. Clearly, $f * g \in C(\mathbb{R}^n)$.

Consider the corresponding distribution F_h,

$$\begin{aligned} F_h(\phi) = F_{f*g}(\phi) &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)g(y)dy \right] \phi(x)dx \\ &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x-y)\phi(x)dx \right] g(y)dy \\ (z := x-y) \implies &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(z)\phi(z+y)dz \right] g(y)dy \\ (x := z) \implies &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x)\phi(x+y)dx \right] g(y)dy \\ (\phi^y(x) := \phi(x+y)) \implies &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(x)\phi^y(x)dx \right] g(y)dy \\ &= \int_{\mathbb{R}^n} \underbrace{F_f(\phi^y)}_{\Phi(y)} g(y)dy \\ &= F_g(\Phi) \\ &= F_g(F_f(\phi^y)) \end{aligned}$$

Here, for any $y \in \mathbb{R}^n$, we denoted by ϕ^y the translation of ϕ,

$$\phi^y(x) = \phi(x+y).$$

Therefore,

$$F_{f*g}(\phi) = F_g\left(F_f(\phi^y)\right). \tag{1.13}$$

It turns out that formula (1.13) makes sense for any $F \in \mathcal{E}'(\mathbb{R}^n)$, $G \in \mathcal{D}'(\mathbb{R}^n)$.

Definition 1.56 For $F \in \mathcal{E}'(\mathbb{R}^n), G \in \mathcal{D}'(\mathbb{R}^n)$, we can define the *convolution*, $H = F * G \in \mathcal{D}'(\Omega)$ by

$$(F * G)(\phi) = G(F(\phi^y)).$$

We will now show that this definition is (1) well-defined, (2) linear and (3) continuous.

1. Well-defined:

   (a) Support:
   Since $F \in \mathcal{E}'(\mathbb{R}^n)$,

   $$\implies \operatorname{supp}(F) \Subset \mathbb{R}^n$$

   i.e. $\exists b > 0$ such that $\mathbb{R}^n \setminus B_b(\mathbf{0}) \subset \mathcal{N}(F)$, where $B_b(\mathbf{0})$ is a ball of radius b centred at $\mathbf{0}$.
   Since $\phi \in \mathcal{D}(\mathbb{R}^n)$,

   $$\implies \exists a > 0 \text{ such that } \operatorname{supp}(\phi^y) \subset B_a(y)$$

   i.e. $\operatorname{supp}(\phi) \subset B_a(\mathbf{0})$
   Thus if $|y| > b + a$,

   $$\operatorname{supp}(F) \cap \operatorname{supp}(\phi^y) = \emptyset$$
   $$\implies F(\phi^y) = 0$$

   Denote $\Phi(y) = F(\phi^y)$. Then $\operatorname{supp}(\Phi) \subset B_{b+a}(\mathbf{0})$.

   (b) Is $\Phi(y) \in C^\infty(\mathbb{R}^n)$?
   We will show that Φ is infinitely differentiable by proving that

   $$\partial^{\boldsymbol{\alpha}} \Phi = F(\partial^{\boldsymbol{\alpha}} \phi^y),$$

   using induction on $|\boldsymbol{\alpha}|$.
   Assume

   $$\partial^{\boldsymbol{\alpha}} \Phi = F(\partial^{\boldsymbol{\alpha}} \phi^y), \quad |\boldsymbol{\alpha}| \leq m$$

   and let us show that

   $$\partial^{\boldsymbol{\beta}} \Phi = F(\partial^{\boldsymbol{\beta}} \phi^y), \quad |\boldsymbol{\beta}| = m + 1.$$

   Let $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{e}_i$, where

   $$\boldsymbol{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$$

   with the 1 in the ith place.
   By our inductive hypothesis,

   $$\partial_i \partial^{\boldsymbol{\alpha}} \Phi = \partial_i F(\partial^{\boldsymbol{\alpha}} \phi^y)$$

   but we need to prove that the right hand side exists. In order to do this, we need to show that:

   $$\psi_s \xrightarrow[s \to 0]{} \psi \text{ in } D(\Omega)$$

Where

$$\psi_s(y) = \frac{\partial^{\boldsymbol{\alpha}} \phi^{y+se_i} - \partial^{\boldsymbol{\alpha}} \phi^y}{s}\ ,\ \psi(y) = \partial^{\boldsymbol{\beta}} \phi^y$$

So we need, for any multi-index $\boldsymbol{\gamma}$:

$$\partial^{\boldsymbol{\gamma}} \psi_s \xrightarrow[s \to 0]{} \partial^{\boldsymbol{\gamma}} \psi$$

For $|\boldsymbol{\gamma}| = 0$, we have:

$$\begin{aligned}
\psi_s(y) = \frac{\partial^{\boldsymbol{\alpha}} \phi^{y+se_i} - \partial^{\boldsymbol{\alpha}} \phi^y}{s} =& \frac{\partial^{\boldsymbol{\alpha}} \phi(x+y+se_i) - \partial^{\boldsymbol{\alpha}} \phi(x+y)}{s} \\
\xrightarrow[s \to 0]{}& \partial^{\boldsymbol{\alpha}+\boldsymbol{e}_i} \phi(x+y) \\
=& \partial^{\boldsymbol{\beta}} \phi^y = \psi(y)
\end{aligned}$$

Now, for an arbitrary multi-index $\boldsymbol{\gamma}$, we want to show that:

$$\partial^{\boldsymbol{\gamma}} \psi_s = \partial^{\boldsymbol{\gamma}} \left[\frac{\partial^{\boldsymbol{\alpha}} \phi(x+y+se_i) - \partial^{\boldsymbol{\alpha}} \phi(x+y)}{s} \right] \xrightarrow[s \to 0]{} \partial^{\boldsymbol{\gamma}+\boldsymbol{\beta}} \phi(x+y) = \partial^{\boldsymbol{\gamma}} \psi$$

Well,

$$\begin{aligned}
\frac{\partial^{\boldsymbol{\gamma}+\boldsymbol{\alpha}} \phi(x+y+se_i) - \partial^{\boldsymbol{\gamma}+\boldsymbol{\alpha}} \phi(x+y)}{s} &= \partial_i \partial^{\boldsymbol{\gamma}+\boldsymbol{\alpha}} \phi(x+y) \\
&= \partial^{\boldsymbol{\gamma}+\boldsymbol{\alpha}+\boldsymbol{e}_i} \phi(x+y) \\
&= \partial^{\boldsymbol{\gamma}+\boldsymbol{\beta}} \phi(x+y)
\end{aligned}$$

So $\psi_s \xrightarrow[s \to 0]{} \psi$ in $D(\Omega)$, which means that $F(\psi_s) \xrightarrow[s \to 0]{} F(\psi)$, now:

$$\begin{aligned}
\partial^{\boldsymbol{\beta}} \Phi = \partial_i F(\partial^{\boldsymbol{\alpha}} \phi^y) &= \lim_{s \to 0} \frac{F(\partial^{\boldsymbol{\alpha}} \phi^{y+se_i}) - F(\partial^{\boldsymbol{\alpha}} \phi^y)}{s} \\
&= \lim_{s \to 0} F\left(\frac{\partial^{\boldsymbol{\alpha}} \phi^{y+se_i} - \partial^{\boldsymbol{\alpha}} \phi^y}{s} \right) \\
&= \lim_{s \to 0} F(\psi_s) \\
&= F(\psi) = F(\partial^{\boldsymbol{\beta}} \phi^y)
\end{aligned}$$

So then, by induction, for any multi-index $\boldsymbol{\alpha}$

$$\partial^{\boldsymbol{\alpha}} F(\phi^y) = F(\partial^{\boldsymbol{\alpha}} \phi^y) \tag{1.14}$$

Therefore $F(\phi^y) \in C^\infty(\mathbb{R}^n)$. Thus $G(F(\phi^y))$ is well defined.

2. Linearity: Clearly, the functional

$$\phi \mapsto G\left(F(\phi^y)\right)$$

is linear.

3. Continuity: This is also a pain, but it's the same type of pain.

   Let $\phi_k \to \phi$ in $\mathcal{D}$. We need

$$(F * G)\phi_k \to (F * G)(\phi)$$

   i.e.

$$G(F(\phi_k^y)) \xrightarrow[k\to\infty]{} G(F(\phi^y)).$$

   It is sufficient to show that

$$F(\phi_k^y) = \Phi_k(y) \to \Phi(y) = F(\phi^y) \qquad \text{in } \mathcal{D}.$$

   So let's check support and differentials

   (a) Check support: Let us show that $\mathrm{supp}(\Phi_k), \mathrm{supp}(\Phi)$ lie in the same ball.
   As $\phi_k \to \phi$ in $\mathcal{D}$, $\exists a > 0$ such that $\mathrm{supp}(\phi_k), \mathrm{supp}(\phi) \subset B_a$.
   Then $\mathrm{supp}(\Phi_k), \mathrm{supp}(\Phi) \subset B_{b+a}$ (where $\mathrm{supp}(F) \subset B_b$).

   (b) Show $\partial^{\boldsymbol{\alpha}}\Phi_k(y) \to \partial^{\boldsymbol{\alpha}}\Phi(y)$
   i.e. show $F(\partial^{\boldsymbol{\alpha}}\phi_k^y) \to F(\partial^{\boldsymbol{\alpha}}\phi^y)$
   Since $\phi_k \to \phi$ in $\mathcal{D}(\mathbb{R}^n)$, this implies that $\phi_k^y \to \phi^y$ in $\mathcal{D}(\mathbb{R}^n)$.
   Then by Proposition 1.7, $\partial^{\boldsymbol{\alpha}}\phi_k^y \to \partial^{\boldsymbol{\alpha}}\phi^y$ in $\mathcal{D}(\Omega) \quad \forall \boldsymbol{\alpha}$.
   Therefore $F(\partial^{\boldsymbol{\alpha}}\phi_k^y) \to F(\partial^{\boldsymbol{\alpha}}\phi^y)$

   Congratulations!!! You did it!

Proposition 1.57 Convolution is a continuous map from $\mathcal{E}'(\Omega) \times \mathcal{D}'(\Omega)$ to $\mathcal{D}'(\Omega)$, i.e. if $F_p \to F$ in $\mathcal{E}'(\Omega)$ and $G_p \to G$ in $\mathcal{D}'(\Omega)$, then

$$F_p * G_p \to F * G.$$

We won't prove this as it's too difficult.

Problem 1.58 $*$ If $F_p \to F$ in $\mathcal{E}'(\Omega)$ and $G \in \mathcal{D}'(\Omega)$, show that

$$F_p * G \to F * G$$

Solution

$$(F_p * G)(\phi) = G(F_p(\phi^y))$$

G is continuous, and $F_p \to F$ in $\mathcal{E}'(\Omega)$, so $F_p(\phi^y) \to F(\phi^y)$, and hence

$$\begin{aligned} &\xrightarrow[p\to\infty]{} G(F(\phi^y)) \\ &= (F * G)(\phi) \end{aligned}$$

✓

Lemma 1.59

$$F * G = G * F. \tag{1.15}$$

Remark 1.60 We prove (1.15) later, although it's obvious for distributions from the functions: if $F = F_f, G = F_g$,

$$G * F = F_h \qquad h(x) = \int g(x-y)f(y)\,dy$$

$$F * G = F_{\widehat{h}} \qquad \widehat{h}(x) = \int f(x-y)g(y)\,dy$$

which are clearly equal.

Although the convolution of distribution is usually defined only for $\Omega = \mathbb{R}^n$, in some cases it may be defined for general Ω. In particular, when $0 \in \Omega$, $F * G$ is defined for $F = \partial^{\boldsymbol{\alpha}}\delta$, where $\boldsymbol{\alpha}$ is arbitrary.

Problem 1.61 Show that

$$(\partial^{\boldsymbol{\alpha}}\delta) * G = \partial^{\boldsymbol{\alpha}} G.$$

Solution

$$\begin{aligned}
[(\partial^{\boldsymbol{\alpha}}\delta) * G](\phi) &= G[(\partial^{\boldsymbol{\alpha}}\delta)(\phi^y)] \\
&= G[(-1)^{|\boldsymbol{\alpha}|}\delta(\partial^{\boldsymbol{\alpha}}\phi^y)] \\
&= (-1)^{|\boldsymbol{\alpha}|}G[\delta(\partial^{\boldsymbol{\alpha}}\phi^y)] \\
&= (-1)^{|\boldsymbol{\alpha}|}G[\partial^{\boldsymbol{\alpha}}\phi] \\
&= (\partial^{\boldsymbol{\alpha}} G)(\phi)
\end{aligned}$$

✓

Now we will show that we can move about derivatives:

Proposition 1.62 Let $F * G \in \mathcal{D}'(\Omega)$ be the convolution of two distributions. Then, for any multi-index $\boldsymbol{\alpha}$,

$$\partial^{\boldsymbol{\alpha}}(F * G) = (\partial^{\boldsymbol{\alpha}} F) * G = F * (\partial^{\boldsymbol{\alpha}} G).$$

Proof: By the definition of differentiation (definition 1.51),

$$\begin{aligned}
\partial^{\boldsymbol{\alpha}}(F * G)(\phi) &= (-1)^{|\boldsymbol{\alpha}|}(F * G)(\partial^{\boldsymbol{\alpha}}\phi) \\
&= (-1)^{|\boldsymbol{\alpha}|}G(F(\partial^{\boldsymbol{\alpha}}\phi^y))
\end{aligned}$$

by the definition of convolution (definition 1.56).

To proceed, notice that

$$\partial^{\boldsymbol{\alpha}}\phi^y(x) = \frac{\partial^{\boldsymbol{\alpha}}\phi(x+y)}{\partial x^{\boldsymbol{\alpha}}} = \frac{\partial^{\boldsymbol{\alpha}}\phi(x+y)}{\partial y^{\boldsymbol{\alpha}}}$$

Therefore

$$\begin{aligned} F(\partial^{\boldsymbol{\alpha}}\phi^y) &= F\left(\frac{\partial^{\boldsymbol{\alpha}}\phi^y}{\partial y^{\boldsymbol{\alpha}}}\right) \\ &= \frac{\partial^{\boldsymbol{\alpha}}}{\partial y^{\boldsymbol{\alpha}}}(F(\phi^y)) \end{aligned}$$

by equation 1.14. So, proceeding,

$$\begin{aligned} \partial^{\boldsymbol{\alpha}}(F*G)(\phi) &= (-1)^{|\boldsymbol{\alpha}|}(F*G)(\partial_x^{\boldsymbol{\alpha}}\phi) \\ &= (-1)^{|\boldsymbol{\alpha}|}G(F(\partial_x^{\boldsymbol{\alpha}}\phi^y)) \\ &= (-1)^{|\boldsymbol{\alpha}|}G(F(\partial_y^{\boldsymbol{\alpha}}\phi^y)) \\ &= (-1)^{|\boldsymbol{\alpha}|}G(\partial_y^{\boldsymbol{\alpha}}F(\phi^y)) \\ &= \partial_y^{\boldsymbol{\alpha}}G(F(\phi^y)) \\ &= (F*\partial^{\boldsymbol{\alpha}}G)(\phi) \end{aligned}$$

Note the importance of which variable we differentiate with respect to. And since $F*G = G*F$, swapping F and G holds. □

Corollary 1.63

$$\partial^{\boldsymbol{\beta}+\boldsymbol{\gamma}}(F*G) = \partial^{\boldsymbol{\beta}}(F)*\partial^{\boldsymbol{\gamma}}(G)$$

1.8 Density

Now for a little notation

Notation 1.64

$$\mathcal{F} : C(\Omega) \mapsto \mathcal{D}'(\Omega) \qquad \mathcal{F}(f) = F_f$$

As we know $\mathcal{F}(C(\Omega)) \neq \mathcal{D}'(\Omega)$. However, we have the following density result

Theorem 1.65

$$\mathrm{cl}\left(\mathcal{F}(C_0(\Omega))\right) = \mathcal{D}'(\Omega),$$

i.e. for any $F \in \mathcal{D}'(\Omega)$, $\Omega \subset \mathbb{R}^n$, there is a sequence $f_p \in \mathcal{D}(\Omega)$ such that

$$F_{f_p} \xrightarrow[p\to\infty]{} F \quad \text{in } \mathcal{D}'(\Omega) \tag{1.16}$$

Sketch of proof: We will prove this theorem in 2 parts:

1. We will show that the distributions with compact support, $\mathcal{E}'(\Omega)$ are dense in the space of all distributions $\mathcal{D}'(\Omega)$, i.e. we show that $\mathrm{cl}(\mathcal{E}'(\Omega)) = \mathcal{D}'(\Omega)$.

2. We will show that the space of distributions of compact support due to continuous functions with compact support is dense in the space of distributions with compact support, i.e. we show that $\mathrm{cl}\left(\mathcal{F}(C_0(\Omega))\right) = \mathcal{E}'(\Omega)$. We will do this in three steps:

   (a) For F, a distribution with compact support, we will construct a sequence of distributions converging to F in $\mathcal{D}'(\Omega)$.

   (b) Then we will show that in fact these distributions all have compact support, lying in the same compact, so that they converge to F in $\mathcal{E}'(\Omega)$.

   (c) Finally we will show that they are in fact all distributions due to continuous functions with compact support.

Once we have proven these facts we see that $\mathcal{F}(C_0(\Omega))$ is dense in $\mathcal{E}'(\Omega)$ which is, in turn, dense in $\mathcal{D}'(\Omega)$. Recall, by the nature of the closure operation, this then means that $\mathrm{cl}(\mathcal{F}(C_0(\Omega))) = \mathcal{D}'(\Omega)$. (Since $\mathcal{D}'(\Omega) = \mathrm{cl}(\mathcal{E}'(\Omega)) = \mathrm{cl}(\mathrm{cl}(C_0(\Omega))) = \mathrm{cl}(C_0(\Omega))$ as the closure of a closed set is itself).

1. We will show that any distribution can be approximated by distributions with compact support, i.e.

$$\forall F \in \mathcal{D}'(\Omega),\ \exists F_p \in \mathcal{E}'(\Omega) \text{ s.t. } F_p \to F$$

First, let K_p, $p = 1, 2, \ldots$, be a sequence of compacts exhausting Ω, i.e.

$$K_p \Subset \Omega, \quad K_p \subset \mathrm{int}(K_{p+1}), \quad \bigcup_{p=1}^{\infty} K_p = \Omega.$$

In fact we can define them as

$$K_p = \left\{ x \in \Omega : |x| \leq p, \quad \mathrm{dist}(x, \partial\Omega) \geq \frac{1}{p} \right\}$$

Now let $\chi_p \in \mathcal{D}(\Omega)$ be functions with the properties described by Proposition 1.7, where $K_1 = K_p$ and $K_2 = K_{p+1}$, i.e.

$$\chi_p(x) = \begin{cases} 1 & \text{if } x \in K_p \\ 0 & \text{if } x \in \Omega \setminus K_{p+1} \end{cases}$$

Let $F \in \mathcal{D}'(\Omega)$, then define F_p by:

$$F_p = \chi_p F$$

$$F_p(\phi) = \chi_p F(\phi) = F(\chi_p \phi).$$

Note that

$$\mathrm{supp}(F_p) \subset K_{p+1}. \tag{1.17}$$

Indeed, if $\phi \in \mathcal{D}(\Omega)$ and $\mathrm{supp}(\phi) \cap K_{p+1} = \emptyset$, then

$$F_p(\phi) = \chi_p F(\phi) = F(\underbrace{\chi_p \phi}_{0}) = 0.$$

Also, $F_p \to F$. To prove this, we want to show that

$$F(\chi_p \phi) = F_p(\phi) \xrightarrow[p\to\infty]{} F(\phi) \qquad \forall \phi \in \mathcal{D}(\Omega)$$

which clearly it does. Indeed, for large p, $K \subset K_p$

Proof:

$$\operatorname{dist}(K, \partial\Omega) = d > 0$$

since $K \Subset \Omega$. As K is a compact, $K \subset B_R$, for some radius $R > 0$.

Hence for $p > \max(R, \frac{1}{d})$, $K \subset K_p$. □

2. (a) Any distribution with compact support can be approximated by distributions of functions with compact support, i.e.

$$\forall F \in \mathcal{E}'(\Omega), \quad \exists f_p \in \mathcal{D}(\Omega) \text{ s.t. } F_{f_p} \to F \in \mathcal{E}'(\Omega)$$

Let us introduce a function $\chi(x) \in \mathcal{D}(\mathbb{R}^n)$ such that

i. $\operatorname{supp}(\chi) \subset B_1(0)$

ii. $\chi(x) = \begin{cases} 1 & \text{if } x \in B_{1/2}(0) \\ 0 & \text{if } x \in \Omega \setminus B_1(0) \end{cases}$

iii. $\displaystyle\int_{\mathbb{R}^n} \chi(x) dx = 1$

iv. $\chi = \chi(|x|)$ (i.e. χ is a radial function)

The existence of such χ follows easily from proposition 1.7.

Now consider, for sufficiently small $\varepsilon > 0$, a function $\chi^\varepsilon(x) \in \mathcal{D}(\mathbb{R}^n)$ such that

i. $\operatorname{supp}(\chi^\varepsilon) \subset B_\varepsilon(0)$

ii. $\chi^\varepsilon(x) = \begin{cases} 1 & \text{if } x \in B_{\varepsilon/2}(0) \\ 0 & \text{if } x \in \Omega \setminus B_\varepsilon(0) \end{cases}$

iii. $\displaystyle\int_{\mathbb{R}^n} \chi^\varepsilon(x) dx = 1$

iv. $\chi^\varepsilon(x) = \dfrac{1}{\varepsilon^n}\chi\left(\dfrac{x}{\varepsilon}\right)$

Now define

$$F^\varepsilon := \chi^\varepsilon * F \in \mathcal{E}'(\Omega) \tag{1.18}$$

then for $\phi \in \mathcal{D}(\Omega)$, by the definition of convolution,

$$\begin{aligned} F^\varepsilon(\phi) = (\chi^\varepsilon * F)(\phi) &= F(\chi^\varepsilon(\phi^y)) \\ &= F_{\chi^\varepsilon}(\phi^y) \\ &:= \int_{\mathbb{R}^n} \chi^\varepsilon(x)\phi(x+y)\, dx \quad \in \mathcal{D}(\mathbb{R}^n) \end{aligned}$$

and we'll show now that

$$\chi^\varepsilon(\phi^y) \to \phi(y) \quad \in \mathcal{D}(\Omega).$$

First note that

$$\begin{aligned}
\chi^\varepsilon(\phi^y) &= \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\phi(x+y)\,dx \\
&= \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\left[\phi(y)+\phi(x+y)-\phi(y)\right]\,dx \\
&= \phi(y)\int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\,dx + \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\left[\phi(x+y)-\phi(y)\right]\,dx \\
&= \phi(y) + \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\left[\phi(x+y)-\phi(y)\right]\,dx
\end{aligned}$$

As $\int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\,dx = 1.$

So we can write

$$\chi^\varepsilon(\phi^y) = \phi(y) + \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)[\phi(x+y)-\phi(y)]\,dx$$

Observe that

$$\sup_{|x|\le\varepsilon} |\phi(x+y)-\phi(y)| \xrightarrow[\varepsilon\to 0]{} 0 \qquad \forall y\in\Omega$$

i.e. pointwise in Ω, which means that $\phi(x+y)-\phi(y) \xrightarrow[\varepsilon\to 0]{} 0$ uniformly on any $K \Subset \Omega$. Hence

$$\begin{aligned}
\psi^\epsilon(y) = \chi^\varepsilon(\phi^y) = \phi(y) + \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\underbrace{[\phi(x+y)-\phi(y)]}_{\to 0\ \forall y\in\Omega}dx \\
\xrightarrow[\varepsilon\to 0]{} \phi(y) \qquad \text{uniformly on any compact}
\end{aligned}$$

i.e.

$$\psi^\epsilon(y) = \chi^\varepsilon(\phi^y) \to \phi(y) \qquad \forall y\in\Omega$$

and uniformly on any compact.
And the same goes for the derivative. If we replace $\chi^\varepsilon(\phi^y)$ by $\partial^{\boldsymbol{\alpha}}\chi^\varepsilon(\phi^y)$, we use the same trick and

$$\partial^{\boldsymbol{\alpha}}\chi^\varepsilon(\phi^y) \to \partial^{\boldsymbol{\alpha}}\phi(y) \qquad \forall y\in\Omega, \boldsymbol{\alpha}$$

This means that $\chi^\varepsilon(\phi^y) = \psi^\varepsilon \to \phi$ in $\mathcal{E}(\Omega) = C^\infty(\Omega)$.

So $(\chi^\varepsilon * F)(\phi) = F(\chi^\varepsilon(\phi^y)) \xrightarrow[\varepsilon\to 0]{} F(\phi)$ since $F\in\mathcal{E}'(\Omega)$ and $\chi^\varepsilon(\phi^y)\to\phi$ in $\mathcal{E}(\Omega)$

So $\chi^\varepsilon * F \to F$ in $\mathcal{D}'(\Omega)$.

(b) Now we show that $\chi^\varepsilon * F \in \mathcal{E}'(\Omega)$, i.e. We show that $\chi^\varepsilon * F$ has compact support. First, we find $\text{supp}(\chi^\varepsilon(\phi^y))$.

Suppose $d(y, \text{supp}(\phi)) > \varepsilon$ then $d(x+y, \text{supp}(\phi)) > 0$ for $|x| \le \varepsilon$, which means that $x + y \notin \text{supp}(\phi)$, so:

$$\chi^\varepsilon(\phi^y) = \int\limits_{|x|\le\varepsilon} \chi^\varepsilon(x)\phi(x+y)\,dx = 0 \quad \forall y \text{ s.t. } d(y, \text{supp}(\phi)) > \varepsilon$$

So, $\text{supp}(\chi^\varepsilon(\phi^y)) \subset \{y : d(y, \text{supp}(\phi)) \le \varepsilon\}$, i.e. $\text{supp}(\chi^\varepsilon(\phi^y))$ lies in an ε-vicinity of $\text{supp}(\phi)$.

Now we claim that $\text{supp}(\chi^\varepsilon * F) \subset \{x : d(x, \text{supp}(F)) \le \varepsilon\} = \text{supp}(F)^\varepsilon$, i.e $\text{supp}(\chi^\varepsilon * F)$ lies in an ε-vicinity of $\text{supp}(F)$.

Let ϕ be such that:

$$\text{supp}(\phi) \cap \text{supp}(F)^\varepsilon = \emptyset$$

Then:

$$d(\text{supp}(\phi), \text{supp}(F)) > \varepsilon$$

i.e. the ε-vicinity of $\text{supp}(\phi)$ lies outside of $\text{supp}(F)$. Thus:

$$\text{supp}(\chi^\varepsilon(\phi^y)) \cap \text{supp}(F) = \emptyset$$

So:

$$(\chi^\varepsilon * F)(\phi) = F(\chi^\varepsilon(\phi^y)) = 0 \quad \forall\phi \text{ s.t. } \text{supp}(\phi) \cap \text{supp}(F)^\varepsilon = \emptyset$$

Which means that:

$$\text{supp}(\chi^\varepsilon * F) \subset \text{supp}(F)^\varepsilon$$

Now take $\varepsilon_0 = \frac{1}{2}d(\text{supp}(F), \partial\Omega)$ so that $\text{supp}(F)^{\varepsilon_0}$ is compact in Ω. Then:

$$\forall\varepsilon \le \varepsilon_0 \quad \text{supp}(\chi^\varepsilon * F) \subset \text{supp}(F)^\varepsilon \subset \text{supp}(F)^{\varepsilon_0} \Subset \Omega$$

So that $\text{supp}(\chi^\varepsilon * F) \Subset \Omega$ which means that:

$$\chi^\varepsilon * F \in \mathcal{E}'(\Omega)$$

(c) Now we show that, for $\varepsilon < \frac{1}{2}d(\text{supp}(F), \partial\Omega)$, $\chi^\varepsilon * F \in C_0^\infty(\Omega)$

In other words, there are functions $h^\varepsilon \in C_0^\infty(\Omega)$ s.t. $\chi^\varepsilon * F = F_{h^\varepsilon}$

We approach this problem as we always do; in order to construct h^ε we first look at the case when $F = F_f$, then:

$$\begin{aligned}
(\chi^\varepsilon * F_f)(\phi) &= F_f(\chi^\varepsilon(\phi^y)) \\
&= \int f(y)\left(\int \chi^\varepsilon(x)\phi(x+y)dx\right)dy \\
&= \underbrace{\int\left(\int f(y)\chi^\varepsilon(z-y)\phi(z)dz\right)dy}_{z=x+y,dz=dx} \\
&= \int\left(\int f(y)\chi^\varepsilon(z-y)\phi(z)dy\right)dz \\
&= \int \phi(z)\left(\int \chi^\varepsilon(z-y)f(y)dy\right)dz \\
&= F_{h^\varepsilon}(\phi) \\
\text{Where:} \qquad h^\varepsilon(z) &= \int f(y)\chi^\varepsilon(y-z)dy \\
&= F_f(\chi^{\varepsilon,-z})
\end{aligned}$$

Where we have denoted $\chi^{\varepsilon,-z}(x) = \chi^\varepsilon(x-z)$. Now we generalise this to the case when F is any distribution, by trying to show that:

$$\text{If } h^\varepsilon = F(\chi^{\varepsilon,-z}) \in C_0^\infty(\Omega), \text{ then } \chi^\varepsilon * F = F_{h^\varepsilon}$$

Now observe that:

$$\begin{aligned}
\chi^\varepsilon(\phi^y) = \int \chi^\varepsilon(x)\phi(x+y)dx &= \int \chi^\varepsilon(z-y)\phi(z)dz \\
&= \lim_{\delta\to 0}\sum_i \chi^\varepsilon(z_i-y)\phi(z_i)\delta^n \\
&:= \lim_{\delta\to 0} R_\delta^\varepsilon(y)
\end{aligned}$$

Also that:

$$\begin{aligned}
\partial_y^{\boldsymbol{\alpha}}\chi^\varepsilon(\phi^y) = \partial_y^{\boldsymbol{\alpha}}\int \chi^\varepsilon(z-y)\phi(z)dz &= \int \partial_y^{\boldsymbol{\alpha}}\chi^\varepsilon(z-y)\phi(z)dz \\
&= \lim_{\delta\to 0}\sum_i \partial_y^{\boldsymbol{\alpha}}\chi^\varepsilon(z_i-y)\phi(z_i)\delta^n \\
&= \lim_{\delta\to 0}\partial_y^{\boldsymbol{\alpha}}\sum_i \chi^\varepsilon(z_i-y)\phi(z_i)\delta^n \\
&= \lim_{\delta\to 0}\partial_y^{\boldsymbol{\alpha}} R_\delta^\varepsilon(y)
\end{aligned}$$

Thus

$$\begin{aligned}
\chi^\varepsilon(\phi^y) &= \lim_{\delta\to 0} R^\varepsilon_\delta \text{ in } \mathcal{D}(\Omega) \\
\implies (\chi^\varepsilon * F)(\phi) &= F(\chi^\varepsilon(\phi^y)) \\
&= \lim_{\delta\to 0} F(R^\varepsilon_\delta) \\
&= \lim_{\delta\to 0} F\left(\sum_i \chi^\varepsilon(z_i - y)\phi(z_i)\delta^n\right) \\
&= \lim_{\delta\to 0}\left(\sum_i \left(F(\chi^{\varepsilon,-z_i})\right)\phi(z_i)\delta^n\right) \\
&= \int F(\chi^{\varepsilon,-z})\phi(z)dz \\
&= \int h^\varepsilon(z)\phi(z)dz \qquad h^\varepsilon(z) = F(\chi^{\varepsilon,-z}) \in C_0^\infty(\Omega) \\
&= F_{h^\varepsilon}(\phi)
\end{aligned}$$

These three parts prove that $\mathrm{cl}(\mathcal{F}(C_0^\infty(\Omega)) = \mathcal{E}'(\Omega)$

So we see that $\mathrm{cl}(\mathcal{F}(C_0^\infty(\Omega)) = \mathcal{D}'(\Omega)$ since $\mathrm{cl}(\mathcal{E}'(\Omega)) = \mathcal{D}'(\Omega)$. □

Problem 1.66 $*$ Prove that $\chi^\varepsilon \xrightarrow[\varepsilon\to 0]{} \delta$ (where χ^ε represents the distribution from the function χ^ε defined in part 2a above).

Solution

$$F_{\chi^\varepsilon}(\phi) = \int_{\mathbb{R}^2} \chi^\varepsilon(x)\phi(x)\, dx \qquad \in \mathcal{D}(\mathbb{R}^2)$$

and since $\chi^\varepsilon(x) = 0$ outside of B_ε,

$$\begin{aligned}
&= \int_{B_\varepsilon} \chi^\varepsilon(x)\phi(x)\, dx \\
&= \int_{B_\varepsilon} \chi^\varepsilon(x)\phi(x)\, dx + \phi(0) - \phi(0)
\end{aligned}$$

and since $\phi(0) = \phi(0)\int_{B_\varepsilon} \chi^\varepsilon\, dx = \int_{B_\varepsilon} \chi^\varepsilon \phi(0)\, dx$,

$$= \int_{B_\varepsilon} \chi^\varepsilon(x)\left[\phi(x) - \phi(0)\right] dx + \phi(0)$$

and as $\varepsilon \to 0$, the integral $\to 0$, leaving us with

$$\to \phi(0)$$

Hence $\chi^\varepsilon \to \delta$. ✓

Now we return to a lemma from earlier:

Lemma 1.59

$$F * G = G * F. \tag{1.15}$$

Proof: Let $F_{f_p} \to F$, where $f_p \in \mathcal{D}(\Omega)$ and $G_{g_p} \to G$, where $g_p \in C(\Omega)$. We know, by remark 1.60, that:

$$F_{f_p} * G_{g_p} = G_{g_p} * F_{f_p}$$

and that, by proposition 1.57,

$$F_{f_p} * G_{g_p} \xrightarrow[p\to\infty]{} F * G$$

so

$$F * G = \lim_{p\to\infty}(F_{f_p} * G_{g_p}) = \lim_{p\to\infty}(G_{g_p} * F_{f_p}) = G * F.$$

□

1.9 Integration of Distributions

We complete this section with the notion of integration of distributions with respect to some parameter $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n) \in A \subset \mathbb{R}^n$. Let $F_{\boldsymbol{\tau}} \in \mathcal{D}'(\Omega)$ be continuous with respect to $\boldsymbol{\tau}$, i.e. $F_{\boldsymbol{\tau}} \to F_{\boldsymbol{\tau}_0}$, in $\mathcal{D}'(\Omega)$, when $\boldsymbol{\tau} \to \boldsymbol{\tau}_0$. Recall that this means that,

$$F_{\boldsymbol{\tau}}(\phi) \to F_{\boldsymbol{\tau}_0}(\phi), \quad \text{as } \boldsymbol{\tau} \to \boldsymbol{\tau}_0; \quad \phi \in \mathcal{D}(\Omega).$$

Then, if $g \in C_0(A)$, we define a distribution $G_g \in \mathcal{D}'(\Omega)$ by

$$G_g = \int_A g(\boldsymbol{\tau}) F_{\boldsymbol{\tau}} d\boldsymbol{\tau}.$$

$$G_g(\phi) = \int_A g(\boldsymbol{\tau}) F_{\boldsymbol{\tau}}(\phi) d\boldsymbol{\tau}. \tag{1.19}$$

We can check that if

$$F_{\boldsymbol{\tau}} = F_{f(\boldsymbol{\tau})}$$

$$\text{where } f(\boldsymbol{\tau}) = f(x, \boldsymbol{\tau}) \in C(\Omega \times A)$$

then

$$G_g = F_h$$

$$h(x) = \int_A g(\boldsymbol{\tau}) f(x, \boldsymbol{\tau})\, d\boldsymbol{\tau}.$$

Anyway, now let's consider $A = \Omega = \mathbb{R}^n$.

As a particular case, consider the operation of *translation*. As usual, we start with $f \in C(\Omega)$, which we then continue by 0 on the whole $\mathbb{R}^n$. Let $y \in \mathbb{R}^n$. By translation of f by y we understand the operation

$$T_y : f \mapsto f_y, \quad f_y(x) = f(x-y).$$

Let us consider the corresponding distribution $T_y F_f = F_{f_y}$,

$$\begin{aligned} T_y F_f(\phi) &= \int_\Omega f(x-y)\phi(x)\, dx \\ &= \int_{\mathbb{R}^n} f(x-y)\phi(x)\, dx \\ &= \int_{\mathbb{R}^n} f(x)\phi(x+y)\, dx \\ &= F_f(\phi^y). \end{aligned} \tag{1.20}$$

We would like to use formula (1.20) to define the translation of distributions. However, in general, $\phi^y \notin \mathcal{D}(\Omega)$ but just in $C^\infty(\Omega)$. Thus, assuming that $F \in \mathcal{E}'(\Omega)$ we find the following definition

Definition 1.67 Let $F \in \mathcal{E}'(\Omega), y \in \mathbb{R}^n$. Then,

$$T_y F(\phi) = F(\phi^y), \quad \phi \in \mathcal{D}(\Omega), \tag{1.21}$$

and we continue it by 0 outside Ω.

There are a few different notations for translation, one of Prof. Kurylev's favourites is $T_y F = F(\cdot - y)$, which the editors wish he wouldn't use because it sucks. Another is $T_y F = F_y$.

Example 1.68

$$T_y(\delta) = \delta_y$$

Proof:

$$\begin{aligned} (T_y\delta)(\phi) &= \delta(\phi^y) \\ &= \phi^y(0) \\ &= \phi(y) \\ &= \delta_y(\phi) \end{aligned}$$

□

Observe that the operation of translation $T_y F$ is a distribution in $\mathcal{D}'(\mathbb{R}^n)$, which depends continuously on y:

$$T_y F \to T_{y_0} F \text{ as } y \to y_0.$$

Is it true that

$$(T_y(F))(\phi) = F(\phi^y) \underset{?}{\to} (T_{y_0}F)(\phi) = F(\phi^{y_0})?$$

Yes it is because if $y \to y_0$,

$$\phi^y = \phi(x+y) \to \phi(x+y_0) = \phi^{y_0} \quad \text{in } \mathcal{D}(\mathbb{R}^n)$$

And clearly

$$\partial^{\boldsymbol{\alpha}}\phi(x+y) \xrightarrow[y\to y_0]{} \partial^{\boldsymbol{\alpha}}\phi(x+y_0) \quad \forall \boldsymbol{\alpha} \quad \forall x \in \mathbb{R}^n$$

since the supports lie in the same compact (see the compacts of ϕ^y and ϕ^{y_0} moving towards each other as $y \to y_0$).

Thus, if $g(y) \in C_0(\mathbb{R}^n)$, we can define using equation 1.19

$$G_g = \int\limits_{\mathbb{R}^n} g(y)(T_yF)\ dy$$

$$G_g(\phi) = \int\limits_{\mathbb{R}^n} g(y)F(\phi^y)\ dy$$

Problem 1.69 Show that $G_g = F_g * F$

Problem 1.70 $*$ Let $f(x) = \frac{3}{4}(1-x^2)$ for $|x| \le 1$, $f(x) = 0$ for $|x| > 1$, $x \in \mathbb{R}$. Denote $f_n(x) = nf(nx)$, $n = 1, 2, \ldots$ and let $F_n \in \mathcal{D}'(\mathbb{R})$ be a distribution given by

$$F_n(\phi) = \int_{-\infty}^{\infty} f_n(x)\phi(x)\,dx, \quad \phi \in C_0^\infty(\mathbb{R}).$$

Show that $F_n \to \delta_0$, as $n \to \infty$, in $\mathcal{D}'(\mathbb{R})$.

Solution

$$F_n(\phi) = \int_{-\infty}^{\infty} f_n(x)\phi(x)\ dx$$

and because $f(x)$ has compact support on $[-1,1]$, $f(nx)$ has compact support on $\left[-\frac{1}{n}, \frac{1}{n}\right]$, and $f_n(x) = nf(nx)$,

$$= \int_{-1/n}^{1/n} f_n(x)\phi(x)\ dx.$$

Now, by the mean value theorem, since $f_n(x)$ does not change sign on $\left[-\frac{1}{n}, \frac{1}{n}\right]$, there exists a point $x_0 \in \left(-\frac{1}{n}, \frac{1}{n}\right)$ such that

$$\begin{aligned}
&= \phi(x_0) \int_{-1/n}^{1/n} f_n(x)\, dx \\
&= \phi(x_0) \int_{-1/n}^{1/n} n\frac{3}{4}(1 - x^2n^2)\, dx \\
&= \phi(x_0) \left[\frac{3n}{4}\left(x - \frac{x^3n^2}{3}\right)\right]_{-1/n}^{1/n} \\
&= \phi(x_0) \left[\frac{3n}{4}\left(\frac{1}{n} - \frac{n^2}{3n^3}\right) + \frac{3n}{4}\left(\frac{1}{n} - \frac{n^2}{3n^3}\right)\right] \\
&= \phi(x_0) \left[2 \cdot \frac{3n}{4}\left(\frac{2}{3n}\right)\right] \\
&= \phi(x_0)
\end{aligned}$$

and since $x_0 \in \left(-\frac{1}{n}, \frac{1}{n}\right)$, as $n \to \infty$, $x_0 \to 0$, so

$$F_n(\phi) \to \phi(0)$$

i.e.

$$F_n \to \delta_0$$

✓

Problem 1.71 $*$ Let

$$g_n(x) = \begin{cases} -\frac{3}{2}n^3x & \text{if } |x| < \frac{1}{n} \\ 0 & \text{if } |x| \geq \frac{1}{n} \end{cases}.$$

Show that $G_n \to \delta_0'$, as $n \to \infty$, in $\mathcal{D}'(\mathbb{R})$.

Solution Note that

$$g_n(x) = \frac{d}{dx} f_n(x)$$

using the definition of f from Problem 1.70

$$\begin{aligned}
&= \frac{d}{dx}\left[\frac{3n}{4}\left(1 - n^2x^2\right)\right] \qquad \text{on } |x| < \frac{1}{n} \\
&= -\frac{3n}{4} 2n^2x \\
&= -\frac{3}{2}n^3x \qquad \text{on } |x| < \frac{1}{n}
\end{aligned}$$

and since $F_n \to \delta_0$, we conclude that $G_n \to \delta_0'$, so long as I can show that $\partial H_n \to \partial H$ for some distribution H. And I can do this by pointing out that

$$\begin{aligned}
\partial H_n(\phi) &= H_n(-\partial\phi) \\
&\xrightarrow[n\to\infty]{} H(-\partial\phi) \\
&= \partial H(\phi)
\end{aligned}$$

since $H_n = F_n$ and $H = \delta_0$. ✓

Problem 1.72 Let $f \in C(\mathbb{R}^n)$ satisfy

$$\int_{\mathbb{R}^n} f(x)dx = 1, \quad \int_{\mathbb{R}^n} |f(x)|dx < \infty.$$

Let

$$f_p(x) = p^n f(px), \quad p = 1, 2, \ldots.$$

Show that $f_p \to \delta$, as $p \to \infty$, in $\mathcal{D}'(\mathbb{R}^n)$. (here f_p stands for the distribution F_{f_p})

Solution

$$F_{f_p}(\phi) = p^n \int_{\mathbb{R}^n} f(px)\phi(x)\, dx$$

Letting $y = px$, then $dy = p^n\, dx$, since x is an n-dimensional vector

$$\begin{aligned}
&= \int_{\mathbb{R}^n} f(y)\phi\left(\frac{y}{p}\right)\, dy \\
&= \int_{\mathbb{R}^n} f(y)\phi(0)\, dy + \int_{\mathbb{R}^n} f(y) \underbrace{\left[\phi\left(\frac{y}{p}\right) - \phi(0)\right]}_{\to 0 \text{ pointwise}} dy \\
&= \phi(0) \underbrace{\int_{\mathbb{R}^n} f(y)\, dy}_{1 \text{ by definition}} \\
&= \phi(0)
\end{aligned}$$

Now, pointwise convergence is OK but we really want uniform convergence for rigour, i.e. we want to show

$$\forall \varepsilon\ \exists P(\varepsilon) \text{ s.t. } p > P(\varepsilon) \implies \left| \int_{\mathbb{R}^n} f(y) \left[\phi\left(\frac{y}{p}\right) - \phi(0)\right] dy \right| < \varepsilon$$

The trick we'll employ is that

$$\int_{\mathbb{R}^n} = \int_{B_A} + \int_{\mathbb{R}^n \backslash B_A}$$

Let's deal with the area outside the ball of radius A first. Using part of the question setup,

$$\int_{\mathbb{R}^n} |f(x)|dx = C < \infty \implies \forall \delta\ \exists A(\delta) \text{ s.t. } \int_{\mathbb{R}^n \backslash B_A} |f(x)|dx < \delta$$

So if we define

$$M = \max |\phi(y)|$$

and take

$$\delta = \frac{\varepsilon}{4M}$$

then we have

$$\int\limits_{\mathbb{R}^n \setminus B_A} \left| f(y) \left[\phi\left(\frac{y}{p}\right) - \phi(0) \right] \right| dy \leq 2M \int\limits_{\mathbb{R}^n \setminus B_A} |f(y)| dy$$
$$\leq \frac{\varepsilon}{2}$$

Now, inside the ball of radius A,

$$\left| \frac{y}{p} \right| \leq \frac{A}{p} \xrightarrow[p\to\infty]{} 0$$

which we can write analysis-style as

$$\left| \phi\left(\frac{y}{p}\right) - \phi(0) \right| \leq \frac{\varepsilon}{2C} \qquad \text{if } p \geq P(\varepsilon)$$

(picking $\frac{\varepsilon}{2C}$ for use later!) Therefore

$$\int\limits_{B_A} |f(y)| \left| \phi\left(\frac{y}{p}\right) - \phi(0) \right| dy \leq \frac{\varepsilon}{2C} \int\limits_{\mathbb{R}^n} |f(y)| dy = \frac{\varepsilon}{2}$$

✓

1.10 The Laplace Operator and Green's Function

Let us have $\phi_{\boldsymbol{\alpha}} \in C^\infty(\Omega)$ and $F \in \mathcal{D}'(\Omega)$, and

$$DF = \sum_{|\boldsymbol{\alpha}| \leq m} \phi_{\boldsymbol{\alpha}}(x) \partial^{\boldsymbol{\alpha}} F.$$

Can we solve the equation $DF = \delta_y$?

If we can, then we have a distribution $G_y \in \mathcal{D}'(\Omega)$ which we call *Green's function.* Typically G_y depends continuously on y.

Recall that the Laplacian, ∇^2, in $\mathbb{R}^n$ is given by

$$\nabla^2 \phi = \sum_{i=1}^n \frac{\partial^2 \phi}{\partial x_i^2}.$$

Definition 1.73 A function (or, more generally, a distribution), $\Phi_y \in \mathcal{D}'(\mathbb{R}^n)$ is called *Green's function* (or a *fundamental solution*) for the Laplacian, ∇^2, at the point $y \in \mathbb{R}^n$, if

$$-\nabla^2 \Phi_y = \delta_y.$$

In the study of electromagnetism, the most important case is $n = 3$ (and, sometimes, $n = 1$ or $n = 2$). So, let us look for Green's function in $\mathbb{R}^3$.

Lemma 1.74 The function

$$\Phi_y(x) = \frac{1}{4\pi|x-y|} \tag{1.22}$$

is the fundamental solution for ∇^2 in $\mathbb{R}^3$.

Before proving the lemma, let us make some remarks. The notation $\Phi_y(x)$ for the Green's function is typical in physics, in mathematics we tend to use $G_y(x)$. Note also that this fundamental solution is a distribution associated with the function (1.22). Note that $\Phi_y(x) \notin C(\mathbb{R}^3)$, as it has a singularity at $x = y$, however, $\Phi_y(x) \in L^1_{\text{loc}}(\mathbb{R}^3)$. Strictly speaking we should write, instead of Φ_y, F_{Φ_y} but from now on we will identify f as representing F_f and just write

$$f(\phi) = F_f(\phi).$$

Proof: We need to show that

$$-(\nabla^2 \Phi_y)(\phi) = \delta_y(\phi) = \phi(y). \tag{1.23}$$

By definition

$$\begin{aligned}(\nabla^2 \Phi_y)(\phi) &= \Phi_y(\nabla^2 \phi) \\ &= \int_{\mathbb{R}^3} \frac{\nabla^2 \phi(x)}{4\pi|x-y|} dx.\end{aligned}$$

Denote by $B_\varepsilon(y)$ a ball of radius ε centred at y and observe that

$$\int_{\mathbb{R}^3} \frac{\nabla^2 \phi(x)}{4\pi|x-y|} dx = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^3 \setminus B_\varepsilon(y)} \frac{\nabla^2 \phi(x)}{4\pi|x-y|} dx. \tag{1.24}$$

Since $\frac{1}{4\pi|x-y|}$ is smooth in $\mathbb{R}^3 \setminus B_\varepsilon(y)$ we can integrate by parts in the above integral. Recall the second of Green's identities from MATH1402,

$$\int_V \left(g\nabla^2 f - f\nabla^2 g\right) dV = \int_S \left(\frac{\partial f}{\partial \mathbf{n}} g - f\frac{\partial g}{\partial \mathbf{n}}\right) dS, \tag{1.25}$$

where S is the surface surrounding V and $\mathbf{n}$ is the exterior unit normal to S.

We take

$$\begin{aligned} V &= \mathbb{R}^3 \setminus B_\varepsilon(y) \\ f &= \phi \\ g &= \frac{1}{4\pi|x-y|}. \end{aligned}$$

Note that since $\phi \in \mathcal{D}(\mathbb{R}^3)$, we can integrate over the infinite domain $\mathbb{R}^3 \setminus B_\varepsilon(y)$. Note also that on the sphere $S_\varepsilon(y)$, the normal vector $\mathbf{n}$ exterior to V looks into $B_\varepsilon(y)$. At last, note that

$$\nabla^2 g = \nabla^2 \left(\frac{1}{4\pi|x-y|} \right) = 0. \tag{1.26}$$

So

$$\int\limits_V \left(g\nabla^2 f - f\nabla^2 g \right) dV = \int\limits_S \left(\frac{\partial f}{\partial \mathbf{n}} g - f \frac{\partial g}{\partial \mathbf{n}} \right) dS,$$

becomes

$$\int\limits_{R^3 \setminus B_\varepsilon(y)} \frac{\nabla^2 \phi}{4\pi|x-y|} dV = \int\limits_{S_\varepsilon} \left(\frac{\partial \phi}{\partial \mathbf{n}} \left(\frac{1}{4\pi|x-y|} \right) - \phi \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{4\pi|x-y|} \right) \right) dS \tag{1.27}$$

Now, plugging this into equation 1.24 gives us

$$\begin{aligned} \lim_{\varepsilon\to 0} \int\limits_{\mathbb{R}^3 \setminus B_\varepsilon(y)} \frac{\nabla^2 \phi(x)}{4\pi|x-y|} dx &= \lim_{\varepsilon \to 0} \int\limits_{S_\varepsilon} \left(\frac{\partial \phi}{\partial \mathbf{n}} \left(\frac{1}{4\pi|x-y|} \right) - \phi \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{4\pi|x-y|} \right) \right) dS \\ &= \frac{1}{4\pi} \lim_{\varepsilon \to 0} \int\limits_{S_\varepsilon} \left(\frac{\partial \phi}{\partial \mathbf{n}} \left(\frac{1}{|x-y|} \right) - \phi \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{|x-y|} \right) \right) dS \end{aligned}$$

Now recall that y is *fixed*—this is really important because it means we can do the following substitutions

$$z := x - y$$

Then convert to *spherical coordinates*

$$(z = \rho, \theta, \varphi)$$

So $\phi(x)$ becomes a function of $\phi(y + \rho\boldsymbol{\omega})$, where $\boldsymbol{\omega}$ is a unit vector made up of θ and φ. In fact,

$$\boldsymbol{\omega} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$

Observe that

$$\frac{\partial}{\partial \mathbf{n}} = -\frac{\partial}{\partial \rho}$$

Giving us

$$\frac{1}{4\pi} \lim_{\varepsilon \to 0} \int\limits_{\rho=\varepsilon} \left(-\frac{\partial \phi}{\partial \rho} \left(\frac{1}{\rho} \right) + \phi \frac{\partial}{\partial \rho} \left(\frac{1}{\rho} \right) \right) dS$$

$$= \frac{1}{4\pi} \lim_{\varepsilon \to 0} \int\limits_{\rho=\varepsilon} \left(-\frac{\partial \phi}{\partial \rho} \left(\frac{1}{\varepsilon} \right) - \phi \frac{1}{\varepsilon^2} \right) dS$$

and $dS = \varepsilon^2 \sin\theta \, d\theta \, d\varphi$ which gives us that

$$(\nabla^2 \Phi_y)(\phi) = -\frac{1}{4\pi} \lim_{\varepsilon \to 0} \int\limits_{\rho=\varepsilon} \left(\frac{\partial \phi}{\partial \rho} \frac{1}{\varepsilon} + \phi \frac{1}{\varepsilon^2} \right) \varepsilon^2 \sin\theta \, d\theta \, d\varphi$$

and of course $\displaystyle\int\limits_{\rho=\varepsilon} = \iint\limits_{\substack{0 \le \theta \le \pi \\ 0 \le \varphi \le 2\pi}}$.

Now note that

$$\iint \frac{\partial \phi}{\partial \rho} \frac{1}{\varepsilon} \varepsilon^2 \sin\theta \, d\theta \, d\varphi \xrightarrow[\varepsilon \to 0]{} 0$$

since $\frac{\partial \phi}{\partial \rho}$ is bounded, and from what's left,

$$-\frac{1}{4\pi} \iint\limits_{\theta,\varphi} \phi(y + \varepsilon \boldsymbol{\omega}) \sin\theta \, d\theta \, d\varphi \xrightarrow[\varepsilon \to 0]{} -\frac{1}{4\pi} \phi(y) \underbrace{\iint\limits_{\theta,\varphi} \sin\theta \, d\theta \, d\varphi}_{\substack{\text{vol. of sphere} \\ \text{of radius } 1 \,=\, 4\pi}}$$

$$= -\phi(y).$$

□

So we have shown that

$$-\nabla^2 \left(\frac{1}{4\pi |x - y|} \right) = \delta_y.$$

We want to let $y = 0$:

$$-\nabla^2 \left(\frac{1}{4\pi |x|} \right) = \delta_0?$$

Well,

$$-\nabla^2 \left(T_y \frac{1}{4\pi |x|} \right) = T_y \delta_0,$$

and noting that $T_y \circ \delta^{\boldsymbol{\alpha}} = \delta^{\boldsymbol{\alpha}} \circ T_y$,

$$T_y \left(-\nabla^2 \frac{1}{4\pi |x|} \right) = T_y \delta_0 = \delta_y,$$

So it's sufficient to prove that

$$-\nabla^2 \left(\frac{1}{4\pi |x|} \right) = \delta_0$$

and use commutativity to solve

$$-\nabla^2 \left(\frac{1}{4\pi |x - y|} \right) = \delta_y$$

1.10.1 Motivation

Why is knowledge of Green's function such a useful thing?

Say we want to solve a differential equation with constant coefficients. Let $c_{\boldsymbol{\alpha}}$ be constants and D be the differential operator:

$$D = \sum_{|\boldsymbol{\alpha}| \leq m} c_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}}.$$

Suppose we wish to solve the differential equation:

$$DF = H$$

where H is a known function (or distribution) and F is our unknown. Now suppose we know G such that $DG = \delta$, i.e. we know the Green's function for our differential operator D. Now construct $F = G * H$ so that:

$$\begin{aligned} DF &= D(G * H) \\ &= \sum_{|\boldsymbol{\alpha}| \leq m} c_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} (G * H). \end{aligned}$$

By Proposition 1.62 and linearity

$$\begin{aligned} &= \left(\sum_{|\boldsymbol{\alpha}| \leq m} c_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} G \right) * H \\ &= \delta * H \\ &= H. \end{aligned}$$

So we see that knowledge of the Green's function for a differential operator, allows us to solve a differential equation with any right hand side, by use of the comparatively simple operation of convolution.

For example if you ever wanted to solve $-\nabla^2 F = g$ in $\Omega \subset \mathbb{R}^3$, then

$$F = \frac{1}{4\pi |x|} * g.$$

Observe that if $g \in C_0(\mathbb{R}^3)$,

$$F(x) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{g(y)}{|x - y|} dy.$$

Chapter 2

Electrostatics

Now finally we can apply this mathematics. In this chapter we apply it to electrostatics and in the following chapter, we'll apply it to magnetism (which is very similar), and in doing so, derive Maxwell's famous equations.

2.1 Introduction to electrostatics

2.1.1 Coulomb's law

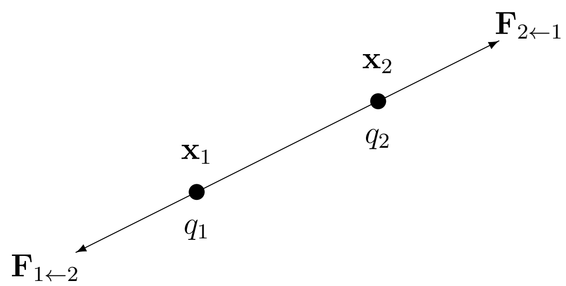

Figure 2.1: Two charged particles exerting a force on each other

If we have two charged particles of charges q_1, q_2 at points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$, then Coulomb's law tells us that

$$F_{1\leftarrow 2} = kq_1q_2\frac{\mathbf{x}_1 - \mathbf{x}_2}{|\mathbf{x}_1 - \mathbf{x}_2|^3} \tag{2.1}$$

where k is a constant to do with our choice of units (typically $\frac{1}{4\pi\varepsilon_0}$).

How do they exert a force on each other? Scientists used to think there was an ether which gave a transfer material for charge to go through. This was later rubbished, of course, and it turns out that when you place a charged particle in space, it propagates a magnetic field.

If we place a charge q_2 in an electric field $\mathbf{E}(\mathbf{x})$, the force exerted on the particle is

$$\mathbf{F}_{q_2}(\mathbf{x}) = q_2\mathbf{E}(\mathbf{x}).$$

Combining this with Coulomb's law (2.1) above,

$$\mathbf{E}(\mathbf{x}) = kq_1 \frac{\mathbf{x}_1 - \mathbf{x}_2}{|\mathbf{x}_1 - \mathbf{x}_2|^3} = kq_1 \frac{\widehat{(\mathbf{x}_1 - \mathbf{x}_2)}}{|\mathbf{x}_1 - \mathbf{x}_2|^2}$$

Electric fields add linearly.

2.1.2 Electric potential

If we have an electric field $\mathbf{E}$, then we define the (scalar) *electric potential* Φ by

$$\mathbf{E}(\mathbf{x}) = -\nabla\Phi(\mathbf{x})$$

(note that $\nabla\times\mathbf{E} = 0$ and the domain is simply connected, so by what we learnt in MATH1402, $\mathbf{E}$ has a potential Φ.)

Obviously Φ is defined up to a constant, so it has no *physical* meaning, but the *difference* of potentials does, of course.

So, taking a path ℓ between our two charges at $\mathbf{x}_1$ and $\mathbf{x}_2$, at any point $\widetilde{\mathbf{x}}$ along the path,

$$\mathbf{F}(\widetilde{\mathbf{x}}) = q\mathbf{E}(\widetilde{\mathbf{x}}).$$

The work done then getting from $\widetilde{\mathbf{x}}$ to $\widetilde{\mathbf{x}} + \delta\widetilde{\mathbf{x}}$ is then $dW = \mathbf{F}(\widetilde{\mathbf{x}}) \cdot d\widetilde{\mathbf{x}}$, so

$$\begin{aligned}
\text{Work} = W &= \int_\ell \mathbf{F}(\widetilde{\mathbf{x}}) \cdot d\widetilde{\mathbf{x}} \\
&= q \int_\ell \mathbf{E}(\widetilde{\mathbf{x}}) \cdot d\widetilde{\mathbf{x}} \\
&= -q \int_\ell \nabla\Phi(\widetilde{\mathbf{x}}) \cdot d\widetilde{\mathbf{x}} \\
&= -q\left[\Phi(\mathbf{x}_2) - \Phi(\mathbf{x}_1)\right] \\
&= q\left[\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)\right]
\end{aligned}$$

The convention is that $\Phi = 0$ at infinity.

2.2 The fundamental equations of electrostatics

We are now going to derive the fundamental equations of electrostatics, first in the case of discrete point charges, and second in the case of a continuous charge distribution.

2.2.1 The discrete case

Definition 2.1 If we have a group of electric charges q_i at positions $\mathbf{y}_i$, for $i = 1, \ldots, p$, then we say that the *electric potential* at the point $\mathbf{x}$ is

$$\Phi^t(\mathbf{x}) = \sum_{i=1}^{p} \frac{q_i}{4\pi\varepsilon_0 |\mathbf{x} - \mathbf{y}_i|}$$

(the t is for total!)

Definition 2.2 The *electric field* created by these charges is defined as

$$\mathbf{E}^t(\mathbf{x}) = -\nabla\Phi^t(\mathbf{x})$$

Definition 2.3 The *charge density distribution* created by these charges is defined as

$$\rho = \sum_{i=1}^{p} q_i \delta_{y_i}$$

Lemma 2.4 If Φ is Green's function and ρ is the discrete distribution as described above, then the electrostatic potential due to the distribution ρ of charges is

$$\Phi^t = \Phi * \frac{\rho}{\varepsilon_0}$$

Proof:

$$\begin{aligned} \Phi * \frac{\rho}{\varepsilon_0} &= \frac{1}{4\pi|\mathbf{x}|} * \frac{1}{\varepsilon_0} \sum_{i=1}^{p} q_i \delta_{\mathbf{y}_i} \\ &= \frac{1}{\varepsilon_0} \sum_{i=1}^{p} \underbrace{\left(\frac{1}{4\pi|\mathbf{x}|} * q_i \delta_{\mathbf{y}_i} \right)}_{\text{What is this?}} \end{aligned}$$

Recall that, for a distribution H,

$$\begin{aligned} (\delta_{\mathbf{z}} * H)(\phi) &:= H(\delta_{\mathbf{z}} \phi^{\mathbf{y}}) \\ &= H(\phi^{\mathbf{y}}(\mathbf{z})) \\ &= H(\phi^{\mathbf{z}}) \qquad (\text{since } \phi^{\mathbf{y}}(\mathbf{z}) = \phi^{\mathbf{z}}(\mathbf{y})) \\ &= (T_{\mathbf{z}} H)(\phi) \end{aligned}$$

so $\delta_{\mathbf{z}} * H = T_{\mathbf{z}} H$.

So

$$\frac{1}{4\pi|\mathbf{x}|} * q_i \delta_{\mathbf{y}_i} = \frac{q_i}{4\pi|\mathbf{x} - \mathbf{y}_i|}$$

and we get

$$\sum_{i=1}^{p} \frac{q_i}{4\pi\varepsilon_0|\mathbf{x}-\mathbf{y}_i|} = \Phi^t(\mathbf{x})$$

□

2.2.2 The continuous case

Now we are *given* a distribution $\rho \in \mathcal{E}'(\mathbb{R}^3)$. So what is Φ^ρ, the electrostatic potential due to ρ?

Well, we want to use

$$\sum_{i=1}^{p} \frac{q_i}{4\pi\varepsilon_0|\mathbf{x}-\mathbf{y}_i|}$$

but we want to make it continuous.

So consider that inside a small cube centred at $\mathbf{x}_i$ in $\mathbb{R}^3$, the approximate charge is

$$q_i = \rho(\mathbf{x}_i)\ \delta x\ \delta y\ \delta z.$$

Now take the limit as $\delta x, \delta y, \delta z \to 0$.

$$\begin{aligned}
\lim_{\left.\begin{smallmatrix}\delta x\\ \delta y\\ \delta z\end{smallmatrix}\right\}\to 0} \sum_i \frac{\rho(\mathbf{x}_i)\ \delta x\ \delta y\ \delta z}{4\pi\varepsilon_0|\mathbf{x}-\mathbf{x}_i|} &= \frac{1}{4\pi\varepsilon_0}\int_{\mathbb{R}^3} \frac{\rho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}d\mathbf{y} \\
&= \frac{1}{4\pi|\mathbf{x}|} * \frac{\rho}{\varepsilon_0} \\
&= \Phi * \frac{\rho}{\varepsilon_0} \\
&= \Phi^\rho
\end{aligned}$$

So we get the same result as we did in the discrete case in Lemma 2.4, i.e.

$$\Phi^\rho = \Phi * \frac{\rho}{\varepsilon_0}$$

2.2.3 The equations

Take the Laplacian of our potential distribution Φ^ρ:

$$\begin{aligned}
-\nabla^2\Phi^\rho &= -\nabla^2\Phi * \frac{\rho}{\varepsilon_0} \\
\textit{(by definition of } \Phi\textit{)} &= \delta * \frac{\rho}{\varepsilon_0} \\
&= \frac{\rho}{\varepsilon_0}
\end{aligned}$$

And recall that the definition of the electric field due to a particle distribution ρ is

$$\mathbf{E}^\rho = -\nabla\Phi^\rho$$

which gives us…

Definition 2.5 The *fundamental equations of electrostatics* are

1. $\nabla \cdot \mathbf{E}^\rho = \dfrac{\rho}{\varepsilon_0}$ (Gauss' law in differential form)
2. $\nabla \times \mathbf{E}^\rho = \mathbf{0}$

2.3 Divergence theorem and Stokes' theorem

Problem 2.6 If $F \in \mathcal{D}'(\Omega)$, where $\Omega \subset \mathbb{R}^3$, then

$$\left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right) = \nabla F \in \mathcal{D}'(\Omega)^3.$$

Show that

$$\nabla \times \nabla F = 0 \qquad \text{and} \qquad \nabla \cdot (\nabla F) = \nabla^2 F$$

Taking the fundamental equations of electrostatics, let us put these into the divergence theorem from MATH1402, for a domain V with boundary S:

$$\int_S \mathbf{E}^\rho \cdot \mathbf{n}\, dS = \int_V \nabla \cdot \mathbf{E}^\rho\, dV$$

But $\nabla \cdot \mathbf{E}^\rho = \frac{\rho}{\varepsilon_0} \implies$

$$= \frac{1}{\varepsilon_0}\int_V \rho\, dV = \frac{Q(V)}{\varepsilon_0} \qquad \text{(Gauss' law in integral form)}$$

where $Q(V)$ is the total charge in V. This is known as *Gauss' law in integral form.*

Now what about putting $\mathbf{E}^\rho$ into Stokes' law? Recall that Stokes' law, for some vector $\mathbf{F}$, is

$$\int_S (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS = \oint_C \mathbf{F} \cdot d\mathbf{r}$$

Substituting $\mathbf{F} = \mathbf{E}^\rho$ gives us

$$0 = \int_S \underbrace{(\nabla \times \mathbf{E}^\rho)}_{0} \cdot \mathbf{n}\, dS = \oint_C \mathbf{E}^\rho \cdot d\mathbf{r}$$

but remember that Stokes' law is only valid for simply connected domains.

2.4 Finding E and Φ for different ρ

Problem 2.7 Suppose that charge distribution is a function of radius alone, i.e.

$$\rho = \rho(r) \qquad r^2 = x^2 + y^2 + z^2 = |\mathbf{x}|^2.$$

1. Find $\mathbf{E}^\rho(\mathbf{x})$.
2. Then using part 1, find $\Phi^\rho(\mathbf{x})$ in terms of $\rho(r)$.

Solution

1. Because of spherical symmetry,

$$\mathbf{E} = \mathbf{E}(r) = \widetilde{E}(r)\widehat{\mathbf{r}}$$

where $\widehat{\mathbf{r}}$ is a unit vector in the radial direction.

By Gauss' law in integral form, we have

$$\begin{aligned}\iint\limits_{S_r} \mathbf{E}\cdot\mathbf{n}\,dS &= \frac{1}{\varepsilon_0}Q(B_r)\\ &= \frac{1}{\varepsilon_0}\int\limits_{B_r}\rho\,dV\\ &= \frac{1}{\varepsilon_0}\underbrace{\int_0^r \rho(r')r'^2\,dr'}_{:=R(r)}\underbrace{\int\limits_{\mathbb{S}^2} d\boldsymbol{\omega}}_{4\pi}\\ &= \frac{4\pi}{\varepsilon_0}R(r)\end{aligned}$$

where $\boldsymbol{\omega} = \sin\theta\,d\theta\,d\varphi$.

But we also have

$$\begin{aligned}\iint\limits_{S_r}\mathbf{E}\cdot\mathbf{n}\,dS &= \iint\limits_{S_r}\mathbf{E}\cdot\mathbf{r}\,dS\\ &= \widetilde{E}(r)\iint\limits_{S_r} dS\\ &= 4\pi r^2\widetilde{E}(r)\end{aligned}$$

Therefore

$$\widetilde{E}(r) = \frac{R(r)}{\varepsilon_0 r^2}.$$

Note that $\widetilde{E}\to 0$ as $r\to 0$ since $R(r)\sim r^3$.

2. Now to find Φ^ρ. We know by definition that

$$-\nabla\Phi(r) = \mathbf{E}$$

Therefore, setting up Φ so that it is 0 at ∞,

$$\begin{aligned}\Phi(r) &= -\int_\infty^r \widetilde{E}(r')\,dr' \\ &= \int_r^\infty \widetilde{E}(r')\,dr' \\ &= \int_r^\infty \frac{R(r')}{r'^2}\,dr'\end{aligned}$$

We could be done at this point but it would be nice to integrate by parts since R itself is an integral. So, letting our 'first' be $R(r)$ and our 'second' be $1/r^2$,

$$\begin{aligned}&= \left[R(r)\frac{-1}{r}\right]_r^\infty + \int_r^\infty R'(r')\frac{1}{r'}\,dr' \\ &= \frac{R(r)}{r} + \int_r^\infty \rho(r')r'\,dr'\end{aligned}$$

✓

Problem 2.8 $*$ Let $\rho = \rho(r)$ in cylindrical coordinates (r, φ, z), i.e. $r^2 = x^2 + y^2$. Find $\mathbf{E}^\rho$ (using Gauss' law).

Solution By symmetry, $\mathbf{E}^\rho = \mathbf{E}(r) = \widetilde{E}(r)\widehat{\mathbf{r}}$, so

$$\begin{aligned}\nabla\cdot\mathbf{E}^\rho &= \nabla\cdot\widetilde{E}(r)\widehat{\mathbf{r}} \\ &= \frac{1}{r}\frac{d}{dr}\left(r\widetilde{E}(r)\right) \\ &= \frac{\rho}{\varepsilon_0}\end{aligned}$$

And so

$$\begin{aligned}\widetilde{E}(r) - \underbrace{\widetilde{E}(0)}_{0} &= \frac{1}{r\varepsilon_0}\int_0^r \widetilde{r}\rho(\widetilde{r})\,d\widetilde{r} \\ \implies \widetilde{E}(r) &= \frac{1}{r\varepsilon_0}\int_0^r \widetilde{r}\rho(\widetilde{r})\,d\widetilde{r}\end{aligned}$$

and $\mathbf{E}^\rho = \widetilde{E}(r)\widehat{\mathbf{r}}$. ✓

2.5 Dipoles

We now introduce the idea of *dipoles*.

Definition 2.9 If we have a charge distribution

$$\rho = \mathbf{p} \cdot \nabla_y \delta_y = -\mathbf{p} \cdot \nabla_x \delta_y \qquad \mathbf{p} = p\mathbf{u}$$

where $\mathbf{u}$ is a unit vector, then $\mathbf{p} \cdot \nabla \delta_y$ is called the *dipole* of size p in the direction $\mathbf{u}$ at the point y.

Recall for maths' sake that

$$\begin{aligned} \mathbf{p} \cdot \nabla \delta_y &= p\left(\frac{\partial \delta_y}{\partial x} u_1 + \frac{\partial \delta_y}{\partial y} u_2 + \frac{\partial \delta_y}{\partial z} u_3\right) \\ &= p\frac{\partial \delta_y}{\partial \mathbf{u}} \end{aligned}$$

Consider placing a charge $\frac{p}{s}$ at the point $\mathbf{y} + s\mathbf{u}$, and a charge $-\frac{p}{s}$ at the point $\mathbf{y}$. This is, physically, a dipole.

Then the electrostatic potential $\Phi^s = \Phi^s_{p,\mathbf{y}}$ due to the two point charges

$$\left(\frac{p}{s}\delta_{\mathbf{y}+s\mathbf{u}} - \frac{p}{s}\delta_{\mathbf{y}}\right)$$

is

$$\Phi^s = \frac{p}{s\varepsilon_0}\left[\frac{1}{4\pi|\mathbf{x} - (\mathbf{y} + s\mathbf{u})|} - \frac{1}{4\pi|\mathbf{x} - \mathbf{y}|}\right]$$

Now take $s \to 0$:

$$\begin{aligned} \lim_{s\to 0} \Phi^s(\mathbf{x}) &= \frac{p}{\varepsilon_0}\frac{\partial}{\partial \mathbf{u}_y}\Phi_y(x) \\ &= -\frac{p}{\varepsilon_0}\frac{\partial}{\partial \mathbf{u}_x}\Phi_y(x) \end{aligned}$$

Now in order to find $\Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{y}}$, we first look at:

$$\nabla\left(\frac{1}{4\pi|\mathbf{x}|}\right)$$

Taking each component of ∇ in turn:

$$\begin{aligned} \frac{\partial}{\partial x}\left(\frac{1}{\sqrt{x^2+y^2+z^2}}\right) &= -\frac{1}{2}\cdot 2x \cdot \frac{1}{(x^2+y^2+z^2)^{3/2}} \\ &= -\frac{x}{|\mathbf{x}|^3} \\ \frac{\partial}{\partial y}\left(\frac{1}{\sqrt{x^2+y^2+z^2}}\right) &= -\frac{y}{|\mathbf{x}|^3} \\ \frac{\partial}{\partial z}\left(\frac{1}{\sqrt{x^2+y^2+z^2}}\right) &= -\frac{z}{|\mathbf{x}|^3} \end{aligned}$$

hence

$$\nabla\left(\frac{1}{4\pi|\mathbf{x}|}\right)=-\frac{\mathbf{x}}{4\pi|\mathbf{x}|^3}$$

But we know that for a dipole (d), of size p in direction $\mathbf{u}$ sitting on the point $\mathbf{y}$,

$$\begin{aligned}
\Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{y}} &= \frac{1}{\varepsilon_0}\Phi*(\mathbf{p}\cdot\nabla_y\delta_y)\\
&= \frac{1}{\varepsilon_0}\Phi*(-\mathbf{p}\cdot\nabla_x\delta_y)\\
&= -\frac{1}{\varepsilon_0}\mathbf{p}\cdot\nabla_x(\Phi*\delta_y)\\
&= -\frac{1}{\varepsilon_0}(\mathbf{p}\cdot\nabla\Phi)*\delta_y\\
&= -\frac{1}{\varepsilon_0}\left(\mathbf{p}\cdot\nabla\left(\frac{1}{4\pi|\mathbf{x}|}\right)\right)*\delta_y\\
&= -\frac{1}{\varepsilon_0}\left(\mathbf{p}\cdot\left(-\frac{\mathbf{x}}{4\pi|\mathbf{x}|^3}\right)\right)*\delta_y\\
&= \frac{1}{4\pi\varepsilon_0}\left(\mathbf{p}\cdot\left(\frac{\mathbf{x}}{|\mathbf{x}|^3}\right)\right)*\delta_y\\
&= \frac{1}{4\pi\varepsilon_0}\frac{\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^3}
\end{aligned}$$

Since convolution with δ_y is the same as translation by $-\mathbf{y}$

And so we get:

Definition 2.10 The potential of a dipole of size p in the direction $\mathbf{u}$ at a point $\mathbf{y}$ is

$$\Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{y}}=\frac{1}{4\pi\varepsilon_0}\frac{\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^3}$$

Now to find the corresponding electric field. $\mathbf{E}^{\mathrm{d}}_{\mathbf{p},\mathbf{y}}$.

We know

$$\mathbf{E}^{\mathrm{d}}_{\mathbf{p},\mathbf{y}}=-\nabla\Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{y}}$$

so let's work it out.

Let $\mathbf{y}=\mathbf{0}$ for ease. Then

$$\Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{0}}=\frac{1}{4\pi\varepsilon_0}\frac{(p_1x_1+p_2x_2+p_3x_3)}{|\mathbf{x}|^3}$$

and working out $\mathbf{E}^{\mathrm{d}}_{\mathbf{p},\mathbf{0}}$ component-wise, we get

$$\mathbf{E}^{\mathrm{d}}_{\mathbf{p},\mathbf{y}} = -\nabla \Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{y}}$$

so

$$\frac{\partial \Phi^{\mathrm{d}}_{\mathbf{p},\mathbf{0}}}{\partial x_1} = -\frac{1}{4\pi\varepsilon_0}\left(\frac{p_1}{|\mathbf{x}|^3} - \frac{3p_1 x_1^2}{|\mathbf{x}|^5}\right).$$

Continuing this process we end up with

$$\mathbf{E}^{\mathrm{d}}_{\mathbf{p},\mathbf{0}} = \frac{1}{4\pi\varepsilon_0}\left(\frac{3(\mathbf{p}\cdot\widehat{\mathbf{x}})\widehat{\mathbf{x}} - \mathbf{p}}{|\mathbf{x}|^3}\right)$$

2.5.1 Jumps

Recall that for a surface S and $h \in C_0(S)$,

$$F_{h,S}(\phi) = \int\limits_S h(\mathbf{x})\phi(\mathbf{x})\, dS_{\mathbf{x}}$$

We want to find, for the h-surface charge distribution, $\Phi_{h,S}$ and $\mathbf{E}_{h,S}$.

Lemma 2.11

$$\Phi_{h,S}(\mathbf{x}) := \frac{1}{\varepsilon_0}\Phi * F_{h,S} = \frac{1}{4\pi\varepsilon_0}\int\limits_S \frac{h(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} dS_{\mathbf{y}}$$

Proof: We have

$$\begin{aligned} \Phi_{h,S}(\phi) &= \frac{1}{\varepsilon_0}\Phi * F_{h,S}(\phi^{\mathbf{y}}) \\ F_{h,S}(\phi^{\mathbf{y}}) &= \int\limits_S h(\mathbf{x})\phi(\mathbf{x}+\mathbf{y}) dS_{\mathbf{x}} \qquad \in \mathcal{D}(\mathbb{R}^3) \end{aligned}$$

(where $\mathbf{y}$ is fixed in $dS_{\mathbf{x}}$). Therefore

$$\begin{aligned} \Phi_{h,S}(\phi) &= \frac{1}{4\pi\varepsilon_0}\int\limits_{\mathbb{R}^3} \frac{1}{|\mathbf{y}|}\left[\int\limits_S h(\mathbf{x})\phi(\mathbf{x}+\mathbf{y}) dS_{\mathbf{x}}\right] d\mathbf{y} \\ &= \frac{1}{4\pi\varepsilon_0}\int\limits_S h(\mathbf{x})\left[\int\limits_{\mathbb{R}^3} \frac{\phi(\mathbf{x}+\mathbf{y})}{|\mathbf{y}|} d\mathbf{y}\right] dS_{\mathbf{x}} \\ &= \frac{1}{4\pi\varepsilon_0}\int\limits_S h(\mathbf{x})\left[\int\limits_{\mathbb{R}^3} \frac{\phi(\mathbf{z})}{|\mathbf{x}-\mathbf{z}|} d\mathbf{z}\right] dS_{\mathbf{x}} \\ &= \int\limits_{\mathbb{R}^3} \phi(\mathbf{z})\left[\frac{1}{4\pi\varepsilon_0}\int\limits_S \frac{h(\mathbf{x})}{|\mathbf{x}-\mathbf{z}|} dS_{\mathbf{x}}\right] d\mathbf{z} \end{aligned}$$

Hence

$$\Phi_{h,S}(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0} \int\limits_S \frac{h(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} dS_{\mathbf{y}}$$

$\square$

Note that $\Phi_{h,S}$ is continuous over S. However, $\mathbf{E}_{h,S}$ has a jump discontinuity across S.

To find $\mathbf{E}$, we need to find $\nabla\Phi$, and $\frac{h(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}$ becomes like $\sim \frac{1}{r^2} r\ dr$ and we have problems. But this is how we get around it:

Imagine we have a surface S upon which we have a charge density h. Take a cylinder of circular area ω and of height 2ε. Push this through the surface at a point so that it's perfectly half-way through the surface. So, denoting the cylinder C,

$$C = \omega \times (-\varepsilon, \varepsilon)\mathbf{n}$$

Gauss' law in integral form, applied to this cylinder, tells us that

$$\int\limits_{\partial C} (\mathbf{E}_{h,S} \cdot \boldsymbol{\nu}) dS = \frac{Q(C)}{\varepsilon_0}$$

where $\boldsymbol{\nu}$ is normal to ∂C

$$= \frac{1}{\varepsilon_0} \int\limits_\omega h\ dS \tag{2.2}$$

because the only charge in the cylinder is in the disc of area ω.

Now, let's take a look at the left hand side.

$$\int\limits_{\partial C_\omega^\varepsilon} = \int\limits_{\omega_\varepsilon^+} + \int\limits_{\omega_\varepsilon^-} + \int\limits_{S_\varepsilon}$$

where

$$\begin{aligned} \omega_\varepsilon^\pm &= \{\omega \pm \varepsilon\mathbf{n}\} && \text{the top and bottom of the cylinder} \\ S_\varepsilon &= \{\partial\omega \times (-\varepsilon, \varepsilon)\mathbf{n}\} && \text{the side surface of the cylinder} \end{aligned}$$

Let's take a look at these three integrals as we take $\varepsilon \to 0$

$$\int\limits_{S_\varepsilon} \mathbf{E} \cdot \boldsymbol{\nu}\ dS \to 0$$

since $\text{Area}(S_\varepsilon) \to 0$

$$\int\limits_{\omega_\varepsilon^+} \mathbf{E} \cdot \boldsymbol{\nu}\, dS = \int\limits_{\omega} \mathbf{E}(x + \varepsilon\mathbf{n}) \cdot \mathbf{n}\, dS_x$$

$$\xrightarrow[\varepsilon \to 0^+]{} \int\limits_{\omega} \lim_{\varepsilon \to 0} \mathbf{E}(x + \varepsilon\mathbf{n}) \cdot \mathbf{n}\, dS_x$$

$$= \int\limits_{\omega} \mathbf{E}^+(x) \cdot \mathbf{n}\, dS_x$$

and similarly

$$\int\limits_{\omega_\varepsilon^-} \mathbf{E} \cdot \boldsymbol{\nu}\, dS = \int\limits_{\omega} \mathbf{E}(x - \varepsilon\mathbf{n}) \cdot (-\mathbf{n})\, dS_x$$

$$\xrightarrow[\varepsilon \to 0^-]{} - \int\limits_{\omega} \mathbf{E}^-(x) \cdot \mathbf{n}\, dS_x$$

Therefore,

$$\lim_{\varepsilon \to 0} \int\limits_{\partial C_\omega^\varepsilon} \mathbf{E}_{h,S} \cdot \boldsymbol{\nu}\, dS = \int\limits_{\omega} \left[\mathbf{E}^+(x) - \mathbf{E}^-(x)\right] \cdot \mathbf{n}\, dS$$

So taking that and comparing it with the right-hand side of equation 2.2, we get

$$\left[\mathbf{E}^+(x) - \mathbf{E}^-(x)\right] \cdot \mathbf{n} = \frac{h(x)}{\varepsilon_0} \tag{2.3}$$

i.e. the normal component of $\mathbf{E}_{h,S}$ has a *jump* across S of size $\dfrac{h}{\varepsilon_0}$.

2.5.2 Tangential continuity

We are now going to show that despite the normal component of the electric field having a jump, the tangential component is continuous.

To do this, let's take a line ℓ from point x_0 to x_1 upon a surface S. Draw similar lines at a distance ε above and below ℓ. Now join lines at the sides connecting them up, forming a loop. Suppose we decide on an anticlockwise direction round this loop. So the four lines, mathematically speaking, are

$$\begin{aligned}
L_\varepsilon &= x_0 \times (\varepsilon, -\varepsilon)\mathbf{n} && \text{(left)} \\
R_\varepsilon &= x_1 \times (-\varepsilon, \varepsilon)\mathbf{n} && \text{(right)} \\
\ell_\varepsilon^- &= \ell \times -\varepsilon\mathbf{n} && \text{(bottom)} \\
\ell_\varepsilon^+ &= \ell \times \varepsilon\mathbf{n} && \text{(top)}
\end{aligned}$$

Now,

$$\oint_{C_\varepsilon^\ell} \mathbf{E} \cdot d\mathbf{r} = 0 \tag{2.4}$$

and, using a similar trick to last time,

$$\oint_{C_\varepsilon^\ell} = \int_{\ell_\varepsilon^-} + \int_{\ell_\varepsilon^+} + \int_{L_\varepsilon} + \int_{R_\varepsilon}$$

Notice that

$$\lim_{\varepsilon \to 0} \left[\int_{L_\varepsilon} + \int_{R_\varepsilon} \right] = 0$$

since the length is $2\varepsilon \to 0$. And for the other two,

$$\begin{aligned} \lim_{\varepsilon \to 0} \left[\int_{\ell_\varepsilon^-} + \int_{\ell_\varepsilon^+} \right] &= \int_\ell \left[\mathbf{E}^-(x) - \mathbf{E}^+(x) \right] \cdot d\mathbf{r} \\ &= 0 \end{aligned} \tag{2.5}$$

by the initial observation in equation 2.4.

Say that $d\mathbf{r} = \dot{\mathbf{r}}\, d\mathbf{t}$, where $\dot{\mathbf{r}}$ is a tangential vector to S, which we'll define as $\mathbf{t}$.

Change ℓ through any $x \in S$ (i.e. rotate ℓ) to have all possible tangential directions. Then 2.5 means that

$$\left[\mathbf{E}^+(x) - \mathbf{E}^-(x) \right] \cdot \mathbf{t} = 0$$

for any tangential $\mathbf{t}$, i.e. *the tangential part of the electric field is continuous!*

2.5.3 Brief recap of definitions

If we have a dipole $\mathbf{p} \cdot \nabla \delta_{\mathbf{y}}$ at $\mathbf{y}$,

$$\Phi_{\mathbf{p},\mathbf{y}}(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0} \frac{\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})}{|\mathbf{x} - \mathbf{y}|^3}$$

and across the whole surface,

$$\Phi^{\mathrm{d}}_{\mathbf{p},S}(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0} \int_S \frac{\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})}{|\mathbf{x} - \mathbf{y}|^3} dS_{\mathbf{y}}$$

Problem 2.12 $*$ Let $S = \{z = 0\}$, the xy-plane. Assume the magnitude of the dipole is constant $p = p_0$.

1. Find $\Phi(x)$, the corresponding electrostatic potential of this surface.
2. Find the jump of $\Phi(x)$ across $z = 0$.

Solution From the above,

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\varepsilon_0}\int\limits_{z=0} \frac{\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^3}\,dS_{\mathbf{y}}$$

Now, letting

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

and using the fact that $\mathbf{p} = p_0\widehat{\mathbf{z}}$,

$$= \frac{1}{4\pi\varepsilon_0}\int\limits_{z=0} \frac{p_0\cdot(x_3-y_3)}{|\mathbf{x}-\mathbf{y}|^3}\,dS_{\mathbf{y}}$$

And since on the surface $z = 0$, $y_3 = 0$,

$$\begin{aligned}
&= \frac{1}{4\pi\varepsilon_0}\int\limits_{z=0} \frac{p_0x_3}{|\mathbf{x}-\mathbf{y}|^3}\,dS_{\mathbf{y}} \\
&= \frac{p_0x_3}{4\pi\varepsilon_0}\int\limits_{z=0} \frac{1}{|\mathbf{x}-\mathbf{y}|^3}\,dS_{\mathbf{y}} \\
&= \frac{p_0x_3}{4\pi\varepsilon_0}\int\limits_{z=0} \frac{1}{\left[(x_1-y_1)^2+(x_2-y_2)^2+(x_3)^2\right]^{3/2}}\,dS_{\mathbf{y}} \\
&= \frac{p_0x_3}{4\pi\varepsilon_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\left[(x_1-y_1)^2+(x_2-y_2)^2+(x_3)^2\right]^{3/2}}\,dy_1\,dy_2
\end{aligned}$$

Using the substitutions

$$y_1' = y_1 - x_1 \qquad y_2' = y_2 - x_2 \qquad dy_1' = dy_1 \qquad dy_2' = dy_2$$

$$= \frac{p_0x_3}{4\pi\varepsilon_0}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\left[(y_1')^2+(y_2')^2+(x_3)^2\right]^{3/2}}\,dy_1'\,dy_2'$$

Then converting to polar coordinates

$$r^2 = (y_1')^2 + (y_2')^2$$

$$\begin{aligned}
&= \frac{p_0x_3}{4\pi\varepsilon_0}\int_0^{2\pi}\int_0^{\infty} \frac{1}{\left[r^2+x_3{}^2\right]^{3/2}} r\,dr\,d\theta \\
&= \frac{p_0x_3}{4\pi\varepsilon_0}2\pi\int_0^{\infty} \frac{r}{\left[r^2+x_3{}^2\right]^{3/2}}\,dr \\
&= \frac{p_0x_3}{2\varepsilon_0}\left[\frac{-1}{\left(r^2+x_3{}^2\right)^{1/2}}\right]_0^{\infty}
\end{aligned}$$

$$= \frac{p_0 x_3}{2\varepsilon_0} \frac{1}{|x_3|}$$
$$= \frac{p_0}{2\varepsilon_0} \operatorname{sgn}(x_3)$$

and the jump, therefore, is $\dfrac{p_0}{\varepsilon_0}$. ✓

2.6 Conductors

A conductor is a very good ~~man~~ metal. Assume we have a conductor in a domain D, placed into an electrostatic field $\mathbf{E}^{\text{ext}}$.

All the positively charged particles want to move along $\mathbf{E}^{\text{ext}}$, but they're too heavy so they don't. All the negatively charged particles want to move along $-\mathbf{E}^{\text{ext}}$ under the action of the electric field, and they can because they're small. (Hand-waving much?)

Eventually then, inside this conductor, the electric field is 0 for our study of electro*statics*, since all the negatively charged particles at the boundary somehow compete with $\mathbf{E}^{\text{ext}}$, creating $\mathbf{E} = \mathbf{0}$.

So electrons move to the boundary on one side.

To put it another way: in conductors, elementary charges can move freely. As a result, there will be redistribution of these elementary charges inside D and the total electric field becomes zero in D.

As $\mathbf{E}^{\text{total}} = \mathbf{E} + \mathbf{E}^{\text{ext}} = \mathbf{0}$ in D, $\implies$ $\Phi = \text{const}$ inside D.

By the Poisson equation,

$$-\nabla^2 \Phi = \frac{\rho}{\varepsilon_0}$$

i.e. $\rho = 0$ inside D.

Therefore the charges are concentrated on $S = \partial D$, i.e. we have a surface charge $F_{S,\sigma}$, where σ is the charge distribution over the surface.

Using $\mathbf{E} := \mathbf{E}^{\text{total}}$ from now on, recall that for jumps of the electric field across a charged surface,

$$\begin{cases} \left[\mathbf{E}^+(y) - \mathbf{E}^-(y)\right] \cdot \mathbf{n}_y = \dfrac{\sigma(y)}{\varepsilon_0} \\ \left[\mathbf{E}^+(y) - \mathbf{E}^-(y)\right] \cdot \mathbf{t} = 0 \end{cases}$$

for any tangential vector $\mathbf{t}$. But $\mathbf{E}^- = 0$ since $\mathbf{E} = 0$ inside D. Therefore

$$\mathbf{E}^+(y) = \frac{1}{\varepsilon_0} \sigma(y)\, \mathbf{n}_y.$$

Also,

$$\Phi^+(y) = \Phi^-(y)$$

because Φ is continuous across the boundary, for y on the surface. And since Φ is constant on the surface,

$$\Phi^+(y) = \Phi^-(y)\big|_{y\in S} = \text{const}.$$

Now, suppose we have a large conductor D_3 with a big hole cut out of it. Inside the cavity are two small conductors D_1, D_2 which do not touch. Call the cavity $\Omega = \mathbb{R}^3 \setminus (D_1 \cup D_2 \cup D_3)$. And let S_i be the boundary of D_i.

For $x \in \Omega$,

$$\begin{aligned} \Phi|_{S_1} &= c_1 \\ \Phi|_{S_2} &= c_2 \\ \Phi|_{S_3} &= c_3 \end{aligned}$$

Now assume we divide D_1 into two by a thin insulator running somewhere through it. Then the electrostatic potential in the two halves are different. Calling the two boundaries $S_1^\pm$,

$$\begin{aligned} \Phi|_{S_1^+} &= c_1^+ \\ \Phi|_{S_1^-} &= c_1^- \end{aligned}$$

which are different. Keep on dividing D_1 into smaller and smaller pieces, each giving a different value of Φ, i.e. we obtain

$$\Phi|_{\partial\Omega} = \phi,$$

an arbitrary continuous function, where ϕ tells you the charge at a point.

Therefore, when considering $\Phi(\mathbf{x})$, $\mathbf{x} \in \Omega$, we can often assume that we know, for a given ϕ,

$$\Phi|_{\partial\Omega} = \phi \qquad \text{(Dirichlet boundary condition)}$$

but we also know that Φ in Ω satisfies

$$-\nabla^2\Phi = \frac{\rho}{\varepsilon_0}. \qquad \text{(Poisson's equation)}$$

The combination of these two equations is called in mathematics, the *Dirichlet boundary value problem for Poisson's equation.* How original.

Problem 2.13 Let us have a paraboloid sitting centrally on top of the xy-plane, and a potential

$$\Phi(\mathbf{x}) = \begin{cases} \sin x & \text{if } z \geq x^2 + y^2 \\ \sin x + e^z(z - x^2 - y^2) & \text{if } z < x^2 + y^2 \end{cases}$$

1. Find the charge density ρ inside the paraboloid.
2. Find the charge density ρ outside the paraboloid.
3. Find the surface charge density σ on the paraboloid $z - (x^2 + y^2) = 0$.

Recall that if your surface is given by the equation

$$S = \{f(x, y, z) = 0\}$$

then

$$\mathbf{n} = \frac{\nabla f}{|\nabla f|}$$

Solution

1. If we ever want to find out ρ then we want to be using the equation
$$-\nabla^2 \Phi = \frac{\rho}{\varepsilon_0}.$$
Inside the paraboloid, $z \geq x^2 + y^2$ and so $\Phi = \sin x$, hence
$$-\nabla^2 \Phi = -(-\sin x)$$
and so
$$\rho = \varepsilon_0 \sin x.$$
2. Outside the paraboloid, $\Phi = \sin x + e^z(z - x^2 - y^2)$, so
$$-\nabla^2 \Phi = -\left(-\sin x + e^z(z - x^2 - y^2 - 2)\right)$$
and so
$$\rho = \varepsilon_0 \left(\sin x - e^z(z - x^2 - y^2 - 2)\right)$$
3. By equation (2.3), for a surface charge distribution σ,
$$(\mathbf{E}^+ - \mathbf{E}^-) \cdot \mathbf{n} = \frac{\sigma}{\varepsilon_0}$$
and of course we know $\mathbf{E}^+$ and $\mathbf{E}^-$ since we know Φ^+ and Φ^-, if we denote the inside as '+' and the outside as '−'.
$$\begin{aligned} \mathbf{E}^+ - \mathbf{E}^- &= -\nabla \Phi^+ + \nabla \Phi^- \\ &= -\nabla \left(\Phi^+ - \Phi^-\right) \\ &= -\nabla \left(-e^z(z - x^2 - y^2)\right) \\ &= \begin{pmatrix} e^z(-2x) \\ e^z(-2y) \\ e^z(z - x^2 - y^2 + 1) \end{pmatrix} \end{aligned}$$

Now what is $\mathbf{n}$? As we're reminded, if the surface is given by the equation $f(x, y, z) = 0$, which in this case is

$$f(x, y, z) = z - x^2 - y^2 = 0$$

then

$$\mathbf{n} = \frac{\nabla f}{|\nabla f|} = \frac{1}{\sqrt{1 + 4(x^2 + y^2)}} \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix}$$

Calculating,

$$\begin{aligned}(\mathbf{E}^+ - \mathbf{E}^-) \cdot \mathbf{n} &= \frac{1}{\sqrt{1 + 4(x^2 + y^2)}} \begin{pmatrix} e^z(-2x) \\ e^z(-2y) \\ e^z(z - x^2 - y^2 + 1) \end{pmatrix} \cdot \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix} \\ &= \frac{e^z[4x^2 + 4y^2 + z - x^2 - y^2 + 1]}{\sqrt{1 + 4(x^2 + y^2)}}\end{aligned}$$

So

$$\begin{aligned}\sigma &= \varepsilon_0(\mathbf{E}^+ - \mathbf{E}^-) \cdot \mathbf{n} \\ &= \frac{\varepsilon_0 e^z[4x^2 + 4y^2 + z - x^2 - y^2 + 1]}{\sqrt{1 + 4(x^2 + y^2)}} \\ \sigma(x, y) &= \frac{\varepsilon_0 e^{x^2+y^2}[4x^2 + 4y^2 + x^2 + y^2 - x^2 - y^2 + 1]}{\sqrt{1 + 4(x^2 + y^2)}} \\ &= \frac{\varepsilon_0 e^{x^2+y^2}[4x^2 + 4y^2 + 1]}{\sqrt{1 + 4(x^2 + y^2)}}\end{aligned}$$

and that's the answer. A bit ugly but there you go. Not everything in mathematics is beautiful.

✓

2.7 Boundary value problems of electrostatics

Say we have a domain (cavity) $\Omega \subset \mathbb{R}^3$, and we have the following three conditions:

$$\text{(O)} \begin{cases} -\nabla^2\Phi = \dfrac{\rho}{\varepsilon} & \text{in } \Omega \\ \Phi|_{\partial\Omega} = \phi & \text{(given } \phi) \\ \Phi(\mathbf{x}) \to 0 & \text{if } |\mathbf{x}| \to \infty \end{cases}$$

The third condition is strictly not necessary in simple cases, so we "have it with a pinch of salt".

Then this problem actually reduces to the following case. Let Φ_1, Φ_2 solve problems (A) and (B) below:

$$\text{(A)} \begin{cases} -\nabla^2\Phi_1 = \dfrac{\rho}{\varepsilon} & \text{in } \Omega \\ \Phi_1|_{\partial\Omega} = 0 & \\ \Phi_1(\mathbf{x}) \to 0 & \text{if } |\mathbf{x}| \to \infty \end{cases}$$

$$\text{(B)} \begin{cases} -\nabla^2\Phi_2 = 0 & \text{in } \Omega \\ \Phi_2|_{\partial\Omega} = \phi & \text{(given } \phi) \\ \Phi_2(\mathbf{x}) \to 0 & \text{if } |\mathbf{x}| \to \infty \end{cases}$$

Then $\Phi = \Phi_1 + \Phi_2$ solves (O). Nice, huh?

Notice that (B).1 is Laplace's equation: Φ_2 is a harmonic function.

Now, let's take a particular case when $\rho = \delta$. So we're looking for Green's function (where we use G notation as not to confuse ourselves!) corresponding to (A). It is a distribution (actually a discontinuous function) in $\mathcal{D}'(\Omega)$

$$G(\mathbf{x}, \mathbf{y}) \qquad \mathbf{y} \in \Omega \text{ fixed}$$

So we have

$$\text{(A}') \begin{cases} -\nabla^2 G(\mathbf{x}, \mathbf{y}) = \delta_{\mathbf{y}}(\mathbf{x}) & \\ G(\mathbf{x}, \mathbf{y})|_{\mathbf{x}\in\partial\Omega} = 0 & \\ G(\mathbf{x}, \mathbf{y}) \to 0 & \text{if } |\mathbf{x}| \to \infty \end{cases}$$

If $\Omega = \mathbb{R}^3$ then

$$G_0 = \frac{1}{4\pi|\mathbf{x} - \mathbf{y}|}$$

where we use the convention of using the subscript 0 to denote free space, i.e. $\mathbb{R}^3$.

Let

$$\Psi_{\mathbf{y}}(\mathbf{x}) = G(\mathbf{x}, \mathbf{y}) - G_0(\mathbf{x}, \mathbf{y}).$$

Then

$$-\nabla^2\Psi_{\mathbf{y}} = -\nabla^2{}_{\mathbf{x}}G - (-\nabla^2{}_{\mathbf{x}}G_0)$$

(where $\mathbf{y}$ is constant)

$$\begin{aligned} &= \delta_{\mathbf{y}}(\mathbf{x}) - \delta_{\mathbf{y}}(\mathbf{x}) \\ &= 0 \end{aligned}$$

So $-\nabla^2\Psi_{\mathbf{y}} = 0$ in Ω.

Meanwhile, let's have a look what happens on the boundary.

$$\Psi_{\mathbf{y}}(\mathbf{x})|_{\mathbf{x}\in\partial\Omega} = G(\mathbf{x}, \mathbf{y})|_{\mathbf{x}\in\partial\Omega} - G_0(\mathbf{x}, \mathbf{y})|_{\mathbf{x}\in\partial\Omega} \tag{2.6}$$

but G on the boundary is 0

$$= -\left.\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|}\right|_{\mathbf{x}\in\partial\Omega} \tag{2.7}$$

So as $\mathbf{x}\to\infty$, $\Psi_{\mathbf{y}}\to 0$.

So we've reduce the problem (A′) to

$$(\mathrm{B}')\begin{cases} -\nabla^2\Psi_{\mathbf{y}} = 0 & \text{in } \Omega \\ \Psi_{\mathbf{y}}(\mathbf{x})|_{\mathbf{x}\in\partial\Omega} = -\left.\dfrac{1}{4\pi|\mathbf{x}-\mathbf{y}|}\right|_{\mathbf{x}\in\partial\Omega} & \\ \Psi(\mathbf{x},\mathbf{y}) \to 0 & \text{if } |\mathbf{x}|\to\infty \end{cases}$$

Now what about solving Φ for any $\rho \in C_0(\Omega)$? Let

$$\begin{aligned} \Phi &= \frac{1}{\varepsilon_0}\int G(\mathbf{x},\mathbf{y})\rho(\mathbf{y})\ d\mathbf{y} \\ -\nabla^2{}_{\mathbf{x}}(\Phi) &= \frac{1}{\varepsilon_0}\int -\nabla^2{}_{\mathbf{x}}G(\mathbf{x},\mathbf{y})\rho(\mathbf{y})\ d\mathbf{y} \\ &= \frac{1}{\varepsilon_0}\int \delta(\mathbf{x}-\mathbf{y})\rho(\mathbf{y})\ d\mathbf{y} \\ &= \frac{1}{\varepsilon_0}\rho(\mathbf{x}) \end{aligned}$$

So we've proved the following lemma!

Lemma 2.14 Let $G(\mathbf{x},\mathbf{y})$ be known, $\mathbf{x},\mathbf{y}\in\Omega$. Then, for any $\rho\in C_0(\Omega)$,

$$\Phi^\rho(\mathbf{x}) = \frac{1}{\varepsilon_0}\int_\Omega G(\mathbf{x},\mathbf{y})\rho(\mathbf{y})\ d\mathbf{y}$$

solves

$$\begin{cases} -\nabla^2\Phi^\rho = \dfrac{1}{\varepsilon_0}\rho \\ \Phi^\rho|_{\partial\Omega} = 0 \end{cases}$$

2.7.1 To summarise

We had

$$\begin{cases} -\nabla^2\Phi = \dfrac{\rho}{\varepsilon} \\ \Phi|_{\partial\Omega} = 0 \\ \Phi(\mathbf{x}) \to 0 \qquad \text{as } |\mathbf{x}|\to\infty \end{cases}$$

Which we reduced to finding the Green's function

$$\begin{cases} -\nabla^2{}_{\mathbf{x}}G = \delta_{\mathbf{y}} \quad \mathbf{y}\in\Omega \\ G|_{\partial\Omega} = 0 \\ G(\mathbf{x}) \to 0 \qquad \text{as } |\mathbf{x}|\to\infty \end{cases}$$

and then since $G(\mathbf{x}, \mathbf{y}) = G_0(\mathbf{x}, \mathbf{y}) + \Psi_{\mathbf{y}}(\mathbf{x})$, it reduces to find Ψ such that

$$\begin{cases} -\nabla^2{}_{\mathbf{x}} \Psi_{\mathbf{y}}(\mathbf{x}) = 0 & \mathbf{y} \in \Omega \\ \Psi_{\mathbf{y}}(\mathbf{x})|_{\mathbf{x} \in \partial\Omega} = -\dfrac{1}{4\pi|\mathbf{x} - \mathbf{y}|} & \\ \Psi_{\mathbf{y}}(\mathbf{x}) \to 0 & \text{as } |\mathbf{x}| \to \infty \end{cases}$$

and this last set of equations dictate the special Dirichlet boundary value problems for Laplace's equation.

2.8 Dirichlet boundary value problems for Laplace's equation

Now we'll use u instead of Φ so as not to imply $\mathbb{R}^3$, since this result holds for any $\mathbb{R}^n$. We have the system

$$\begin{cases} -\nabla^2 u = 0 & \\ u|_{\mathbf{x} \in \partial\Omega} = \phi & \in C(\partial\Omega) \\ u(\mathbf{x}) \to 0 & \text{as } |\mathbf{x}| \to \infty \end{cases}$$

Recalling that $\nabla^2 u = 0$ in $\Omega \iff u(\mathbf{x})$ is a harmonic function in Ω.

Problem 2.15 $*$ Assume that $u \in C^2(\Omega)$ is harmonic. Show that, in 3D (i.e. $\Omega \subset \mathbb{R}^3$),

$$\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \, dS = 0$$

(hint: use Green's identities)

Solution Green's identity in 3D is, for $\Omega \subset \mathbb{R}^3$,

$$\int_{\Omega} [f\nabla^2 g - g\nabla^2 f] \, dV = \oint_{\partial\Omega} \left[f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right] \, dS$$

Letting $f = 1, g = u$,

$$\int_{\Omega} \nabla^2 u \, dV = \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \, dS$$

And u is harmonic $\iff \nabla^2 u = 0$, so

$$\implies \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} \, dS = 0$$

which is what we were looking for. ✓

Problem 2.16 Assume that $u \in C^2(S)$ is harmonic. Show that, in 2D (i.e. $S \subset \mathbb{R}^2$),

$$\int_{\partial S} \frac{\partial u}{\partial \mathbf{n}} \, d\ell = 0$$

Solution Again, we need to use Green's identity, but we first need to find Green's identity in 2D. Let me write down the identity in 3D again:

$$\int_{\Omega} [f\nabla^2 g - g\nabla^2 f] \, dV = \oint_{\partial\Omega} \left[f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right] \, dS$$

Now let

$$\Omega = S \times \{-h, h\} \qquad f = f(x, y) \qquad g = g(x, y)$$

i.e., turn Ω into a cylinder of height $2h$. What does this mean for our integrals?

$$\oint_{\partial\Omega} = 2h \int_{\partial S} + 2 \int_{S}$$

$$\oint_{\Omega} = 2h \int_{S}$$

So Green's theorem becomes

$$2h \int_{S} [f\nabla^2 g - g\nabla^2 f] \, dS = 2h \int_{\partial S} \left[f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right] \, d\ell + 2 \int_{S} \left[f\frac{\partial g}{\partial \mathbf{z}} - g\frac{\partial f}{\partial \mathbf{z}} \right] \, dS$$

where we've made the observation that $\mathbf{n} = \mathbf{z}$ on the top and bottom of the cylinder. This means that $\frac{\partial g}{\partial \mathbf{z}} = \frac{\partial f}{\partial \mathbf{z}} = 0$, since f and g are functions of x, y alone.

$$= 2h \int_{\partial S} \left[f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right] \, d\ell$$

$$\implies \int_{S} [f\nabla^2 g - g\nabla^2 f] \, dS = \int_{\partial S} \left[f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right] \, d\ell$$

which gives us Green's identity for 2 dimensions.

Now, to solve the problem, just like last time we let $f = 1$ and $g = u$ to give

$$0 = \int_{S} \nabla^2 u \, dS = \int_{\partial S} \frac{\partial u}{\partial \mathbf{n}} \, d\ell$$

once again. ✓

Let us introduce a theorem to help us do more stuff.

Theorem 2.17 (The Mean Value Theorem)

Let u be harmonic in $\Omega \subset \mathbb{R}^3$ (i.e. $\nabla^2 u = 0$), and $B_R(\mathbf{x}_0) \subset \Omega$. Then in the 3D case,

$$u(\mathbf{x}_0) = \frac{1}{4\pi R^2} \int\limits_{\substack{\partial B_R(\mathbf{x}_0) \\ = \mathbb{S}_R(\mathbf{x}_0)}} u(\mathbf{x})\, dS_{\mathbf{x}}$$

and in the 2D case,

$$u(\mathbf{x}_0) = \frac{1}{2\pi R} \int\limits_{C_R(\mathbf{x}_0)} u(\mathbf{x})\, d\ell_{\mathbf{x}}.$$

Proof: (The 3D case)

$$0 = \int\limits_{B_R(\mathbf{x}_0)} \overbrace{\nabla^2 u}^{=0} \left(\frac{-1}{4\pi|\mathbf{x}-\mathbf{x}_0|} \right) d\mathbf{x}$$

which, by Green's formula,

$$\begin{aligned}
&= \underbrace{\int\limits_{B_R(x_0)} u\, \overbrace{\nabla^2 \left(\frac{-1}{4\pi|\mathbf{x}-\mathbf{x}_0|} \right)}^{\delta(\mathbf{x}-\mathbf{x}_0)} d\mathbf{x}}_{u(\mathbf{x}_0)} + \int\limits_{\mathbb{S}_R(\mathbf{x}_0)} \frac{\partial u}{\partial \mathbf{n}} \left(\frac{-1}{4\pi|\mathbf{x}-\mathbf{x}_0|} \right) dS \\
&\quad + \int\limits_{\mathbb{S}_R(x_0)} u \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{4\pi|\mathbf{x}-\mathbf{x}_0|} \right) dS \\
&= u(\mathbf{x}_0) - \frac{1}{4\pi R} \underbrace{\int\limits_{\mathbb{S}_R(\mathbf{x}_0)} \frac{\partial u}{\partial \mathbf{n}} dS}_{=0 \text{ by problem}} + \int\limits_{\mathbb{S}_R(\mathbf{x}_0)} u \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{4\pi|\mathbf{x}-\mathbf{x}_0|} \right) dS \\
&= u(\mathbf{x}_0) + \frac{1}{4\pi} \int\limits_{\mathbb{S}_R(\mathbf{x}_0)} u(\mathbf{x}) \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{|\mathbf{x}-\mathbf{x}_0|} \right) dS \\
&= u(\mathbf{x}_0) + \frac{1}{4\pi} \int\limits_{\mathbb{S}_R(\mathbf{x}_0)} u(\mathbf{x}) \frac{\partial}{\partial r} \left(\frac{1}{r} \right) \bigg|_{r=|\mathbf{x}-\mathbf{x}_0|} dS \\
0 &= u(\mathbf{x}_0) + \frac{1}{4\pi} \int\limits_{\mathbb{S}_R(\mathbf{x}_0)} u(\mathbf{x}) \left(-\frac{1}{r^2} \right) dS
\end{aligned}$$

and, by rearranging

$$\begin{aligned}
\implies u(\mathbf{x}_0) &= \frac{1}{4\pi} \int u(\mathbf{x}) \frac{1}{R^2} dS \\
&= \frac{1}{4\pi R^2} \int u(\mathbf{x})\, dS
\end{aligned}$$

□

Problem 2.18 $*$ Prove the theorem for the 2D case.

Solution The fundamental solution to the equation $\nabla^2 G = \delta$ is:

$$G(\mathbf{x}) = \frac{1}{2\pi} \ln|\mathbf{x}|$$

So we shall use Green's identity in 2D as proved in problem 2.16. Once again, that is

$$\int_S [f\nabla^2 g - g\nabla^2 f]\, dS = \int_{\partial S} \left[f\frac{\partial g}{\partial \mathbf{n}} - g\frac{\partial f}{\partial \mathbf{n}} \right] d\ell$$

Letting $g = u$, $f = G$,

$$\begin{aligned}
0 &= \int_{S_R(\mathbf{x}_0)} \overbrace{\nabla^2 u}^{=0} \left(\frac{\ln|\mathbf{x}-\mathbf{x_0}|}{2\pi} \right) d\mathbf{x} \\
&= \underbrace{\int_{S_R} u\, \overbrace{\nabla^2 \left(\frac{\ln|\mathbf{x}-\mathbf{x_0}|}{2\pi} \right)}^{\delta(\mathbf{x}-\mathbf{x}_0)} d\mathbf{x}}_{u(\mathbf{x}_0)} + \int_{\partial S} \frac{\partial u}{\partial \mathbf{n}} \left(\frac{\ln|\mathbf{x}-\mathbf{x_0}|}{2\pi} \right) d\ell \\
&\quad - \int_{\partial S} u \frac{\partial}{\partial \mathbf{n}} \left(\frac{\ln|\mathbf{x}-\mathbf{x_0}|}{2\pi} \right) d\ell \\
&= u(\mathbf{x}_0) - \frac{1}{2\pi}\ln(R) \underbrace{\int_{\partial S} \frac{\partial u}{\partial \mathbf{n}} d\ell}_{=0 \text{ by prob 2.16}} - \int_{\partial S} u \frac{\partial}{\partial \mathbf{n}} \left(\frac{\ln|\mathbf{x}-\mathbf{x_0}|}{2\pi} \right) d\ell \\
&= u(\mathbf{x}_0) - \frac{1}{2\pi} \int_{\partial S} \frac{u(\mathbf{x})}{|\mathbf{x}-\mathbf{x_0}|} d\ell \\
&= u(\mathbf{x}_0) - \frac{1}{2\pi R} \int_{\partial S} u(\mathbf{x})\, d\ell
\end{aligned}$$

and so, by rearranging

$$\implies u(\mathbf{x}_0) = \frac{1}{2\pi R} \int_{C_R(\mathbf{x}_0)} u(\mathbf{x})\, d\ell$$

which proves the theorem. ✓

Problem 2.19 Prove this version of the Mean Value Theorem: Let u be harmonic in $\Omega \subset \mathbb{R}^3$ (i.e. $\nabla^2 u = 0$), and $B_R(\mathbf{x}_0) \subset \Omega$. Then

$$u(\mathbf{x}_0) = \frac{1}{\frac{4\pi}{3}R^3} \int_{B_R(\mathbf{x_0})} u(\mathbf{x})\, dV$$

Solution If you consider integrating over the volume as being equivalent to integrating over all the surfaces of balls with radius 0 through to R, we have

$$\begin{aligned}
\int\limits_{B_R(\mathbf{x_0})} u(\mathbf{x})\,dV &= \int_{r=0}^{R}\left[\,\int\limits_{S_r(\mathbf{x}_0)} u(\mathbf{x})\,dS\right]dr \\
&= \int_{r=0}^{R} 4\pi r^2 u(\mathbf{x}_0)\,dr \\
&= 4\pi u(\mathbf{x}_0)\frac{R^3}{3} \\
\implies u(\mathbf{x}_0) &= \frac{1}{\frac{4\pi}{3}R^3}\int\limits_{B_R(\mathbf{x}_0)} u(\mathbf{x})\,dV
\end{aligned}$$

as desired. ✓

Time for another theorem.

Theorem 2.20 (The Maximum Principle) Let us have a bounded domain $D \subset \mathbb{R}^3$. On this domain, let us have a function $u \in C^2(D^{\text{int}}) \cap C(D)$ which has the property $\nabla^2 u = 0$. That is to say, let u be harmonic and D-bounded.

If we let $M = \max\limits_{\mathbf{x}\in\partial D} u(\mathbf{x})$, then

$$u(\mathbf{x}) \leq M \quad \forall \mathbf{x} \in D$$

Moreover, if D is connected, and there exists a point $\mathbf{x}_0 \in D^{\text{int}}$ such that $u(\mathbf{x}_0) = M$, then

$$u(\mathbf{x}) \equiv M \quad \forall \mathbf{x} \in D.$$

Proof: Assume that there exists a point $\mathbf{x}_0 \in D^{\text{int}}$ such that

$$u(\mathbf{x}_0) = \max_{\mathbf{x}\in D} u(\mathbf{x}) := A$$

Let $d = \operatorname{dist}(\mathbf{x}_0, \partial D) > 0$. Now consider a ball of this radius at the point $\mathbf{x_0}$, i.e. $B_d(\mathbf{x}_0) \subset D$. By Problem 2.19,

$$A = u(\mathbf{x}_0) = \frac{1}{\frac{4\pi}{3}d^3}\int\limits_{B_d} u(\mathbf{x})\,d\mathbf{x}.$$

By the definition of A, we have that $u(\mathbf{x}) \leq A$ in B_d. Therefore

$$\frac{1}{\frac{4\pi}{3}d^3}\int\limits_{B_d} u(\mathbf{x})\,d\mathbf{x} \leq A \cdot \underbrace{\frac{1}{\frac{4\pi}{3}d^3}\int\limits_{B_d} d\mathbf{x}}_{=\frac{\text{vol}}{\text{vol}}=1} = A$$

If $u(\mathbf{y}) < A$ for some $y \in B_d$ (and therefore nearby), then

$$\frac{1}{\frac{4\pi}{3}d^3}\int\limits_{B_d} u(\mathbf{x})\,d\mathbf{x} < A$$

Therefore, assuming that there exists a point $\mathbf{y} \in B_d$ such that $u(\mathbf{y}) < A$ we get

$$A = u(\mathbf{x}_0) = \frac{1}{\frac{4\pi}{3}d^3} \int\limits_{B_d} u(\mathbf{x})\, d\mathbf{x} < A$$

which is clearly a **contradiction**.

Therefore $u(\mathbf{x}) \equiv A$ inside B_d. Now for a bit of tidying up to complete the proof of the first part.

As $d = \mathrm{dist}(\mathbf{x}_0, \partial B)$, this implies that $S_d \cap \partial D \neq \emptyset$. Therefore there exists a point $\mathbf{z}_0 \in \partial D$ such that $u(\mathbf{z}_0) = A = \max_{\mathbf{x} \in D} u(\mathbf{x})$.

Therefore

$$M = \max_{\mathbf{z} \in \partial D} u(\mathbf{z}) = \max_{\mathbf{x} \in D} u(\mathbf{x}) = A$$

which proves the first part. ■ (part 1)

Now, let's again have

$$u(\mathbf{x}_0) = \max_{\mathbf{x} \in D} u(\mathbf{x}) = A$$

Take any $\mathbf{y} \in D$ and connect $\mathbf{x}_0$ and $\mathbf{y}$ by a curve γ lying in D^{int}. Then let

$$\rho = \mathrm{dist}(\gamma, \partial D) > 0.$$

Assume $\mathbf{y}$ is not in the ball, since this would makes things trivial: $\mathbf{y} \in B_d \implies u(\mathbf{y}) = A$.

Let $\mathbf{x}_1$ be the last point on γ and S_d. Then

$$u(\mathbf{x}_1) = A = \max_{\mathbf{x} \in D} u(\mathbf{x}).$$

So then take a ball of radius ρ (so as it doesn't touch the boundary) centred at $\mathbf{x}_1$, $B_\rho(\mathbf{x}_1)$. The same arguments show that $u(\mathbf{x}) = A$ in $B_\rho(\mathbf{x}_1)$.

We continue this process until $\mathbf{y} \in B_\rho(\mathbf{x}_n)$. We need only a finite number of steps as γ has finite length. This proves the second part which proves the whole thing. □

Problem 2.21 Prove, either by rewriting the above proof, or as a corollary, the Minimum Principle Theorem, i.e.

$$\min_{\mathbf{z} \in \partial D} u(\mathbf{z}) = \min_{\mathbf{x} \in D} u(\mathbf{x})$$

Solution It falls trivially out of the Maximum Principle Theorem if we say

$$\widetilde{u}(\mathbf{x}) = -u(\mathbf{x}).$$

✓

There is, however, a stronger form of the theorem, but it needs these definitions first:

Definition 2.22 A point $\mathbf{x}_0$ is a *local maximum* of $u(\mathbf{x})$ if there exists $B_\varepsilon(\mathbf{x}_0)$ such that

$$u(\mathbf{x}_0) = \sup_{\mathbf{x}\in B_\varepsilon(x_0)} u(\mathbf{x})$$

Definition 2.23 A point $\mathbf{x}_0$ is a *local minimum* of $u(\mathbf{x})$ if there exists $B_\varepsilon(\mathbf{x}_0)$ such that

$$u(\mathbf{x}_0) = \inf_{\mathbf{x}\in B_\varepsilon(x_0)} u(\mathbf{x})$$

And the strong version of the theorem is:

Theorem 2.24 (Strong version of the theorem) If D is connected and $u(\mathbf{x}) \not\equiv$ const and $\nabla^2 u = 0$ then u has no local maxima and minima inside D^{int}.

We shan't prove this theorem.

2.8.1 Now, gentlemen, we move to physics

Let's have a domain D with no charges inside it. So

$$\nabla^2 \Phi = 0 \quad \text{in } D$$

Introduce a point charge e into the domain. Then $e\Phi$ is the electrostatic (potential) energy of the charge e.

Now recall that if we have Φ, then we also have $\mathbf{E}$ such that $\mathbf{E} = -\nabla\Phi$.

Then there is a force $\mathbf{F} = e\mathbf{E}$. This electrostatic force tries to decrease the electrostatic energy of e. And since, by problem 2.21, $\min e\Phi$ is achieved on ∂D, the charge ends up on the boundary.

2.8.2 Unbounded domains

Now we shall take a look at what happens in unbounded domains. Let's have

$$\begin{cases} \nabla^2 u = 0 & \text{in } D \\ u\big|_{\mathbf{x}\in\partial\Omega} = \phi & \in C(\partial\Omega) \\ u(\mathbf{x}) \to 0 & \text{as } |\mathbf{x}| \to \infty \end{cases}$$

We shall consider the unbounded domain as the limit of the following set. Let

$$D_R = D \cap B_R.$$

By the maximum principle theorem,

$$u(\mathbf{x})\big|_{\mathbf{x}\in D_R} \le \max_{\mathbf{z}\in\partial D_R} u(\mathbf{z}).$$

Now what is the boundary of D_R?

$$\partial D_R = (\partial D \cap B_R) \cup (S_R \cap D)$$

So

$$\begin{aligned} u(\mathbf{x})\big|_{\mathbf{x}\in D_R} &\le \max\left[\max_{\mathbf{z}\in\partial D\cap B_R} u(\mathbf{z}),\ \max_{\mathbf{x}\in S_R\cap D} u(\mathbf{x})\right] \\ &\le \max\left[\max_{\mathbf{z}\in\partial D} u(\mathbf{z}),\ \max_{\mathbf{x}\in S_R\cap D} u(\mathbf{x})\right] \end{aligned}$$

since the first term is now the maximum of a smaller set. Note that the first term then *does not depend* on R. And what happens as $R \to \infty$? The second term goes to 0, since $u \to 0$ as $\mathbf{x} \to \infty$.

$$\implies u(\mathbf{x})\big|_{\mathbf{x}\in D} \le \max\left[\max_{\mathbf{z}\in\partial D} u(\mathbf{z}), 0\right].$$

But we can do the same thing using the minimum principle theorem, which gives us

$$u(\mathbf{x})\big|_{\mathbf{x}\in D} \ge \min\left[\min_{\mathbf{z}\in\partial D} u(\mathbf{z}), 0\right].$$

So in the case when $D = \mathbb{R}^3$, then $\partial D = \emptyset$, which implies that

$$u(\mathbf{x})\big|_{\mathbf{x}\in\mathbb{R}^3} = 0.$$

2.8.3 Uniqueness of solutions for Dirichlet boundary conditions in bounded domains

Let us have a bounded domain D once again. Then the Dirichlet boundary conditions are

$$\begin{cases} \nabla^2 u = 0 & \text{in } D \\ u\big|_{\mathbf{x}\in\partial\Omega} = \phi & \in C(\partial\Omega) \end{cases}$$

Proposition 2.25 For any bounded D any any $\phi \in C(\partial D)$, there exists a unique $u^\phi \in C^2(D^{\text{int}}) \cap C(D)$ which solves the Dirichlet boundary conditions.

We can't prove the existence: this requires potential theory for Dirichlet problems for Laplace's equation. But we can prove uniqueness:

Proof: Uniqueness Assume we have two functions u_1^ϕ, u_2^ϕ which solve the Dirichlet boundary conditions. Let

$$u = u_1^\phi - u_2^\phi.$$

Then

$$\begin{cases} \nabla^2 u = 0 & \text{in} D \\ u|_{\mathbf{x} \in \partial\Omega} = 0 \end{cases}$$

By the maximum principle, $u(\mathbf{x}) \leq 0$ for $\mathbf{x} \in D$.

By the minimum principle, $u(\mathbf{x}) \geq 0$ for $\mathbf{x} \in D$. Therefore

$$u \equiv 0 \implies u_1^\phi = u_2^\phi$$

which is what we want. □

Now what happened to distributions, you ask. Let's look at $\mathcal{D}'_\phi(D) \subset \mathcal{D}'(D)$, i.e. those distributions which have a 'meaningful' restriction on ∂D which is equal to ϕ.

If $u \in \mathcal{D}'_\phi(D)$ is harmonic, then $\nabla^2 u = 0$, i.e. $u \in C^\infty(D^{\text{int}})$.

Chapter 3

Magnetism

3.1 The laws for magnetostatics

The Danish physicist Hans Christian Ørsted noticed during a lecture in 1820 that if we have an electric current flowing through a circuit, it affects a magnetic dipole, i.e. there is a force which acts on magnets.

So let us have two electric circuits with two currents I_1 and I_2. Turn on I_1, then turn on I_2. Then there is a force applied to the first circuit.

Taking a small part of each circuit,

$$I_1 \, d\mathbf{l}_1 \qquad I_2 \, d\mathbf{l}_2$$

and defining $\mathbf{x}_{12}$ as the vector distance from piece 2 to piece 1, they found the relation for the force by 2 on 1 as

$$\mathbf{F}_{1\leftarrow 2} = \frac{\mu_0}{4\pi} \frac{(I_1 \, d\mathbf{l}_1) \times (I_2 \, d\mathbf{l}_2 \times \mathbf{x}_{12})}{|\mathbf{x}_{12}|^3}$$

3.1.1 Gauss's law for magnetism

Definition 3.1 The *Biot-Savart law* is an equation in electromagnetism that describes the magnetic field $\mathbf{B}$ generated by an electric current. (In English literature, the term 'magnetic field' is commonplace, foreign literature uses the phrase 'magnetic flux density'.) It's given by

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{I_2 \, d\mathbf{l}_2 \times (\mathbf{x} - \mathbf{x}_2)}{|\mathbf{x} - \mathbf{x}_2|^3}$$

Definition 3.2 Suppose we have a flow of charges $\mathbf{J}(\mathbf{y})$. Then

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{y}) \times (\mathbf{x} - \mathbf{y})}{|\mathbf{x} - \mathbf{y}|^3} \, dV_{\mathbf{y}}$$

where we can say

$$\mathbf{J}(\mathbf{y}) = \dot{\mathbf{y}}(s)\delta_{I_2,C}$$

where

$$\begin{aligned} C &= \{\mathbf{y} = \mathbf{y}(s)\}, \text{ the curve around the circuit with current } I_2 \\ \dot{\mathbf{y}}(s) & \text{ is tangent to the curve} \\ \delta_{I_2,C} & \text{ is like } \delta_{h,L} \text{ in Problem 1.42} \\ I_2 & \text{ is constant} \end{aligned}$$

so

$$\dot{\mathbf{y}}(s) \, ds = d\mathbf{l}$$

Then

$$\mathbf{B}(\mathbf{x}) = \mathbf{J} *_{\times} \nabla\Phi$$

(this is vector-valid convolution, i.e. it's valid for each component of $\mathbf{J}_i$.)

Let us have an electric field $\mathbf{E}(\mathbf{x})$. If we have a charge q, in electrostatics,

$$\mathbf{F}(\mathbf{x}) = q\mathbf{E}(\mathbf{x}).$$

If $\rho(\mathbf{x})$ is an electric charge distribution (i.e. we don't have point charges but a continuous distribution instead), then

$$\mathbf{F}^{\text{total}} = \int_D \rho(\mathbf{x})\mathbf{E}(\mathbf{x}) \, dV.$$

Now suppose we have a body with some $\widetilde{\mathbf{J}}(\mathbf{x})$ inside. Then

$$\mathbf{F}^{\text{total}}_{\text{mag}} = \int_D \widetilde{\mathbf{J}}(\mathbf{x}) \times \mathbf{B}(\mathbf{x}) \, dV$$

which is *Ampère's law of magnetic forces*. (Ampère was a big boy: he had lots of laws)

But first we need to review some vector identities.

Problem 3.3 Show that

$$\nabla\times(\mathbf{a}f(\mathbf{x})) = -\mathbf{a}\times\nabla f$$

Problem 3.4 Show that

$$\nabla\times\nabla\times\mathbf{F} = \nabla(\nabla\cdot\mathbf{F}) - \nabla^2\mathbf{F}$$

Solution In fact, this is just a rearrangement of the *definition* of the vector Laplacian $\nabla^2\mathbf{F}$. But you can show it explicitly in 3D by letting $\mathbf{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$, if you like wasting your time. ✓

Problem 3.5 Show that

$$\nabla \cdot (\mathbf{a}f(\mathbf{x})) = \mathbf{a} \cdot \nabla f$$

Problem 3.6 Show that

$$\int \mathbf{F} \cdot \nabla f \; dV = -\int (\nabla \cdot \mathbf{F}) f \; dV$$

(hint: use the divergence theorem, and that $\nabla \cdot (\mathbf{F}f) = \mathbf{F} \cdot \nabla f + (\nabla \cdot \mathbf{F})f$).

Now we're going to take $\mathbf{B}(\mathbf{x})$ from Definition 3.2 and manipulate it. First, note that

$$-\frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^3} = \nabla \frac{1}{|\mathbf{x}-\mathbf{y}|}$$

and that $\mathbf{J}(\mathbf{y})$ is constant with respect to $\mathbf{x}$. Let $\mathbf{J}(\mathbf{y}) \in \mathcal{D}(\mathbb{R}^3)$. Then

$$\begin{aligned} \mathbf{B}(\mathbf{x}) &= \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{y}) \times (\mathbf{x}-\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^3} \, dV_{\mathbf{y}} \\ &= \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \mathbf{J}(\mathbf{y}) \times \nabla \frac{-1}{|\mathbf{x}-\mathbf{y}|} \, dV_{\mathbf{y}} \end{aligned}$$

and by Problem 3.3,

$$= \frac{\mu_0}{4\pi} \nabla \times \int_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \, dV_{\mathbf{y}} \tag{3.1}$$

$$= \nabla \times \left[\frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \, dV_{\mathbf{y}} \right]$$

so $\mathbf{B}$ is the curl of some magnetic field. So it follows, since $\operatorname{div}(\operatorname{curl}(\cdot)) = 0$,

$$\nabla \cdot \mathbf{B} = 0$$

Definition 3.7 Gauss's law for magnetism is

$$\nabla \cdot \mathbf{B} = 0$$

3.1.2 Ampère's law

But what is $\nabla\times\mathbf{B}$? Combining the identity from Problem 3.4 with equation line (3.1) above,

$$\begin{aligned}\nabla\times\mathbf{B} &= \frac{\mu_0}{4\pi}\int\limits_{\mathbb{R}^3}\left[\nabla_x\times\nabla_x\left(\frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\right)\right]dV_{\mathbf{y}}\\ &= \underbrace{\frac{\mu_0}{4\pi}\int\limits_{\mathbb{R}^3}\left[\nabla_x\left(\nabla_x\cdot\frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\right)\right]dV_{\mathbf{y}}}_{B}\underbrace{-\frac{\mu_0}{4\pi}\int\limits_{\mathbb{R}^3}\left[{\nabla^2}_x\left(\frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\right)\right]dV_{\mathbf{y}}}_{A}\end{aligned}$$

Looking at the terms A and B in turn,

$$\begin{aligned}A &= -\frac{\mu_0}{4\pi}\int\limits_{\mathbb{R}^3}\left[\mathbf{J}(\mathbf{y}){\nabla^2}_x\left(\frac{1}{|\mathbf{x}-\mathbf{y}|}\right)\right]dV_{\mathbf{y}}\\ &= \mu_0\int\limits_{\mathbb{R}^3}\mathbf{J}(\mathbf{y})\delta_{\mathbf{x}-\mathbf{y}}dV_{\mathbf{y}}\\ &= \mu_0\mathbf{J}(\mathbf{x})\end{aligned}$$

and

$$\begin{aligned}B &= \int\limits_{\mathbb{R}^3}\left[\nabla_x\left(\nabla_x\cdot\frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\right)\right]dV_{\mathbf{y}}\\ &= \nabla_x\left[\int\limits_{\mathbb{R}^3}\nabla_x\cdot\frac{\mathbf{J}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}dV_{\mathbf{y}}\right]\end{aligned}$$

which, by Problem 3.5 is

$$\begin{aligned}&= \nabla_x\int\limits_{\mathbb{R}^3}\mathbf{J}(\mathbf{y})\cdot\nabla_x\left(\frac{1}{|\mathbf{x}-\mathbf{y}|}\right)dV_{\mathbf{y}}\\ &= -\nabla_x\int\limits_{\mathbb{R}^3}\mathbf{J}(\mathbf{y})\cdot\nabla_y\left(\frac{1}{|\mathbf{x}-\mathbf{y}|}\right)dV_{\mathbf{y}}\end{aligned}$$

which, by Problem 3.6 is

$$= \nabla_x\left[\int\limits_{\mathbb{R}^3}(\underbrace{\nabla\cdot\mathbf{J}}_{0})(\mathbf{y})\cdot\frac{1}{|\mathbf{x}-\mathbf{y}|}dV_{\mathbf{y}}\right] \tag{3.2}$$

claiming that in magneto-electrostatics, $\nabla\cdot\mathbf{J}=0$, (we prove this ahead in Lemma 3.9)

$$= 0$$

which gives us

$$\nabla\times\mathbf{B} = \mu_0\mathbf{J}(\mathbf{x})$$

Definition 3.8 Ampère's law is

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}(\mathbf{x})$$

Now we prove the claim:

Lemma 3.9 In magneto-electrostatics,

$$\nabla \cdot \mathbf{J} = 0$$

Proof: This claim really comes from the philosophy of life: nothing comes from nothing. We don't want *deus ex machina* here, else we should shut our doors and go to church.

Take a region V and look for the total charge Q in the region

$$Q(t) = \int_V \rho(\mathbf{x}, t)\, dV_{\mathbf{x}}.$$

We believe that charge doesn't come from nothing. If the charge changes, then there is some charge coming in and out of the region: i.e. if Q as a function of time changes, there should be an in/outflux of charge in the region. Mathematically, if $\frac{\partial Q}{\partial t} \neq 0$, there should be a *current* flux through $S = \partial V$. But this current flux is

$$\int_S \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}\, dS.$$

So the continuity law is essentially, before we tidy it up,

$$\int_V \frac{\partial \rho}{\partial t}(\mathbf{x}, t)\, dV = -\int_S \mathbf{J}(t) \cdot \mathbf{n}\, dS$$

where we have a minus sign since if ρ increases, $\mathbf{J}$ should be pointing inwards. Using the divergence theorem on the right-hand side,

$$\int_V \frac{\partial \rho}{\partial t}(\mathbf{x}, t)\, dV = -\int_V \nabla \cdot \mathbf{J}(t)\, dV$$

and since V is arbitrary,

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{J}(\mathbf{x}, t)$$

which of course we can rewrite in the more usual form,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \qquad \text{(Continuity equation)}$$

Now since we're talking about *statics* at the moment, $\frac{\partial \rho}{\partial t} = 0$, so we get

$$\nabla \cdot \mathbf{J} = 0$$

which is what we claimed. $\square$

So we've derived the four laws of electromagnetostatics. In differential form, they are

1. $\nabla \cdot \mathbf{E} = \dfrac{\rho}{\varepsilon_0}$ Gauss's law (Def$^{\text{n}}_{\cdot}$ 2.5)
2. $\nabla \times \mathbf{E} = 0$ (Def$^{\text{n}}_{\cdot}$ 2.5)
3. $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ Ampère's law (Def$^{\text{n}}_{\cdot}$ 3.8)
4. $\nabla \cdot \mathbf{B} = 0$ Gauss's law for magnetism (Def$^{\text{n}}_{\cdot}$ 3.7)

In integral form, they are

1. $\displaystyle\int\limits_{S=\partial V} \mathbf{E} \cdot \mathbf{n}\, dS = \frac{1}{\varepsilon_0} \int\limits_V \rho\, dV$
2. $\displaystyle\oint\limits_C \mathbf{E} \cdot d\mathbf{r} = 0$
3. $\displaystyle\oint\limits_{C=\partial S} \mathbf{B} \cdot d\mathbf{r} = \mu_0 \int\limits_S \mathbf{J} \cdot \mathbf{n}\, dS$
4. $\displaystyle\int_S \mathbf{B} \cdot \mathbf{n}\, dS = 0$

3.2 The laws of magnetodynamics

3.2.1 Maxwell–Ampère law

In the dynamical case, $\nabla \cdot \mathbf{J} \neq 0$ like we used in Lemma 3.9. Instead,

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}.$$

What does this mean for our equations? Recall in equation (3.2) we had

$$\nabla \times \mathbf{B} = -\frac{\mu_0}{4\pi} \nabla_{\mathbf{x}} \frac{\partial}{\partial t} \int \frac{\rho(\mathbf{y}, t)}{|\mathbf{x} - \mathbf{y}|}\, dV_{\mathbf{y}} + \mu_0 \mathbf{J}(\mathbf{x}).$$

But what is

$$\frac{1}{4\pi\varepsilon_0} \int \frac{\rho(\mathbf{y}, t)}{|\mathbf{x} - \mathbf{y}|}\, dV_{\mathbf{y}}?$$

What is that? *What is that?* It's $\Phi^\rho(\mathbf{x}, t)$! So

$$\begin{aligned}\nabla \times \mathbf{B} &= \varepsilon_0 \mu_0 \frac{\partial}{\partial t}(-\nabla_{\mathbf{x}} \Phi^\rho) + \mu_0 \mathbf{J}(\mathbf{x}) \\ &= \varepsilon_0 \mu_0 \frac{\partial}{\partial t} \mathbf{E}(\mathbf{x}, t) + \mu_0 \mathbf{J}(\mathbf{x})\end{aligned}$$

and so by rearranging we get

$$\nabla\times(\mu_0^{-1}\mathbf{B}) = \frac{\partial}{\partial t}(\varepsilon_0\mathbf{E}) + \mathbf{J} \qquad \text{(Maxwell-Ampère eqn in vacuum)}$$

What are μ_0, ε_0? In the vacuum (or any *isotropic* material), they are just some constants. But if we go to the *brick*, we have to replace them by *matrices* ε, μ. These

Figure 3.1: ε and μ are no longer constants if we go to the brick

matrices are different for each material, and they depend on $\mathbf{x}$ and t. Pretty cool, huh?

We can define two new things to make our equations a bit nicer.

Definition 3.10 The *magnetic* **H** *field*, or *magnetic field strength* is defined as

$$\mathbf{H} := \mu^{-1}\mathbf{B}$$

Definition 3.11 The *electric displacement in a vacuum* is

$$\mathbf{D} := \varepsilon\mathbf{E}$$

So the Maxwell–Ampère equation becomes...

Definition 3.12 The Maxwell–Ampère equation is

$$\nabla\times\mathbf{H} = \frac{\partial\mathbf{D}}{\partial t} + \mathbf{J}$$

and this is for *any* material.

3.2.2 Maxwell–Faraday equation

How does the statics equation $\nabla\times\mathbf{E} = \mathbf{0}$ change for dynamics? In the 1830s, Michael Faraday (of 1991 £20 note fame) carried out his famous experiment. He took a closed (i.e. circular) piece of wire and he produced a magnetic flux through it by placing another wire with a current flowing through it next to it (by the Biot-Savart law). If he changed the magnetic field, there appeared a current in the wire. We change the flux, which creates a force through it, proportional to the flux. What produces forces? Electric fields! So he concluded that when you change an electric field in time, you change a magnetic field. After all these fantastic experiments he found that

Definition 3.13 The Maxwell–Faraday equation is

$$\nabla\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

Now, gentlemen and ladies, we shall show that all of Faraday's work (and the two knighthoods which he declined!) was unnecessary since we shall derive the Maxwell–Faraday law from Ampère's law alone.

Let us have a circuit $\mathcal{C}$ as given in Figure 3.2. The top bar (U) is moving upwards at

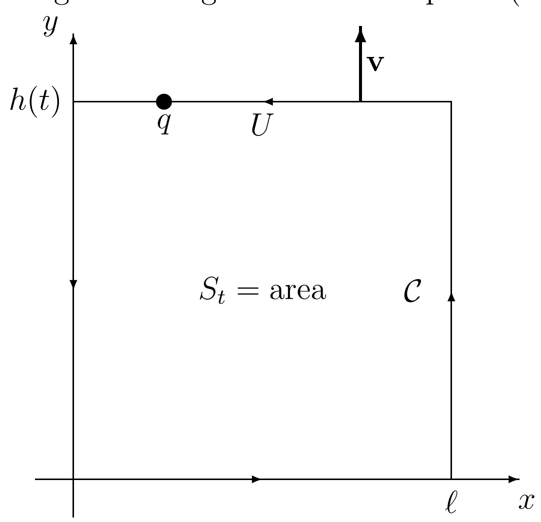

Figure 3.2: The circuit we're looking at

a speed $v = h'(t)$. Let $\mathbf{B} = B_0\widehat{\mathbf{k}}$, where B_0 is the magnetic field of the Earth. Take a charge q along the bar U at height $h(t)$.

$$\mathbf{J} = q\mathbf{v} = qv\widehat{\mathbf{j}}$$

Then by Ampère's law of forces,

$$\mathbf{F} = q\mathbf{v}\times\mathbf{B} = qvB_0\widehat{\mathbf{i}}.$$

But we know that $\mathbf{F} = q\mathbf{E}$, so

$$\mathbf{E} = \mathbf{v}\times\mathbf{B} = vB_0\widehat{\mathbf{i}}.$$

Now let's have a look at the integral around the circuit.

$$\oint_{\mathcal{C}(t)} \mathbf{E}\cdot d\mathbf{r} = \int_U vB_0\widehat{\mathbf{i}}\cdot d\mathbf{r}$$
$$= -vB_0\ell \tag{3.3}$$

and

$$\begin{aligned}
\int_{S_t} \mathbf{B} \cdot \mathbf{n} \, dS &= \int_{S_t} \mathbf{B} \cdot \widehat{\mathbf{k}} \, dS \\
&= \int_{S_t} B_0 \widehat{\mathbf{k}} \cdot \widehat{\mathbf{k}} \, dS \\
&= B_0 \int_{S_t} dS \\
&= B_0 \, \text{area}(S_t) \\
&= B_0 \ell h(t) & (3.4) \\
\implies \frac{\partial}{\partial t} \int_{S_t} \mathbf{B} \cdot \mathbf{n} \, dS &= \frac{d}{dt} B_0 \ell h(t) \\
&= B_0 \ell v & (3.5)
\end{aligned}$$

And by equating (3.3) and (3.5),

$$-\oint_{\mathcal{C}(t)} \mathbf{E} \cdot d\mathbf{r} = \frac{\partial}{\partial t} \int_{S_t} \mathbf{B} \cdot \mathbf{n} \, dS$$

or putting in the dependencies,

$$-\oint_{\mathcal{C}(t)} \mathbf{E}(\mathbf{x}, t) \cdot d\mathbf{r} = \frac{\partial}{\partial t} \int_{S_t} \mathbf{B}(t) \cdot \mathbf{n} \, dS. \tag{3.6}$$

Now here comes the trick. Say the change is not because of area, but instead because of changing flux. By Stokes' law, the left hand side of equation (3.6) is equal to

$$-\oint_{\mathcal{C}(t)} \mathbf{E}(\mathbf{x}, t) \cdot d\mathbf{r} = \int_{S} (-\nabla \times \mathbf{E}) \cdot \mathbf{n} \, dS$$

and because S is not dependent on t, the right hand side of equation (3.6) is equal to

$$\frac{\partial}{\partial t} \int_{S_t} \mathbf{B}(t) \cdot \mathbf{n} \, dS = \int_{S_t} \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n} \, dS$$

i.e.

$$\int_{S} (-\nabla \times \mathbf{E}) \cdot \mathbf{n} \, dS = \int_{S_t} \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n} \, dS$$

And by the lemma we're about to prove, since S is arbitrary, it's safe to say that

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

Lemma 3.14 The step above is valid.

Proof: Basically this step is valid because it's a version of the DuBois-Reymond lemma. We start with

$$\int \left(\nabla\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t}\right)\cdot\mathbf{n}\,dS = 0 \tag{3.7}$$

and say that this is

$$\iff \mathbf{F} := \nabla\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = \mathbf{0}.$$

Why? Assume it is not zero at some point. Then there exists a point $\mathbf{x}_0$ such that $\mathbf{F}(\mathbf{x}_0) \neq 0$.

Take a plane Π perpendicular to $\mathbf{F}$ through the point $\mathbf{x}_0$, and call the disc of radius ε in Π, S_ε. We know that, by equation (3.7),

$$\int\limits_{S_\varepsilon} \mathbf{F}\cdot\widehat{\mathbf{n}}\,dS = 0.$$

But by our clever choice of plane,

$$\widehat{\mathbf{n}} = \frac{\mathbf{F}(\mathbf{x}_0)}{|\mathbf{F}(\mathbf{x}_0)|}$$

And so

$$\mathbf{F}\cdot\widehat{\mathbf{n}}\big|_{\mathbf{x}=\mathbf{x}_0} = |\mathbf{F}(\mathbf{x}_0)| > 0.$$

But $\mathbf{F}$ is continuous on Π, so

$$\mathbf{F}\cdot\widehat{\mathbf{n}} > \frac{1}{2}|\mathbf{F}(\mathbf{x}_0)|$$

if ε is sufficiently small. Thus

$$\begin{aligned}\int_{S_\varepsilon} \mathbf{F}\cdot\widehat{\mathbf{n}}\,dS &> \frac{1}{2}|\mathbf{F}(\mathbf{x}_0)|\,\mathrm{area}(S_\varepsilon)\\ &> 0\end{aligned}$$

which is a **contradiction**. Hence $\mathbf{F}(\mathbf{x}) = \mathbf{0}\ \forall\mathbf{x}$. ☐

So we've proved the Maxwell–Faraday equation another way.

3.3 Maxwell's equations

And so we have derived Maxwell's equations in any material. Recalling that, for matrices ε and μ,

$$\mathbf{D} = \varepsilon\mathbf{E} \qquad \text{and} \qquad \mathbf{B} = \mu\mathbf{H},$$

In *differential form:*

1. $\nabla \cdot \mathbf{D} = \rho$ — Gauss's law (Def$^{\underline{n}}$ 2.5)
2. $\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$ — Maxwell–Faraday equation (Def$^{\underline{n}}$ 3.13)
3. $\nabla \cdot \mathbf{B} = 0$ — Gauss's law for magnetism (Def$^{\underline{n}}$ 3.7)
4. $\nabla \times \mathbf{H} = \dfrac{\partial \mathbf{D}}{\partial t} + \mathbf{J}$ — Maxwell–Ampère equation (Def$^{\underline{n}}$ 3.12)

In *integral form*:

1. $\displaystyle\int_S \mathbf{D} \cdot \mathbf{n}\, dS = \int_V \rho\, dV$
2. $\displaystyle\int_C \mathbf{E} \cdot d\mathbf{r} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}\, dS$
3. $\displaystyle\int_S \mathbf{B} \cdot \mathbf{n}\, dS = 0$
4. $\displaystyle\int_C \mathbf{H} \cdot d\mathbf{r} = \int_S \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \right) \cdot \mathbf{n}\, dS$

3.4 Étude on differential forms

Coordinates do not come from God. What we want to be able to say is that Maxwell's equations (particularly in integral form) are valid for any set of coordinates that we choose to integrate along. So let's have a brief discussion about differential forms.

$$\int_C \mathbf{E} \cdot d\mathbf{r} = \int_C E_1 dx^1 + E_2 dx^2 + E_3 dx^3$$

Let $x^1 = x^1(t)$, $x^2 = x^2(t)$, $x^3 = x^3(t)$, then

$$\begin{aligned}
&= \int_a^b \left[E_1(\mathbf{x}(t)) \frac{\partial x^1}{\partial t} dt + E_2(\mathbf{x}(t)) \frac{\partial x^2}{\partial t} dt + E_3(\mathbf{x}(t)) \frac{\partial x^3}{\partial t} dt \right] \\
&= \int_a^b \left[E_1(\mathbf{x}(t)) \frac{\partial x^1}{\partial t} + E_2(\mathbf{x}(t)) \frac{\partial x^2}{\partial t} + E_3(\mathbf{x}(t)) \frac{\partial x^3}{\partial t} \right] dt
\end{aligned}$$

So what if $x^i = x^i(y^k)$, where $i, k = 1, 2, 3$, i.e. our x-coordinates are in terms of y-coordinates, e.g. polars in Cartesians. Then

$$dx^1 = \frac{\partial x^1}{\partial y^1} dy^1 + \frac{\partial x^1}{\partial y^2} dy^2 + \frac{\partial x^1}{\partial y^3} dy^3$$

and similarly for dx^2 and dx^3 so in this case,

$$\begin{aligned}\int_C \mathbf{E}\cdot d\mathbf{r} &= \int_C E_1 dx^1 + E_2 dx^2 + E_3 dx^3 \\ &= \int \left[E_1\frac{\partial x^1}{\partial y^1} + E_2\frac{\partial x^2}{\partial y^1} + E_3\frac{\partial x^3}{\partial y^1}\right] dy^1 \\ &\quad + \int \left[E_1\frac{\partial x^1}{\partial y^2} + E_2\frac{\partial x^2}{\partial y^2} + E_3\frac{\partial x^3}{\partial y^2}\right] dy^2 \\ &\quad + \int \left[E_1\frac{\partial x^1}{\partial y^3} + E_2\frac{\partial x^2}{\partial y^3} + E_3\frac{\partial x^3}{\partial y^3}\right] dy^3 \\ &= \int \widehat{\mathbf{E}}\cdot d\mathbf{y}\end{aligned}$$

where $\widehat{\mathbf{E}}$ is $\mathbf{E}$ in y-coordinates, i.e.

$$\widehat{E}_1 = \frac{\partial \mathbf{x}}{\partial y^1}\cdot\mathbf{E} = \sum_{i=1}^{3}\frac{\partial x^i}{\partial y^1}E_i$$

or generally,

$$\widehat{E}_j = \sum_{i=1}^{3}\frac{\partial x^i}{\partial y^j}E_i$$

so we can interchange

$$\sum_{i=1}^{3} E_i(\mathbf{x})dx^i \longleftrightarrow \sum_{j=1}^{3}\widehat{E}_j(\mathbf{y})dy^j$$

i.e. Maxwell's equations are valid for any choice of coordinates. This means that we can apply them, with suitable considerations of course, in any manifold.

3.4.1 Vectors and covectors

Velocity is a real vector.

$$\begin{aligned}v^i &= \frac{dx^i}{dt} \\ &= \frac{dx^i(\mathbf{y})}{dt} \\ &= \sum_{j=1}^{3}\frac{\partial x^i}{\partial y^j}\frac{\partial y^j}{\partial t} \\ &= \sum_{j=1}^{3}\frac{\partial x^i}{\partial y^j}\widehat{v}^j.\end{aligned}$$

See how we're summing over j, in contrast to $\widehat{\mathbf{E}}$, where we're summing over i. So

$$\widehat{\mathbf{v}} = \left(\frac{\partial x}{\partial y}\right)^{-1} \mathbf{v}$$
$$\widehat{\mathbf{E}} = \left(\frac{\partial x}{\partial y}\right)^{\mathrm{T}} \mathbf{E}$$

So when you're looking at a point, $\mathbf{v}$ is a *vector*, and $\mathbf{E}$ is a *covector*.

Generally, $\mathbf{E}, \mathbf{H}$ are one-forms, and $\mathbf{B}, \mathbf{D}$ are two-forms.

Chapter 4

Electromagnetic Waves

Solutions to Maxwell's equations in the absence of sources and sinks, i.e. with $\rho = 0$, $\mathbf{J}^{\text{ext}} = \mathbf{0}$, are called *electromagnetic waves*.

Definition 4.1 Ohm's law, as generalised by Gustav Kirchhoff, says that

$$\mathbf{J} = \mathbf{J}^{\text{ext}} + \sigma \mathbf{E}$$

where σ is the conductivity matrix, a function of $\mathbf{x}$ and t. σ has to be positive definite.

In our study of (the very beginnings!) of the theory of electromagnetic waves we will also assume that

$$\sigma = 0,$$

i.e. the conductivity is 0.

We will also assume that $\varepsilon = \varepsilon(\mathbf{x})$, $\mu = \mu(\mathbf{x})$, i.e. the electric and magnetic permitivities are time-independent. These reduce Maxwell's equations to

1. $\nabla \cdot \mathbf{D} = 0$
2. $\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$
3. $\nabla \cdot \mathbf{B} = 0$
4. $\nabla \times \mathbf{H} = \dfrac{\partial \mathbf{D}}{\partial t}$

4.1 Electromagnetic energy

Electromagnetic waves carry with them electromagnetic energy. The energy density $\mathcal{E}(\mathbf{x}, t)$ of this energy is given by

$$\begin{aligned}\mathcal{E}(\mathbf{x}, t) &= \frac{1}{2}\left[\mathbf{D}\cdot\mathbf{E} + \mathbf{B}\cdot\mathbf{H}\right] \\ &= \frac{1}{2}\left[(\varepsilon\mathbf{E})\cdot\mathbf{E} + (\mu\mathbf{H})\cdot\mathbf{H}\right].\end{aligned} \tag{4.1}$$

Now we look at how $\mathcal{E}$ changes with time t in some domain $V \subset \mathbb{R}^3$,

$$\begin{aligned}\frac{d}{dt}\int_V \mathcal{E}(\mathbf{x}, t)\, dV &= \int_V \frac{\partial \mathcal{E}(\mathbf{x}, t)}{\partial t}\, dV \\ &= \int_V \frac{\partial}{\partial t}\left(\frac{1}{2}\left[(\varepsilon\mathbf{E})\cdot\mathbf{E} + (\mu\mathbf{H})\cdot\mathbf{H}\right]\right) dV\end{aligned}$$

and using $\frac{\partial}{\partial t}\left[(\varepsilon\mathbf{E})\cdot\mathbf{E}\right] = 2\left[\frac{\partial}{\partial t}(\varepsilon\mathbf{E})\cdot\mathbf{E}\right]$,

$$\begin{aligned}&= \int_V \left[\left(\varepsilon\frac{\partial\mathbf{E}}{\partial t}\right)\cdot\mathbf{E} + \left(\mu\frac{\partial\mathbf{H}}{\partial t}\right)\cdot\mathbf{H}\right] dV \\ &= \int_V \left[\frac{\partial\mathbf{D}}{\partial t}\cdot\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t}\cdot\mathbf{H}\right] dV.\end{aligned} \tag{4.2}$$

Using what's left of Maxwell's equations, rewrite (4.2) as

$$\begin{aligned}&= \int_V \left[(\nabla\times\mathbf{H})\cdot\mathbf{E} - \mathbf{H}\cdot(\nabla\times\mathbf{E})\right] dV \\ &= -\int_V \nabla\cdot(\mathbf{E}\times\mathbf{H})\, dV,\end{aligned} \tag{4.3}$$

where we use that $-\nabla\cdot(\mathbf{a}\times\mathbf{b}) = (\nabla\times\mathbf{b})\cdot\mathbf{a} - \mathbf{b}\cdot(\nabla\times\mathbf{a})$ for any vector fields $\mathbf{a}, \mathbf{b}$.

Problem 4.2 Show that

$$-\nabla\cdot(\mathbf{a}\times\mathbf{b}) = (\nabla\times\mathbf{b})\cdot\mathbf{a} - \mathbf{b}\cdot(\nabla\times\mathbf{a})$$

Solution Let $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, and just do the computation. Boring but trivial. ✓

Let's now introduce the *Poynting* vector field.

Definition 4.3 The *Poynting* vector field $\mathbf{S}$, named after its inventor John Henry Poynting, who was a professor at the University of Birmingham from 1880, is defined as

$$\mathbf{S} = \mathbf{E}\times\mathbf{H}.$$

With this definition, (4.3) gives rise to

$$\int_V \frac{\partial \mathcal{E}(\mathbf{x},t)}{\partial t}\,dV = -\int_V \nabla\cdot\mathbf{S}\,dV$$

and then using the divergence theorem

$$= -\int_{\partial V} \mathbf{S}\cdot\mathbf{n}\,dA, \tag{4.4}$$

where dA is an area element and $\mathbf{n}$ is the exterior unit normal to ∂V.

Equation (4.4) implies that the Poynting vector $\mathbf{S}$ describes the flow of electromagnetic energy. Rewritten in the differential form it gives rise to the following continuity equation

$$\frac{\partial \mathcal{E}(\mathbf{x},t)}{\partial t} + \nabla\cdot\mathbf{S}(\mathbf{x},t) = 0,$$

cf. the charge continuity equation $\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{J} = 0$.

Problem 4.4 Show that

$$\frac{\partial \mathcal{E}(\mathbf{x},t)}{\partial t} + \nabla\cdot\mathbf{S}(\mathbf{x},t) = 0.$$

Solution

$$\begin{aligned}
\nabla\cdot\mathbf{S} &= \nabla\cdot(\mathbf{E}\times\mathbf{H})\\
&= (\nabla\times\mathbf{E})\cdot\mathbf{H} - \mathbf{E}\cdot(\nabla\times\mathbf{H})\\
&= \left(-\frac{\partial\mathbf{B}}{\partial t}\right)\cdot\mathbf{H} - \mathbf{E}\cdot\left(\frac{\partial\mathbf{D}}{\partial t}\right)\\
&= -\left[\left(\frac{\partial\mathbf{B}}{\partial t}\right)\cdot\mathbf{H} + \mathbf{E}\cdot\left(\frac{\partial\mathbf{D}}{\partial t}\right)\right]
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial\mathcal{E}}{\partial t} &= \frac{\partial}{\partial t}\frac{1}{2}\left[\mathbf{D}\cdot\mathbf{E} + \mathbf{B}\cdot\mathbf{H}\right]\\
&= \frac{\partial}{\partial t}\frac{1}{2}\left[(\varepsilon\mathbf{E})\cdot\mathbf{E} + (\mu\mathbf{H})\cdot\mathbf{H}\right]\\
&= \frac{\partial}{\partial t}(\varepsilon\mathbf{E})\cdot\mathbf{E} + \frac{\partial}{\partial t}(\mu\mathbf{H})\cdot\mathbf{H}\\
&= \frac{\partial\mathbf{D}}{\partial t}\cdot\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t}\cdot\mathbf{H}
\end{aligned}$$

$\Longrightarrow \frac{\partial\mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S} = 0$, as desired. ✓

4.2 Electromagnetic waves in an homogeneous isotropic medium

In the following we assume that the electric and magnetic permittivities ε and μ are isotropic and homogeneous (i.e. independent of $\mathbf{x}$),

$$\varepsilon_{ij}(\mathbf{x}) = \varepsilon\delta_{ij} \quad \mu_{ij}(\mathbf{x}) = \mu\delta_{ij},$$

An example of such a medium is the vacuum, where $\varepsilon = \varepsilon_0, \mu = \mu_0$. An example of one that is not is the brick that we introduced in section 3.2.1.

Rewriting Maxwell's equations, taking into account the above,

1. $\nabla \cdot \mathbf{D} = 0$
2. $\nabla\times\mathbf{E} = -\mu\dfrac{\partial \mathbf{H}}{\partial t}$
3. $\nabla \cdot \mathbf{B} = 0$
4. $\nabla\times\mathbf{H} = \varepsilon\dfrac{\partial \mathbf{E}}{\partial t}$

Differentiating equations 2 and 4 with respect to t, we get

$$\nabla\times\left(\frac{\partial \mathbf{E}}{\partial t}\right) = -\mu\frac{\partial^2 \mathbf{H}}{\partial t^2}$$
$$\nabla\times\left(\frac{\partial \mathbf{H}}{\partial t}\right) = \varepsilon\frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

Using Maxwell's equations as rewritten above, for $\frac{\partial \mathbf{E}}{\partial t}, \frac{\partial \mathbf{H}}{\partial t}$, we get

$$\nabla\times\left(\nabla\times\mathbf{H}\right) = -\varepsilon\mu\frac{\partial^2 \mathbf{H}}{\partial t^2} \tag{4.5}$$
$$\nabla\times\left(\nabla\times\mathbf{E}\right) = -\varepsilon\mu\frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{4.6}$$

Recall that, by the definition of the vector Laplacian $\nabla^2\mathbf{H}$,

$$\nabla\times\left(\nabla\times\mathbf{H}\right) = \nabla\left(\nabla\cdot\mathbf{H}\right) - \nabla^2\mathbf{H},$$

where

$$\nabla^2\mathbf{H} = (\nabla^2 H_1, \nabla^2 H_2, \nabla^2 H_3).$$

Then using equation 3 in Maxwell's equations, $\nabla \cdot \mathbf{H} = 0$, so we get

$$\nabla\times\left(\nabla\times\mathbf{H}\right) = -\nabla^2\mathbf{H}$$

and (4.5) becomes

$$\frac{\partial^2 \mathbf{H}}{\partial t^2} = \frac{1}{\varepsilon\mu}\nabla^2\mathbf{H}.$$

Similarly we can also derive

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\varepsilon\mu}\nabla^2\mathbf{E}.$$

Problem 4.5 Derive
$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\varepsilon\mu}\nabla^2\mathbf{E}.$$

Component wise, this is equivalent to six scalar wave equations for the components $E_i(\mathbf{x}, t)$, $H_i(\mathbf{x}, t)$, where $i = 1, 2, 3$,

$$\frac{\partial^2 E_i(\mathbf{x}, t)}{\partial t^2} - c^2\nabla^2 E_i(\mathbf{x}, t) = 0$$
$$\frac{\partial^2 H_i(\mathbf{x}, t)}{\partial t^2} - c^2\nabla^2 H_i(\mathbf{x}, t) = 0, \tag{4.7}$$

where
$$c = \frac{1}{\sqrt{\varepsilon\mu}} \tag{4.8}$$

is the speed of electromagnetic waves (the speed of light) in the medium. (Note that in vacuum, $c_0 = \frac{1}{\sqrt{\varepsilon_0\mu_0}}$).

4.3 Plane waves

An important special case of electromagnetic waves are the *plane* waves. In this case, $\mathbf{E}$ and $\mathbf{H}$ depend, in addition to t, only on one space variable $x_e = \mathbf{x} \cdot \widehat{\mathbf{e}}$, where $\widehat{\mathbf{e}} = (e_1, e_2, e_3)$ is a unit vector. Since we're in isotropic material, we can take $\widehat{\mathbf{e}} = \widehat{\mathbf{i}}$, and $x_e = x_1$. Explicitly,

$$\mathbf{E} = \mathbf{E}(x_1, t) \qquad \mathbf{H} = \mathbf{H}(x_1, t) \tag{4.9}$$

Clearly, in this case, $\mathbf{E}$ and $\mathbf{H}$ remain the same, at any fixed time t, in any plane orthogonal to $\widehat{\mathbf{i}}$, which explains the name "plane waves". Then, the 3D wave equations (4.7) become 1D wave equations (which we studied in MATH1302*),

$$\frac{\partial^2 E_i(x_1, t)}{\partial t^2} - c^2\frac{\partial^2 E_i(x_1, t)}{\partial x_1^2} = 0$$
$$\frac{\partial^2 H_i(x_1, t)}{\partial t^2} - c^2\frac{\partial^2 H_i(x_1, t)}{\partial x_1^2} = 0 \tag{4.10}$$

We also do know the general solution in this case,

$$E_i(\mathbf{x}, t) = E_i^+(x_1 + ct) + E_i^-(x_1 - ct)$$
$$H_i(\mathbf{x}, t) = H_i^+(x_1 + ct) + H_i^-(x_1 - ct). \tag{4.11}$$

i.e. in each case we have the superposition of two waves, one going to the right and one going to the left.

*LOL

Problem 4.6 Show that if $\mathbf{E} = \mathbf{E}(x_1, t)$, $\mathbf{H} = \mathbf{H}(x_1, t)$, then

$$\begin{aligned} E_i(\mathbf{x}, t) &= E_i^+(x_1 + ct) + E_i^-(x_1 - ct) \\ H_i(\mathbf{x}, t) &= H_i^+(x_1 + ct) + H_i^-(x_1 - ct) \end{aligned}$$

for $i = 1, 2, 3$ and some functions $E_i^\pm(x_1)$, $H_i^\pm(x_1)$.

Solution Look at the wave equation

$$\frac{\partial^2 f}{\partial t^2} = c^2 \nabla^2 f,$$

assuming that $f = f(x_1, t)$.

Use the substitution

$$\zeta = x_1 - ct \qquad \eta = x_1 + ct$$

Then

$$\begin{aligned} \frac{\partial f}{\partial \eta} &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial \eta} + \frac{\partial f}{\partial t}\frac{\partial t}{\partial \eta} \\ &= \frac{\partial f}{\partial x_1} + \frac{1}{c}\frac{\partial f}{\partial t} \\ &:= F(x_1, t) \\ \frac{\partial^2 f}{\partial \zeta \partial \eta} &= \frac{\partial}{\partial \zeta}\left(\frac{\partial f}{\partial x_1} + \frac{1}{c}\frac{\partial f}{\partial t}\right) \\ &= \frac{\partial F}{\partial x_1}\frac{\partial x_1}{\partial \zeta} + \frac{\partial F}{\partial t}\frac{\partial t}{\partial \zeta} \\ &= \left(\frac{\partial^2 f}{\partial x_1^2} + \frac{1}{c}\frac{\partial^2 f}{\partial x_1 \partial t}\right) + \left(\frac{\partial^2 f}{\partial x_1 \partial t} + \frac{1}{c}\frac{\partial^2 f}{\partial t^2}\right) \cdot \left(-\frac{1}{c}\right) \\ &= \frac{\partial^2 f}{\partial x_1^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} \\ &= \frac{\partial^2 f}{\partial x_1^2} - \frac{1}{c^2}(c^2 \nabla^2 f) \\ &= \frac{\partial^2 f}{\partial x_1^2} - \frac{\partial^2 f}{\partial x_1^2} \\ &= 0. \end{aligned}$$

So

$$\frac{\partial F}{\partial \zeta} = 0$$

since $F = \frac{\partial f}{\partial \eta}$. Therefore F is an arbitrary function of η: $F = F(\eta)$. That is to say

$$\frac{\partial f}{\partial \eta} = F(\eta),$$

an arbitrary function of η. So

$$f = \alpha(\eta) + \beta(\zeta)$$

where

$$\alpha(\eta) = \int_0^\eta F(\widetilde{\eta})\, d\widetilde{\eta}$$

hence

$$f = \alpha(x_1 - ct) + \beta(x_1 + ct)$$

✓

This formula shows that, indeed, the waves propagate with speed c in the $\widehat{\mathbf{i}}$-direction, or, more precisely, $\mathbf{E}(x_1 - ct), \mathbf{H}(x_1 - ct)$ is the wave propagating in the $\widehat{\mathbf{i}}$-direction while $\mathbf{E}(x_1 + ct), \mathbf{H}(x_1 + ct)$ is the wave propagating in the $-\widehat{\mathbf{i}}$-direction. Note that to satisfy Maxwell's equations, each of these two waves should satisfy these equations inpedendently.

Using the original Maxwell equations, rather than the 1D wave equation (4.7), we can further analyse $\mathbf{E}, \mathbf{H}$ in this case. However, we will concentrate on one special type of plane waves.

4.4 Harmonic plane waves

An electromagnetic wave is *harmonic* if its dependence on t is periodic and of the form $e^{\pm i\omega t}$, where ω is the frequency of oscillation. Thus, a harmonic electromagnetic wave has the form

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= \mathbf{E}^+(\mathbf{x})e^{i\omega t} + \mathbf{E}^-(\mathbf{x})e^{-i\omega t}, \\ \mathbf{H}(\mathbf{x}, t) &= \mathbf{H}^+(\mathbf{x})e^{i\omega t} + \mathbf{H}^-(\mathbf{x})e^{-i\omega t}. \end{aligned}$$

If the electromagnetic wave is harmonic and plane, then the above equation, together with (4.11), implies that

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= \mathbf{E}^+e^{i(kx_1+\omega t)} + \mathbf{E}^-e^{i(kx_1-\omega t)} + \text{complex conjugate} \\ \mathbf{H}(\mathbf{x}, t) &= \mathbf{H}^+e^{i(kx_1+\omega t)} + \mathbf{H}^-e^{i(kx_1-\omega t)} + \text{complex conjugate}, \end{aligned} \tag{4.12}$$

where $\mathbf{E}^\pm, \mathbf{H}^\pm$ are (complex) constant vectors and $\frac{\omega}{k} = c$.

Again, to satisfy Maxwell's equations, each of the terms in these two sums of four terms should satisfy them. Thus, it is sufficient to concentrate on, say, $\mathbf{E}^-e^{i(kx_1-\omega t)}$ and $\mathbf{H}^-e^{i(kx_1-\omega t)}$. We'll drop the minus sign to signal our generalisation.

Recall the divergence pair of Maxwell's equations: $\nabla \cdot \mathbf{B} = \nabla \cdot \mathbf{D} = 0$. Obviously then $\nabla \cdot \mathbf{H} = \nabla \cdot \mathbf{E} = 0$ so

$$\begin{aligned} 0 &= \nabla \cdot \left(\mathbf{E}e^{i(kx_1-\omega t)}\right) \\ &= (\mathbf{E} \cdot \widehat{\mathbf{i}})ike^{i(kx_1-\omega t)} \\ \implies 0 &= (\mathbf{E} \cdot \widehat{\mathbf{i}}) \end{aligned}$$

and similarly

$$\implies 0 = (\mathbf{H} \cdot \widehat{\mathbf{i}}) \tag{4.13}$$

i.e. $\mathbf{E}, \mathbf{H}$ lie in the plane orthogonal to the direction of the wave propagation $\widehat{\mathbf{i}}$.

Next, the Maxwell-Faraday equation tells that

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = -\mu \frac{\partial \mathbf{H}}{\partial t}$$

and (4.12) tells us that

$$-\mu \frac{\partial \mathbf{H}}{\partial t} = ick\mu \mathbf{H}$$

And we can work out that

$$\nabla \times \mathbf{E} = ik(\widehat{\mathbf{i}} \times \mathbf{E})$$

Problem 4.7 Show that

$$\nabla \times \mathbf{E} = ik(\widehat{\mathbf{i}} \times \mathbf{E})$$

Solution Recall that from equation (4.9), $\mathbf{E} = \mathbf{E}(x_1, t)$. So

$$\begin{aligned}
\nabla \times \mathbf{E} &= \begin{vmatrix} \widehat{\mathbf{i}} & \widehat{\mathbf{j}} & \widehat{\mathbf{k}} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ E_1(x_1, t) & E_2(x_1, t) & E_3(x_1, t) \end{vmatrix} \\
&= \frac{\partial E_2}{\partial x_1} \widehat{\mathbf{k}} - \frac{\partial E_3}{\partial x_1} \widehat{\mathbf{j}}
\end{aligned}$$

which by equation (4.12),

$$= ik(E_2 \widehat{\mathbf{k}} - E_3 \widehat{\mathbf{j}})$$

and on the other side,

$$\begin{aligned}
ik(\widehat{\mathbf{i}} \times \mathbf{E}) &= ik \begin{vmatrix} \widehat{\mathbf{i}} & \widehat{\mathbf{j}} & \widehat{\mathbf{k}} \\ 1 & 0 & 0 \\ E_1(x_1, t) & E_2(x_1, t) & E_3(x_1, t) \end{vmatrix} \\
&= ik(E_2 \widehat{\mathbf{k}} - E_3 \widehat{\mathbf{j}})
\end{aligned}$$

which is clearly equivalent. ✓

so this implies

$$\begin{aligned}
\mathbf{H} &= \frac{1}{c\mu} (\widehat{\mathbf{i}} \times \mathbf{E}) \qquad (4.14) \\
&= \sqrt{\frac{\varepsilon}{\mu}} (\widehat{\mathbf{i}} \times \mathbf{E}) \\
&= \frac{1}{Z} (\widehat{\mathbf{i}} \times \mathbf{E})
\end{aligned}$$

where

$$Z = \sqrt{\frac{\mu}{\varepsilon}}$$

is called the *electromagnetic impedance*.

Similarly, the Maxwell-Ampère equation implies that

$$\mathbf{E} = -Z\widehat{\mathbf{i}} \times \mathbf{H}. \tag{4.15}$$

Note that (4.15) follows from (4.14) if we apply $\widehat{\mathbf{i}}\times$ to both parts.

Let us analyse these equations. In principle, $\mathbf{E}$ is a complex vector. By shifting t, we can assume

$$\mathbf{E} = E_1\mathbf{u}_1 + E_2 e^{i\alpha}\mathbf{u}_2,$$

where E_1, E_2 are positive and $\mathbf{u}_1, \mathbf{u}_2, \widehat{\mathbf{i}}$ form a right-hand triple of orthonormal vectors. Then,

$$Z\mathbf{H} = E_1\mathbf{u}_2 - E_2 e^{i\alpha}\mathbf{u}_1.$$

This means that the real parts of $\mathbf{E}$ and $\mathbf{H}$ are orthogonal, and the same is true for their imaginary parts.

The last remark is about *polarisation*. When $\alpha \in \{0, \pm\pi\}$, then the direction of $\mathbf{E}\exp i(kx_1 - \omega t)$ does not depend on $(\mathbf{x}, t)$: we have linearly polarised waves since $e^{i\alpha} = \pm 1$. However, for other α, if we fix any $\mathbf{x}$, the direction of *physical* electric field (which is the real part of $\mathbf{E}\exp i(kx_1 - \omega t)$) will rotate with period $T = \frac{2\pi}{\omega}$.

For example, if $E_1 = E_2$ and $\alpha = \frac{\pi}{2}$

$$\mathbf{E}^{\text{physical}}(\mathbf{x}, t) = E_1\left(\cos(kx_1 - \omega t)\mathbf{u}_1 - \sin(kx_1 - \omega t)\mathbf{u}_2\right).$$

which gives you an ellipse.

Problem 4.8 Let $\alpha \neq \frac{\pi}{2}, \mathbf{u}_2 \neq \mathbf{0}$. So $e^{i\alpha}$ is not real. Find the formula of $\mathbf{E}^{\text{physical}}$ and show that we get a slanted ellipse.

This problem is left as an exercise for the reader!

And that concludes our study of Maxwell's Theory of Electrodynamics. The exam in 2011, for which these notes were given, featured two questions from chapter 1, two questions from chapters 2 and 3, and one question from chapter 4. Please report any errors you have noticed in the document to the editors, whose details can be found on page 4. Chocolate, beer and any other form of appreciation will be warmly received.

谢谢

Figure 4.1: Class of 2011