Partial Differential Equations

T. W. Körner after Joshi and Wassermann

October 12, 2002

Small print These notes are a digest of much more complete notes by M. S. Joshi and A. J. Wassermann which are also being issued for this course. I should **very much** appreciate being told of any corrections or possible improvements and might even part with a small reward to the first finder of particular errors. This document is written in LaTeX2e and stored in the file labelled `~twk/IIB/PDE.tex` on emu in (I hope) read permitted form. My e-mail address is `twk@dpmms`.

Contents

1 Introduction

1.1 Generalities

When studying ordinary (i.e. not partial) differential equations we start with linear differential equations with constant coefficients. We then study linear differential equations and then plunge into a boundless ocean of non-linear differential equations. The reader will therefore not be surprised if most of a first course on the potentially much more complicated study of partial differential equations should limit itself essentially to linear partial differential equations.

A linear partial differential equation is an equation of the form $Pu = f$ where u and f are suitable functions from $\mathbb{R}^n$ to $\mathbb{R}$ and

$$P = \sum_{|\alpha|\leq k} a_\alpha(x)\partial^\alpha.$$

Our definition introduces some of the condensed notation conventions which make the subject easier for the expert and harder for the beginner. The *multi-index*

$$\alpha = (\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n) \in \mathbb{N}^n.$$

We write

$$|\alpha| = \sum_{j=1}^n \alpha_j$$

and

$$\partial^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}}.$$

Sometimes we write $\partial^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x}$.

Although the first part of the course will deal with 'first order' linear partial differential equations without restrictions on the coefficients (and indeed even with slightly more general partial differential equations) the main part of the course will deal with linear partial differential equations with constant coefficients. The main tools used will be Laurent Schwartz's theory of distributions and the Fourier transform.

The fact that we do not deal with non-linear equations does not mean that they are not important. The equations of general relativity are non-linear and at a more mundane level the Navier-Stokes equation of fluid dynamics

$$\frac{\partial u}{\partial t} - \Delta u + u.\Delta u = f - \Delta p,\ \ \Delta.u = 0$$

is non-linear.

We call $P = \sum_{|\alpha| \leq k} a_\alpha(x) \partial^\alpha$ a differential operator (or just an 'operator'). We look at three such operators in detail. The first which must be most studied non-trivial differential operator in mathematics is the Laplacian Δ known in more old fashioned texts as ∇^2 and defined by

$$\Delta u = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2}$$

where $u : \mathbb{R}^n \to \mathbb{R}$ is a suitable function. The second is the wave operator $\Box$ given by

$$\Box u(t, \mathbf{x}) = \frac{\partial^2 u}{\partial t^2}(t, \mathbf{x}) - \Delta_{\mathbf{x}} u(t, \mathbf{x})$$

where $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ and $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is a suitable function. The third is the heat operator J given by

$$Ju(t, \mathbf{x}) = \frac{\partial u}{\partial t}(t, \mathbf{x}) - \Delta_{\mathbf{x}} u(t, \mathbf{x}).$$

The notations Δ and $\Box$ are standard but the notation J is not.

1.2 The symbol

A key concept in more advanced work is the *total symbol* $\sigma(P)$ of a linear partial differential operator

$$P(x, \partial) = \sum_{|\alpha| \leq k} a_\alpha(x) \partial^\alpha$$

obtained by replacing $\frac{\partial}{\partial x_j}$ by $i\xi_j$ so that

$$\sigma(P) = p(x, \xi) = \sum_{|\alpha| \leq k} a_\alpha(x)(i\xi)^\alpha = e^{-ix.\xi} P(e^{ix.\xi}).$$

Note that x and ξ are n dimensional vectors and that we use the convention

$$y^\alpha = \prod_{j=1}^{n} y_j^{\alpha_j}.$$

To see where the symbol comes from, observe that taking Fourier transforms

$$\widehat{Pu}(\xi) = p(x, \xi)\hat{u}(\xi)$$

and so taking inverse Fourier transforms

$$P(x, \partial)u = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{ix.\xi} p(x, \xi)\hat{u}(\xi)\, d\xi.$$

Applying the Leibnitz rule we get the following result.

Lemma 1.1 (Proposition 1). *If P and Q are linear partial differential operators*

$$\sigma(PQ)(x, \xi) = \sum \frac{(-i)^{|\alpha|}}{\alpha!} \partial_\xi^\alpha \sigma(P)(x, \xi) \partial_x^\alpha \sigma(Q)(x, \xi).$$

Here as one might expect $\alpha! = \prod_{j=1}^{n} \alpha_j!$.

We say that the differential operator

$$P(x, \partial) = \sum_{|\alpha| \leq k} a_\alpha(x) \partial^\alpha$$

has order k and that it has principal symbol

$$\sigma_k(P)(x, \xi) = \sum_{|\alpha| = k} a_\alpha(x)(i\xi)^\alpha.$$

Notice that if P has degree k and Q degree l then the principal symbol of PQ is given by the simple formula

$$\sigma_{l+k}(PQ) = \sigma_k(P)\sigma_l(Q).$$

If the principal symbol is never zero or only vanishes to first order then the operator is said to be of *principal type*. In more advanced work it is shown that when the operator is of principal type the lower order terms have little effect on the qualitative behaviour of the associated partial differential equation.

We define the *characteristic set* to be the subset of $\mathbb{R}^n \times \mathbb{R}^n$ where the principal symbol vanishes.

$$\mathrm{char}(P) = \{(x, \xi) \,:\, \sigma_k(P)(x, \xi) = 0\}.$$

We say that P is *elliptic* at x if the principle symbol $\sigma_k(P)(x, \xi) \neq 0$ for $\xi \neq 0$. If P is *elliptic* at x for all x we say that P is elliptic.

The reason for the use of the word 'elliptic' may be traced to the symbols of the three special operators.

$$\begin{aligned} \sigma(\Delta)(\xi) &= -|\xi|^2 && \text{(Laplacian)} \\ \sigma(\Delta)(\tau, \xi) &= -\tau^2 + |\xi|^2 && \text{(Wave)} \\ \sigma(J) &= i\tau + |\xi|^2 && \text{(Heat)} \end{aligned}$$

Traditionally second order operators which behaved like the Laplacian were called elliptic, those which behaved like the wave operator were called hyperbolic and those that behaved like the heat operator were called parabolic. The distinction is very useful but the reference to conic sections is not.

Example 1.2 (Example 1). *(i) We could have considered complex valued u in place of real valued u. If we do this the operator*

$$P = \frac{\partial}{\partial x} + i\frac{\partial}{\partial y}$$

has principal symbol

$$\sigma_1(P)(x, y : \xi, \eta) = i\xi - \eta$$

and so is elliptic.

(ii) The Laplacian Δ is elliptic but the wave operator $\Box$ and the heat operator J are not.

$$\mathrm{char}(\Box) = \{(x,t,\tau,\xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \,:\, \tau^2 - |\xi|^2 = 0\}$$

$$\mathrm{char}(J) = \{(x,t,\tau,\xi) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \,:\, \tau - |\xi|^2 = 0\}$$

2 Ordinary differential equations

2.1 The contraction mapping theorem

Hadamard introduced the idea of a well posed problem into the study of partial differential equations. According to Hadamard a well posed problem must have a solution which exists, is unique and varies continuously on the given data. Without going too deeply into the matter we may agree that these are reasonable matters to investigate.

Thus given a partial differential equation with boundary conditions we shall study the following problems.

Existence Can we show that there is a solution in a neighbourhood of a given point? Can we show that there exists a solution everywhere?

Uniqueness Is the solution unique?

Continuity Does u depend continuously on the boundary conditions? Does u depend continuously on on other elements of the problem?

Smoothness How many times is u differentiable? Does u have points of singularity? Does the solution blow up in some way after a finite time?

We may illustrate these ideas in the case of ordinary differential equations. Some of the material in this section will be familiar from examination questions on the first analysis course in 1B but you should consider it all with the exception of the contraction mapping theorem (Theorem 2.1) itself to be examinable.

In 1B we proved Banach's contraction mapping theorem.

Theorem 2.1 (Theorem 1). *Let (X,d) be a complete non-empty metric space and $T : X \to X$ a map such that $d(Tx,Ty) \le k(x,y)$ for all $x,\ y \in X$ and some k with $0 \le k < 1$. Then there exists a unique $x_0 \in X$ such that $Tx_0 = x_0$. If $x \in X$ then $T^n x \to x_0$ as $n \to \infty$.*

This can be strengthened as follows.

Theorem 2.2 (Corollary 1). *Let (X,d) be a complete non-empty metric space and $T : X \to X$ a map. Suppose further that there exists an integer*

$N \geq 1$ and a k with $0 \leq k < 1$ such that $d(T^N x, T^N y) \leq k(x, y)$ for all $x,\ y \in X$. Then there exists a unique $x_0 \in X$ such that $Tx_0 = x_0$. If $x \in X$ then $T^n x \to x_0$ as $n \to \infty$.

Note that these theorems not only give a fixed point but also give a method for finding it.

For the rest of section 2.1 f will be a function from $\mathbb{R} \times \mathbb{R}^n$. Let

$$E = \{t \in \mathbb{R} : |t - t_0| \leq a\} \times \{x \in \mathbb{R}^n : \|x - x_0\| \leq b\}.$$

We assume that f satisfies the Lipschitz condition

$$\|f(t, x_1) - f(t, x_2)\| \leq c\|x_1 - x_2\|$$

on E.

Exercise 2.3. *(i) Show that a function that satisfies a Lipschitz condition is continuous.*

(ii) Show that any continuously differentiable function satisfies a Lipschitz condition.

(iii) Show that a function that satisfies a Lipschitz condition need not be differentiable everywhere.

(iv) Show that the function f considered above is bounded on E.

We set $M = \sup_{(t,x) \in E} \|f(t, x)\|$ and $h = \min(a, bM^{-1})$.

Theorem 2.4. *If f is as above the differential equation*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0, 0) = x_0 \qquad \bigstar$$

has a unique solution for $|t - t_0| \leq h$.

Example 2.5. *(Here $x : \mathbb{R} \to \mathbb{R}$.)*

(i) The differential equation

$$\frac{dx}{dt} = 0,\ x(0) = 0,\ x(1) = 1$$

has no solution.

(ii) Show that the differential equation

$$\frac{dx}{dt} = x^{\frac{2}{3}},\ x(0) = 0$$

has at least two solutions.

From now on we move away from 1B.

Theorem 2.6 (Theorem 3). *The solution of ★ in Theorem 2.4 depends continuously on x_0. More formally, if we define $T : \mathbb{R} \to C([t-h, t+h])$ by taking Ty to be the solution of*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = y$$

and give $C([t-h, t+h])$ the uniform norm then T is continuous.

In the special case of *linear* ordinary differential equations we can give a rather strong perturbation result.

Theorem 2.7 (Theorem 4). *We use the standard operator norm on the space $\mathcal{L} = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ of linear maps. Suppose that $A, B : \mathbb{R} \times \mathbb{R}^n \to \mathcal{L}$ are continuous and that $M \geq \|A(t,x)\|, \ \|B(t,x)\|$ for all t and x.*

$$\frac{d\xi}{dt}(t,x) = A(t,x)\xi(t,x), \ \xi(t_0, x) = a(x)$$
$$\frac{d\eta}{dt}(t,x) = B(t,x)\eta(t,x), \ \eta(t_0, x) = b(x)$$

then if $|t - t_0| \leq K$

$$\|\xi(t,x) - \eta(t,x)\| \leq C(a, K, M)\|A - B\|(e^{M|t-t_0|} - 1) + \|a - b\|e^{M|t-t_0|}$$

where $C(a, K, M)$ depends only on a, M and K.

2.2 Vector fields, integral curves and flows

Let U be an open subset of $\mathbb{R}^n$. A *time dependent vector field* on U is a map

$$f : (-\epsilon, \epsilon) \times U \to \mathbb{R}^n$$

which associates a vector $f(t, x)$ to each time t and point x.

Let $x_0 \in U$. An *integral curve* for the vector field f with starting point x_0 is a map

$$\phi : (-\delta, \delta) \to U$$

such that

$$\frac{d\phi}{dt} = f(t, \phi(t)) \text{ and } \phi(0) = x_0$$

so that the tangent vectors to ϕ are just the values of the vector field at that point at time.

Lemma 2.8. *With the notation above, if f has a continuous derivative then an integral curve always exists.*

A *local flow* for f at x_0 is a map

$$\alpha : (-\delta, \delta) \times U_0 \to U$$

where $\epsilon > \delta > 0$, U_0 is open and $x_0 \in U_0 \subseteq U$, such that

$$\frac{d}{dt}\alpha(t, x) = f(t, \alpha(t, x)), \alpha(0, x) = x.$$

Thus if x is fixed $\alpha_x(t) = \alpha(t.x)$ is an integral curve with starting point x. In some sense a flow is a collection of integral curves.

We note but shall not prove that the smoother the vector field the smoother the associated flow.

Theorem 2.9 (Theorem 5). *If f is C^k and*

$$\frac{d}{dt}\alpha(t, x) = f(t, \alpha(t, x)), \alpha(0, x) = x.$$

then $\alpha \in C^k$ $[1 \leq k]$.

Now suppose f does not depend on t. Let $\alpha_x(t) = \alpha(t, x)$ as before.

Lemma 2.10. *With the notation just established there exists an $\epsilon > 0$ such that*

$$\alpha_{t+s}(u) = \alpha_t(\alpha_s(u))$$

for all $|t|, |s|, |u| < \epsilon$.

It is clear from the proof of Lemma 2.10 that the relation

$$\alpha_{t+s} = \alpha_t \circ \alpha_s$$

will hold whenever it is reasonable for it to hold and in particular if the flow is determined for all time we have a group action. This kind of group action is called a dynamical system (see Course O6).

Definition 2.11. *Let $f : U \to \mathbb{R}^n$ be a vector field. We say that $x_0 \in U$ is a critical point of f if $f(x_0) = 0$.*

Lemma 2.12. *Let $f : U \to \mathbb{R}^n$ be a (time independent) vector field.*

(i) If ϕ is an integral curve passing through a critical point x_0 then ϕ is constant.

(ii) Suppose ϕ is an integral curve defined for all times t and $\phi(t) \to x_0 \in U$ as $t \to \infty$. Then x_0 is a critical point.

2.3 First order linear and semi-linear PDEs

Integral curves give us a method for solving first order partial differential equations

$$\sum_{j=1}^{n} a_j(x)\frac{\partial u}{\partial x_j} = f(x).$$

In fact the method works for the slightly more general *semi-linear* first order partial differential equation

$$\sum_{j=1}^{n} a_j(x)\frac{\partial u}{\partial x_j} = f(x,u). \tag{*}$$

Here we assume that a_j is a real once continuously differentiable function and that f is a real or complex continuously differentiable function.

The idea is to consider the vector field

$$A(x) = (a_1(x),\ a_2(x),\ \ldots,\ a_n(x)).$$

If γ is an integral curve of A then

$$\frac{d}{dt}u(\gamma(t)) = \sum_{j=1}^{n} a_j(\gamma(t))\frac{\partial u}{\partial x_j}(\gamma(t)).$$

Thus solving (*) is equivalent to solving the ordinary differential equation

$$\frac{d}{dt}u(\gamma(t)) = f(\gamma(t), u(\gamma(t))) \tag{**}$$

This not only gives us a method for solving (*) but tells us the appropriate boundary conditions to use. For example, specifying $u(0,y)$ along an integral curve will, in general, give a problem with no solution. What we need to do is to specify $u(0,y)$ on a hypersurface S and solve along each integral curve.

Recall that that we defined the *characteristic set* of a linear differential operator $P(x,\partial)$ to be the subset of $\mathbb{R}^n \times \mathbb{R}^n$ where the principal symbol vanishes.

$$\mathrm{char}(P) = \{(x,\xi) \,:\, \sigma_k(P)(x,\xi) = 0\}.$$

In the case of the linear differential operator

$$L(x,\partial) = \sum_{j=1}^{n} a_j(x)\frac{\partial}{\partial x_j}$$

we have

$$\operatorname{char}(L) = \left\{ (x, \xi) \,:\, \sum_{j=1}^{n} a_j(x)\xi_j = 0 \right\}.$$

The hypersurface S is said to be *characteristic* for P at x if the normal vector is a characteristic vector for P. The hypersurface S is called *non-characteristic* if it is not characteristic at any point. In the case of the linear differential operator L which we are considering, S is non-characteristic if at each $x \in S$ the normal ζ to S satisfies

$$\sum_{j=1}^{n} a_j \zeta_j \neq 0.$$

Retracing our steps it is clear that we have essentially (there is a slight subtlety which we shall deal with when we prove Theorem 2.15) proved the following theorem. (Note the use of the word 'locally'.)

Theorem 2.13 (Theorem 6). *Locally, there is a unique solution of* $(*)$ *with* $u(x, 0)$ *given on a non-characteristic hypersurface* S.

Working through a few examples will make the theory much clearer.

Example 2.14 (Example 2). *Solve the partial differential equation*

$$\frac{\partial u}{\partial x} + 2x \frac{\partial u}{\partial y} = u^2$$

subject to $u(0, y) = f(y)$.

2.4 First order quasi-linear PDEs

An extension of the previous technique enables us to solve the slightly more general first order *semi-linear* partial differential equation

$$\sum_{j=1}^{n} a_j(x, u) \frac{\partial u}{\partial x_j} = b(x, u). \qquad (***)$$

subject to $u|_S = \phi$ where S is a non-characteristic hypersurface. Here we assume that the a_j is a real once continuously differentiable function and that b is a real continuously differentiable function. (Note that the technique does not apply if B is complex valued and that a counter-example of Hans Lewy shows that a solution may not exist.)

The solution technique relies on regarding u as a variable just like x. We give the technique as a recipe. Suppose that S is parameterised by a function g so that

$$S = \{x = g(s) \,:\, s \in \mathbb{R}^{n-1}\}.$$

We work with the vector field

$$(a_1, a_2, \ldots, a_n, b)$$

on $\mathbb{R}^{n+1}$ and solve for the integral curves

$$\begin{aligned}
\frac{dx_s}{dt} &= a(x_s(t), y_s(t)) \qquad x_s(0) = g(s)\\
\frac{dy_s}{dt} &= b(x_s(t), y_s(t)) \qquad y_s(0) = \phi(s).
\end{aligned}$$

Our solution is basically $y(s,t)$ but we want it as a function of x not (s,t). The map

$$(s,t) \mapsto x(s,t)$$

will have an invertible derivative at $t = 0$ provided that the vector

$$(a_1(g(s), \phi(s)), a_2(g(s), \phi(s)), \ldots, a_n(g(s), \phi(s)))$$

is not tangent to S and since we have specified that S is a non-characteristic this will always be the case. A theorem of differential geometry (proved in Course B4 Differential Manifolds, but in any case very plausible) called the ‘inverse function theorem’ asserts that under these circumstances (continuous derivative invertible at a point) the map

$$(s,t) \mapsto x(s,t)$$

is locally invertible. We may thus define

$$u(x) = y(s(x), t(x))$$

and direct calculation shows that it is a solution. Checking reversibility we see that we have proved the following theorem.

Theorem 2.15 (Theorem 7). *Equation (***) has a unique solution locally.*

Once again, working through a few examples will make the theory much clearer.

Example 2.16 (Example 3). *Solve the partial differential equation*

$$u\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 1$$

subject to $u = s/2$ on $x = y = s$.

If a semi-linear equation (with real coefficients) has real right hand side it can be solved by the technique of this subsection but if the right hand side is complex we must use the previous method.

3 Distributions

3.1 Introduction

Distribution theory is a synthesis of ideas coming from partial differential equations (e.g. weak solutions), physics (the Dirac delta function) and functional analysis but the whole is much more than the sum of its parts.

We shall need the notion of the support of a continuous function.

Definition 3.1. *If $f : \mathbb{R}^n \to \mathbb{C}$ is continuous we define the support* $\operatorname{supp} f$ *of f by*

$$\operatorname{supp} f = \operatorname{closure}\{x \,:\, f(x) \neq 0\}.$$

Lemma 3.2. *If $f : \mathbb{R}^n \to \mathbb{C}$ is continuous the support of f is compact if and only if there exists an $R \geq 0$ such that $f(x) = 0$ if $\|x\| > R$.*

Remember that, for $\mathbb{R}^n$, a set is compact if and only if it is closed and bounded.

We use the space of *test functions* $\mathcal{D} = C_0^\infty(\mathbb{R}^n)$ consisting of smooth (that is infinitely differentiable functions) functions of compact support. This is a very restricted space of functions but fortunately non-trivial.

Lemma 3.3 (Lemma 1). *(i) If we define $E : \mathbb{R} \to \mathbb{R}$ by*

$$\begin{aligned} E(t) &= \exp(-1/t^2) && \textit{for } t > 0 \\ E(t) &= 0 && \textit{otherwise} \end{aligned}$$

then E is infinitely differentiable.

(ii) Given $\delta > \eta > 0$ we can find $F : \mathbb{R} \to \mathbb{R}$ such that

$$\begin{aligned} 1 \geq F(t) &\geq 0 && \textit{for all } t \\ F(t) &= 1 && \textit{if } |t| \leq \eta \\ F(t) &= 0 && \textit{if } |t| \geq \delta. \end{aligned}$$

(iii) Given $\delta > \eta > 0$ we can find $G : \mathbb{R}^n \to \mathbb{R}$ such that

$$\begin{aligned} 1 \geq G(x) &\geq 1 && \textit{for all } x \\ G(x) &= 1 && \textit{if } \|x\| \leq \eta \\ G(x) &= 0 && \textit{if } \|x\| \geq \delta. \end{aligned}$$

Functions like G are called bump functions.

We equip $\mathcal{D}$ with a notion of convergence.

Definition 3.4. *Suppose that $f_j \in \mathcal{D}$ for each j and $f \in \mathcal{D}$. We say that*

$$f_j \underset{\mathcal{D}}{\to} f$$

if the following two conditions hold.

(i) There exists an R such that $\operatorname{supp} f_j \subseteq B(0, R)$ *for all j.*

(ii) For each α we have $\partial^\alpha f_n \to \partial^\alpha f$ uniformly as $n \to \infty$.

This is a definition that needs to be thought about carefully.

We can now define a distribution to be a linear map $T : \mathcal{D} \to \mathbb{C}$ which is continuous in the sense that

$$f_n \underset{\mathcal{D}}{\to} f \text{ implies } Tf_n \to Tf.$$

We write $\mathcal{D}'$ for the set of distributions. we shall often write

$$Tf = \langle t, f \rangle.$$

Lemma 3.5. *The set $\mathcal{D}'$ is a vector space if we use the natural definitions*

$$\langle T + S, f \rangle = \langle T, f \rangle + \langle S, f \rangle, \ \langle \lambda T, f \rangle = \lambda \langle T, f \rangle.$$

Our first key insight is the following.

Lemma 3.6. *If $g \in C(\mathbb{R}^n)$ then if we define T_g by*

$$\langle T_g, f \rangle = \int_{\mathbb{R}^n} g(t) f(t) \, dt$$

we have $T_g \in \mathcal{D}'$.

We usually write $T_g = g$, i.e.

$$\langle T_g, f \rangle = \int_{\mathbb{R}^n} g(t) f(t) \, dt$$

and say that every continuous function is a distribution. (Hence the name 'generalised function' sometimes given to distributions.)

The second insight is more in the nature of a recipe than a theorem.

Lemma 3.7 (The standard recipe). *Suppose that $A : \mathcal{D} \to \mathcal{D}$ is a linear map which is continuous in the sense that*

$$f_n \underset{\mathcal{D}}{\to} f \text{ implies } Af_n \underset{\mathcal{D}}{\to} Af \tag{1}$$

and that $A^t : \mathcal{D} \to \mathcal{D}$ is a linear map which is continuous in the sense that

$$f_n \underset{\mathcal{D}}{\to} f \text{ implies } A^t f_n \underset{\mathcal{D}}{\to} A^t f. \tag{2}$$

Suppose in addition

$$\langle Ag, f\rangle = \langle g, A^t f\rangle \tag{3}$$

Then, if S is any distribution we can define a distribution $T_A S$ by the equation

$$\langle T_A S, f\rangle = \langle S, A^t f\rangle.$$

The mapping $T_A : \mathcal{D}' \to \mathcal{D}'$ is linear and whenever $g \in \mathcal{D}$ we will have $Ag = T_A g$.

We usually write $T_A = A$, so that

$$\langle AT, f\rangle = \langle T, A^t f\rangle.$$

Exercise 3.8. *Although we shall not use this in the course, convergence in $\mathcal{D}'$ is defined as follows. If S_j $[j \geq 1]$ are distributions we say that $T_j \underset{\mathcal{D}'}{\to} S$ if*

$$\langle S_n, f\rangle \to \langle S, f\rangle$$

for all $f \in \mathcal{D}$. Show that if A satisfies the conditions of Lemma 3.7, the map $T_A : \mathcal{D}' \to \mathcal{D}'$ is continuous in the sense that

$$S_j \underset{\mathcal{D}'}{\to} S \text{ implies } T_A S \underset{\mathcal{D}'}{\to} T_A S.$$

Lemma 3.9. *If T is a distribution we can define associated distributions by following Lemma 3.7 as follows.*

(i) (Multiplication by a smooth function.) If $\psi \in C^\infty(\mathbb{R}^n)$

$$\langle \psi T, f\rangle = \langle T, \psi f\rangle$$

for all $f \in \mathcal{D}$.

(ii) (Differentiation)

$$\langle \partial_{x_j} T, f\rangle = -\langle T, \partial_{x_j} f\rangle$$

for all $f \in \mathcal{D}$.

(iii) (Reflection) If we write $Rf(x) = f(-x)$ then

$$\langle RT, f\rangle = \langle T, Rf\rangle$$

for all $f \in \mathcal{D}$.

(iv) (Translation) If we write $\tau_a f(x) = f(x-a)$ then

$$\langle \tau_a T, f\rangle = \langle T, \tau_{-a} f\rangle$$

for all $f \in \mathcal{D}$.

Here is another example along the same lines.

Lemma 3.10. *If A is a linear map on $\mathbb{R}^n$ then it induces a linear change of coordinate map on functions A^* by the formula*

$$A^* f(x) = f(Ax).$$

If A is invertible we can define a transpose map $(A^*)^t$ by $(A^*)^t = |\det A|^{-1}(A^{-1})^*$ i.e. by

$$((A^*)^t(f))(x) = |\det A|^{-1} f(T^{-1}x).$$

If T is a distribution we can define an associated distribution (i.e. obtain a change of variables result for distributions) by following Lemma 3.7 to obtain

$$\langle A^* T, f\rangle = \langle T, |\det A|^{-1}(A^{-1})^* f\rangle.$$

[Joshi and Wasserman think in terms of 'signed areas' and have $\det A$ where we have $|\det A|$ in accordance with the conventions of applied mathematics.]

As an indication that we are on the right road, the next example shows consistency with the Dirac deltas of mathematical methods in 1B.

Example 3.11. *(i) If we write*

$$\langle \delta_a, f\rangle = f(a)$$

for all $f \in \mathcal{D}$ then δ_a is a distribution.

(ii) With the notation of Lemma 3.9 (iv),

$$\tau_b \delta_a = \delta_{a+b}.$$

(iii) We have

$$\langle \partial_{x_j} \delta_a, f \rangle = -(\partial_{x_j} f)(a)$$

for all $f \in \mathcal{D}$.

(iv) If we work on $\mathbb{R}$ and define the Heaviside function $H : \mathbb{R} \to \mathbb{R}$ by

$$\begin{aligned} H(t) &= 1 \qquad \textit{for } t \geq 0, \\ H(t) &= 0 \qquad \textit{otherwise} \end{aligned}$$

then, in the distributional sense, $H' = \delta$.

3.2 The support of a distribution

We start with a definition.

Definition 3.12. *If $T \in \mathcal{D}'$ then the support* $\operatorname{supp} T$ *of T is defined as follows. A point $x \notin \operatorname{supp} T$ if we can find an open set U such that $x \in U$ and whenever $f \in \mathcal{D}$ is such that $\operatorname{supp} f \subseteq U$ we have $\langle T, f \rangle = 0$.*

Exercise 3.13. *(i) Check that if f is a continuous function its support as a continuous function is the same as its support when it is considered as a distribution.*

(ii) Show that if T is a distribution $\operatorname{supp} T$ is closed.

The following theorem is extremely important as is the method (partition of unity) by which we obtain it.

Theorem 3.14. *If $f \in \mathcal{D}$ and $T \in \mathcal{D}'$ then*

$$\operatorname{supp} T \cap \operatorname{supp} f = \emptyset \textit{ implies } \langle T, f \rangle = 0.$$

In my opinion the partition of unity is more an idea rather than a theorem but here is one form expressed as a theorem.

Exercise 3.15 (Theorem 12). *Let K be a compact subset of $\mathbb{R}^n$ and U_1, U_2, … U_m open sets in $\mathbb{R}^n$ such that $K \subseteq \bigcup U_j$. Then we can find $f_j \in \mathcal{D}$ with $0 \leq f_i(x) \leq 1$ for all x, $\operatorname{supp} f_j \subseteq U_j$ and $\sum f_j(x) = 1$ for $x \in K$, $\sum f_j(x) \leq 1$ everywhere.*

When looking at Theorem 3.14 you should keep the following important example in mind.

Exercise 3.16. *We work in* $\mathbb{R}$.

(i) Show that there exists an $f \in \mathcal{D}$ *with* $f(0) = 0$, $f'(0) = 1$.

(ii) Observe that $f(x) = 0$ *when* $x \in \operatorname{supp} \delta_0$ *but* $\langle \delta_0', f \rangle = -1$.

(iii) Why does this not contradict Theorem 3.14?

The support tells us where a distribution lives. A related concept is the singular support.

Definition 3.17. *If* $T \in \mathcal{D}'$ *then the singular support* $\operatorname{singsupp} T$ *of* T *is defined as follows. A point* $x \notin \operatorname{singsupp} T$ *if we can find an* $f \in \mathcal{D}$ *such that* $x \notin \operatorname{supp} T - f$.

A careful argument using partitions of unity shows that if $\operatorname{singsupp} T = \emptyset$ then T is in fact a smooth function.

When working with distributions it is helpful to recall the following results.

Lemma 3.18. *If* T *is a distribution and* $\partial_j T = 0$ *for all* j *then* $T = c$ *where* c *is a constant.*

(We shall not prove this plausible result but like much else it may be found in Friedlander's beautiful little book, see section 2.1 [2].)

Lemma 3.19 (Theorem 28, A version of Taylor's theorem). *If* $f \in C^\infty(\mathbb{R})$ *and* $f(0) = f'(0) = \cdots = f^{m-1}(0) = 0$ *then we can find* $g \in C^\infty(\mathbb{R})$ *such that*

$$f(x) = x^m g(x)$$

for all x.

There is no difficulty (apart from notation) in extending Lemma 3.19 to higher dimensions. In the next section we shall make use of the following special case.

Lemma 3.20. *If* $f \in C^\infty(\mathbb{R}^n)$ *and* $f(0) = 0$ *then we can find* $g_j \in C^\infty(\mathbb{R}^n)$ *such that*

$$f(x) = \sum g_j x_j$$

for all x.

As an example of the use of Lemma 3.19 we do the following exercise.

Exercise 3.21. *Show that the general solution of*

$$xT = 0$$

is $T = c\delta_0$.

3.3 Fourier transforms and the Schwartz space

We would like to take Fourier transforms of distributions but if we look at the recipe of Lemma 3.7 we see that it fails since $f \in \mathcal{D}$ does not imply $\hat{f} \in \mathcal{D}$. (In fact, unless $f = 0$, $f \in \mathcal{D}$ implies $\hat{f} \notin \mathcal{D}$ but we do not need this result.)

Schwartz introduced another space of functions $\mathcal{S}(\mathbb{R}^n$ which behaves extraordinarily well under Fourier transforms.

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) \,:\, \sup_x (1 + \|x\|)^m |\partial^\alpha f(x)| < \infty \text{ for all } m \geq 0 \text{ and } \alpha\}.$$

Thus $\mathcal{S}$ 'consists of all infinitely differentiable functions all of whose derivative tend to zero faster than polynomial towards infinity'. Since $\mathcal{S} \supseteq \mathcal{D}$ we have plentiful supply of such functions.

If $f \in \mathcal{S}(\mathbb{R}^n)$ then the Fourier transform

$$\hat{f}(\xi) = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-ix.\xi} f(x)\, dx$$

We sometimes write $\mathcal{F}f = \hat{f}$. We also write

$$D_j f = -i \frac{\partial}{\partial x_j}$$

and abusing notation take $x_j f$ to be the function

$$x \mapsto x_j f(x).$$

It is clear that

$$\begin{aligned} D_j &: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \\ x_j &: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \end{aligned}$$

Familiar computations now give our first result.

Lemma 3.22 (Lemma 2). *The Fourier transform $f \mapsto \hat{f}$ takes $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$ and*

$$\begin{aligned} \widehat{D_j f} &= \xi_j \hat{f} \\ \widehat{x_j f} &= -D_j \hat{f} \end{aligned}$$

Lemma 3.22 can be used to give a neat proof of the Fourier inversion formula for $\mathcal{S}$. We use the following lemma.

Lemma 3.23 (Lemma 4). *If $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a linear map which commutes with x_j and D_j for all j then, writing I for the identity map,*

$$T = cI$$

for some constant c.

Our proof of Lemma 3.23 depends on a slight improvement of the Taylor series result Lemma 3.20.

Lemma 3.24 (Corollary 2). *If $f \in \mathcal{S}(\mathbb{R}^n)$ and $f(a) = 0$ then we can find $f_j \in \mathcal{S}(\mathbb{R}^n)$ such that*

$$f(x) = \sum_{j=1}^{n}(x_j - a_j)f_j(x)$$

for all x.

We can now prove our inversion theorem.

Theorem 3.25. *If $R : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is the map given by $Rf(x) = f(-x)$ then*

$$\mathcal{F}^2 = R.$$

Observe that this means that $\mathcal{F}^4 = I$ so $\mathcal{F}$ is invertible. Stating our result in a more traditional but equivalent manner we have the following theorem.

Theorem 3.26 (Theorem 9). *The Fourier transform*

$$f \mapsto \hat{f}$$

is an isomorphism of $\mathcal{S}$ onto itself with inverse given by

$$f(x) = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{ix.\xi} \hat{f}(\xi)\, dx.$$

If $f, g \in \mathcal{F}$, or more generally if the integral makes sense, we define the convolution $f * g$ by

$$f * g(x) = \int_{\mathbb{R}^n} f(x-t)g(t)\, dt$$

The standard results that we recall from mathematical methods hold.

Lemma 3.27 (Theorem 10). *If $f, g \in \mathcal{F}$*

(i) $\int f\hat{g} = \int \hat{f}g$,
(ii) $\int fg^* = \int \hat{f}(\hat{g})^*$, *(Parseval)*
(iii) $\widehat{f * g} = \hat{f}\hat{g}$,
(iv) $\widehat{fg} = \hat{f} * \hat{g}$,
(v) $f * g \in \mathcal{F}$.

3.4 Tempered distributions

Just as we constructed a space of distributions $\mathcal{D}'$ using functions from $\mathcal{D}$ as *test functions* so we can construct a space of distributions $\mathcal{S}'$ using functions from $\mathcal{S}$ as test functions.

Of course, we need to equip $\mathcal{S}$ with a notion of convergence.

Definition 3.28. *Suppose that $f_j \in \mathcal{S}$ for each j and $f \in \mathcal{S}$. We say that*

$$f_j \underset{\mathcal{S}}{\rightarrow} f$$

if for each $m \geq 0$ and α we have $(1 + \|x\|)^m \partial^\alpha (f_j(x) - f(x)) \rightarrow 0$ uniformly as $n \rightarrow \infty$.

We can now say that $T \in \mathcal{S}'$ to be a linear map $T : \mathcal{D} \rightarrow \mathbb{C}$ which is continuous in the sense that

$$f_n \underset{\mathcal{S}}{\rightarrow} f \text{ implies} T f_n \rightarrow T f.$$

As before we shall often write

$$T f = \langle T, f \rangle.$$

We can now develop the theory of *tempered distributions* $T \in \mathcal{S}'$ just as we did that of distributions $T \in \mathcal{D}'$. Using results in the previous subsection it is easy to check that $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ is a linear map which is continuous in the sense that

$$f_n \underset{\mathcal{S}}{\rightarrow} f \text{ implies } \mathcal{F} f_n \underset{\mathcal{S}}{\rightarrow} \mathcal{F} f \tag{1}$$

and that

$$\langle \mathcal{F} g, f \rangle = \langle g, \mathcal{F} f \rangle \tag{3}$$

for any $f, g \in \mathcal{S}$. (Thus, in the language of Lemma 3.7, $\mathcal{F}^t = \mathcal{F}$.) Thus, applying the result corresponding to Lemma 3.7 if $T \in \mathcal{S}'$ we can define $\mathcal{F} T$ by the equation

$$\langle \mathcal{F} T, f \rangle = \langle T, \mathcal{F} f \rangle,$$

or in the obvious corresponding notation

$$\langle \hat{T}, f \rangle = \langle T, \hat{f} \rangle,$$

There is no difficulty in proving results corresponding to Lemma 3.22.

Lemma 3.29. *(i) If $R : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is the extension to $\mathcal{S}'$ of the reflection function on $\mathcal{S}$ given by $Rf(x) = f(-x)$ then, working in $\mathcal{S}'$*

$$\mathcal{F}^2 = R.$$

(ii) The Fourier transform

$$f \mapsto \hat{f}$$

is an isomorphism of $\mathcal{S}'$ onto itself with inverse given by $\mathcal{F}^3$.

$$f(x) = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{ix.\xi} \hat{f}(\xi)\, dx$$

(iii) If $T \in \mathcal{S}'$ then, in an obvious notation,

$$\widehat{D_j T} = \xi_j \hat{T}$$
$$\widehat{x_j T} = -D_j \hat{T}$$

However we do not have to develop the theory of $\mathcal{S}'$ and $\mathcal{D}'$ separately since $\mathcal{D}$ sits nicely in $\mathcal{S}$ and so $\mathcal{S}'$ sits nicely in $\mathcal{D}'$.

Theorem 3.30. *(i) We have $\mathcal{D} \subseteq \mathcal{S}$. Moreover if that $f_j \in \mathcal{D}$ for each j and $f \in \mathcal{D}$ then*

$$f_j \underset{\mathcal{D}}{\to} f \text{ implies } f_j \underset{\mathcal{S}}{\to} f$$

(ii) We may consider any element $T \in \mathcal{S}'$ as an element of $\mathcal{D}$ in a natural manner.

Theorem 3.31 (Theorem 13). *(i) Given any $f \in \mathcal{S}$ we may find a sequence $f_j \in \mathcal{D}$ such that*

$$f_j \underset{\mathcal{S}}{\to} f$$

as $j \to \infty$.

(ii) If $T \in \mathcal{F}$ and $\langle T, f\rangle = 0$ whenever $f \in \mathcal{D}$ then $\langle T, f\rangle = 0$ whenever $f \in \mathcal{S}$.

(iii) Let $T,\ S \in \mathcal{F}$. If T and S are unequal when considered as members of $\mathcal{F}$ they remain unequal when considered as members $\mathcal{D}$.

Thus we may, and shall consider $\mathcal{S}'$ as a subspace of $\mathcal{D}'$. When we talk about distributions we shall mean members of $\mathcal{D}'$. When we talk about *tempered distributions* we shall mean members of $\mathcal{S}'$.

Example 3.32 (Example 6). *The delta function δ_0 is a tempered distribution. Its Fourier transform $\hat{\delta_0}$ is the constant function $(2\pi)^{-n/2}$.*

The constant function 1 is a tempered distribution. Its Fourier transform $\hat{1}$ is $(2\pi)^{n/2}\delta_0$.

The way we have defined $\hat{T}$ by Lemma 3.7 ensures that

$$\text{distributional Fourier transform}(f) = \text{classical Fourier transform}(f)$$

whenever $f \in \mathcal{F}$. In fact more is true.

Lemma 3.33 (Simple version of Proposition 2). *If f is continuous and $\int |f(x)|\,dx < \infty$ then*

$$\textit{distributional Fourier transform}(f) = \textit{classical Fourier transform}(f)$$

i.e.

$$\hat{f}(\xi) = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-ix.\xi} f(x)\,dx.$$

We shall also need the following simple result which the reader may already have met.

Lemma 3.34. *If f is continuous and $\int |f(x)|\,dx < \infty$ then $\hat{f}$ is continuous and bounded.*

As a corollary we have the following useful lemma.

Lemma 3.35 (Corollary 5). *If q is smooth and*

$$|q(\xi)| \leq C((1 + \|\xi\|^2)^{1/2})^{-n-l-\epsilon}$$

for some $\epsilon > 0$ then the Fourier transform and Fourier inverse transform of q is C^l.

4 Convolution and fundamental solutions

In this chapter we shall return at last to the topic of partial differential equations but first we need to study convolution of distributions.

4.1 Convolution

Unfortunately the notion of convolution of distributions is not as ‘clean’ as the rest of the theory but it is essential.

Recall that if f and g are ‘nice functions’ (e.g. if they are both Schwartz functions) we define the convolution $f * g$ by

$$f * g(x) = \int_{\mathbb{R}^n} f(x-t)g(t)\,dt.$$

Convolution plays an essential role in probability theory, signal processing, linear partial differential equations (as we shall see) and many other parts of analysis. We proved in Lemma 3.27 (iii) that $\widehat{f * g} = \hat{f}\hat{g}$, and it can be argued that importance of the Fourier transform is that it converts convolution into multiplication.

However if we try out this formula on the tempered distribution 1 we get

$$\widehat{1 * 1} = \hat{1}\hat{1} = (2\pi)^n \delta_0 \delta_0$$

an ‘equation’ in which neither side makes sense (at least within distribution theory as developed here). *In general we can not multiply two distributions and we can not convolve two distributions.*

Having said this, there are many circumstances in which it is possible to convolve two distributions. To see this observe that under favourable circumstances, if f, g and h are all nice functions

$$\begin{aligned}
\langle f * g, h\rangle &= \int \left(\int_{\mathbb{R}^n} f(s-y)g(y)\,dy\right) h(s)\,ds \\
&= \int\int f(s-y)g(y)h(y)\,dy\,ds \\
&= \int\int f(s-y)g(y)h(s)\,ds\,dy \\
&= \int\int f(x)g(y)h(x+y)\,dx,\,dy \\
&= \langle f(x), \langle g(y), h(x+y)\rangle\rangle.
\end{aligned}$$

We shall attempt to generalise this. Our treatment is not quite as general as that given by Joshi and Wassermann in section 4.4 of their notes but follows the treatment given by Friedlander in his book [2]. The underlying idea is identical and either treatment can be used in examination questions.

Looking at the discussion of convolution in ‘favourable circumstances’ we observe that we have made use of the ‘exterior product’ $(f,g) \mapsto f \otimes g$ where

$$f \otimes g(x,y) = f(x)g(y).$$

We observe that f, g and $f \otimes g$ are considered as distributions then

$$\langle f \otimes g, h \rangle = \int \int f(x)g(y)h(x,y)\,dx\,dy. \qquad \dagger$$

The exterior product carries over easily to distributions. If $S \in \mathcal{D}(\mathbb{R}^n)$ $T \in \mathcal{D}(\mathbb{R}^m)$ we set

$$\langle S \otimes T, h \rangle = \langle S, \langle T, h_x \rangle \rangle$$

where $h_x(y) = h(x,y)$ and $h \in \mathbb{R}^{n+m}$. Using dummy variables

$$\langle S \otimes T(x,y), h(x,y) \rangle = \langle S(x), \langle T(y), h(x,y) \rangle \rangle.$$

Lemma 4.1. *With the notation and hypotheses of the previous paragraph, $S \otimes T$ is a well defined member of $\mathcal{D}(\mathbb{R}^{m+n})$.*

Suppose now $S,\ T \in \mathcal{D}(\mathbb{R}^n)$. If $h \in \mathcal{D}(\mathbb{R}^n)$ the discussion of convolution in 'favourable circumstances' suggests that we should look at the function $\tilde{h} : \mathbb{R}^{2n} \to \mathbb{C}$ given by $\tilde{h}(x,y) = h(x+y)$ and apply $S \otimes T$ to $\tilde{h}$. Unfortunately, although $\tilde{h} \in C^\infty(\mathbb{R}^{2n})$, $\tilde{h}$ is not of compact support unless $h = 0$.

Exercise 4.2. *Prove the two statements made in the last sentence.*

So far we have only allowed elements of $\mathcal{D}'$ to operate on elements of $\mathcal{D}$ so we appear to be stuck.

However, *under certain circumstances*, elements of $\mathcal{D}'$ will operate on more general C^∞ functions in a natural manner.

Lemma 4.3. *(i) Suppose $T \in \mathcal{D}'(\mathbb{R}^n)$ and $f \in C^\infty(\mathbb{R}^n)$. If $\operatorname{supp} f \cap \operatorname{supp} T$ is compact then there is a natural definition of $\langle T, f \rangle$.*

(ii) Suppose $T \in \mathcal{D}'(\mathbb{R}^n)$ and $f_j \in C^\infty(\mathbb{R}^n)$. Suppose further that there exists a compact set K such that $K \supseteq \operatorname{supp} f_j \cap \operatorname{supp} T$ for all j. Then if $\partial^\alpha f_j \to 0$ uniformly on K for each α it follows that $\langle T, f_j \rangle \to 0$ as $j \to \infty$.

We have arrived at our goal.

Theorem 4.4. *If $S,\ T \in \mathcal{D}'(\mathbb{R}^n)$ and $\operatorname{supp} S$ or $\operatorname{supp} T$ is compact then we may define $S * T$ as follows. If $h \in \mathcal{D}(\mathbb{R}^n)$ define $\tilde{h} : \mathbb{R}^{2n} \to \mathbb{C}$ by $\tilde{h}(x,y) = h(x+y)$, the expression $\langle S \otimes T, \tilde{h} \rangle$ is well defined in the sense of Lemma 4.3 so we set*

$$\langle S * T, h \rangle = \langle S \otimes T, \tilde{h} \rangle$$

The map $S * T$ so defined is a distribution.

Lemma 4.5. *With the notation and hypotheses of Theorem 4.4*

$$\operatorname{supp} S * T \subseteq \operatorname{supp} S + \operatorname{supp} T.$$

(Here $\operatorname{supp} S + \operatorname{supp} T = \{x + y \,:\, x \in \operatorname{supp} S,\ y \in \operatorname{supp} T\}$.)

Note that if the hypotheses of Lemma 4.4 hold then using dummy variables we have the elegant formula

$$\langle S * T(t), f(t)\rangle = \langle S(x)\langle T(y), f(x+y)\rangle.$$

The hypotheses of Theorem 4.4 can clearly be relaxed somewhat, so Joshi and Wassermann refer to 'convolvable distributions'.

We summarise some of the useful facts about convolutions.

Lemma 4.6. *(i) If T and S are distributions at least one of which has compact support then*

$$T * S = S * T.$$

(ii) We recall that δ_0 has compact support. If T is any distribution

$$\delta_0 * T = T.$$

(iii) If R, T and S are distributions at least two of which have compact support then

$$R * (S * T) = (R * S) * T.$$

(iv) If T and S are distributions at least one of which has compact support then

$$\partial^\alpha (T * S) = (\partial^\alpha T) * S = T * (\partial^\alpha S).$$

Note that this gives $\partial^\alpha u = \partial^\alpha \delta_0 * u$, so applying a differential operator is equivalent to convolving with some distribution.

One of the key facts about convolution is that it smooths.

Lemma 4.7. *If T is a distribution and f an m times continuously differentiable function and at least one of the two has compact support then $\langle T_y, f(x-y)\rangle$ is a function and, as distributions,*

$$\langle T_y, f(x-y)\rangle = T * f(x).$$

Thus $T * f$ is an m times continuously differentiable function.

We shall not need the next result so I leave it as an exercise but the discussion would seem incomplete without it.

Exercise 4.8. *(i) If T is a distribution of compact support then, using the extension of Lemma 4.3,*

$$\hat{T}(\xi) = (2\pi)^{-n/2}\langle T(x), \exp(i\xi.x)\rangle.$$

Further $\hat{T}$ is a continuous function.

(ii) If S and T are distributions of compact support

$$\widehat{S * T}(\xi) = \widehat{T}(\xi)\widehat{S}(\xi).$$

4.2 Fundamental solutions

We now have the tools to make substantial progress with the study of linear partial differential equations with constant coefficients. The key notion is that of a fundamental solution.

Definition 4.9. *A* fundamental solution *of $P(D)$ where $P(\xi_1, \ldots, \xi_n)$ is a polynomial with constant coefficients and $P(D) = P(-i\partial_1, \ldots, -i\partial_n)$ is a distributional solution, T, of*

$$P(D)T = \delta_0.$$

Note that we do not say that a fundamental solution is unique.

Example 4.10. *If $n = 1$ and $P(x) = ix$ then $E_c = H + c$ is a fundamental solution where H is the Heaviside function and c is any constant.*

The interest of fundamental solutions is shown by the next lemma.

Lemma 4.11. *We use the notation of Definition 4.9*

(i) If T is a distribution of compact support and E is a fundamental solution of $p(D)$ then $S = E * T$ is a solution of $P(D)S = T$.

(ii) If f is an m times continuously differentiable function of compact support and E is a fundamental solution of $P(D)$ then $S = E * f$ is an m times continuously differentiable solution of $P(D)S = T$.

Example 4.12. *(i) We continue with Example 4.10. Suppose f is a continuous function of compact support and we wish to solve*

$$u'(x) = f(x).$$

Applying Lemma 4.11 we obtain

$$u(x) = c\int_{-\infty}^{\infty} f(t)\,dt + \int_{-\infty}^{x} f(t)\,dt.$$

(ii) Particular interest attaches to the two cases $c = 0$ when the solution is

$$u(x) = \int_{-\infty}^{x} f(t)\,dt \tag{A}$$

and $c = -1$ when the solution is

$$u(x) = -\int_{x}^{\infty} f(t)\,dt \tag{B}$$

since equation (A) produces a solution for $u'(x) = f(x)$ valid whenever $f(x) = 0$ for x large and positive and equation (B) produces a solution for $u'(x) = f(x)$ valid whenever $f(x) = 0$ for x large and negative.

One of the points to note about Example 4.12 (ii) is that it shows the advantage of extending the notion of convolution as far as as we can.

Example 4.13. *If S and T are distributions on $\mathbb{R}$ such that we can find a real number R such that*

$$\operatorname{supp} S,\ \operatorname{supp} T \subseteq (-R, \infty).$$

Explain in as much detail as you consider desirable why we can define $S * T$.

It also shows that it may be desirable to consider different fundamental solutions for different kinds of problems. We can think of fundamental solutions as inverses on different classes of functions in accordance with the formula

$$E * P(D) = P(D) * E = \delta_0.$$

4.3 Our first fundamental solution

If E is a fundamental solution of $P(D)$ then by definition

$$P(D)E = \delta_0$$

so, provided that E is a tempered distribution, we may take Fourier transforms to obtain

$$P(\xi)\hat{E}(\xi) = (2\pi)^{-n/2},$$

so, at least formally,

$$\hat{E}(\xi) = \frac{1}{(2\pi)^{n/2} P(\xi)}.$$

If $P(\xi)$ is never zero the argument can easily be reversed in rigorous manner to give the following theorem.

Theorem 4.14. *If P is a polynomial with no real zeros then there is a fundamental solution given by the tempered distribution u with*

$$\hat{u}(\xi) = \frac{1}{(2\pi)^{n/2} P(\xi)}.$$

There is only one fundamental solution which is also a tempered distribution.

Exercise 4.15. *Show that Theorem 4.14 applies to the operator $-\Delta + 1$. Use Theorem 4.14 to solve the one dimensional case*

$$-u'' + u = \delta_0$$

and verify by direct calculation that you have indeed got a solution.

A large part of the theory of partial differential equations consists of different ways of deal with the zeros of $P(\xi)$.

4.4 The parametrix

If $P(D)$ is an elliptic operator, $P(\xi)$ may have real zeros but they must lie in a bounded set. Using this we can obtain something 'almost as good' as a fundamental solution.

Theorem 4.16 (Theorem 14). *If $P(D)$ is elliptic, ϕ a smooth function of compact support identically 1 on a neighbourhood of the zero set of $P(\xi)$ then we can find a tempered distribution E with*

$$\hat{E}(\xi) = \left(\frac{1}{2\pi}\right)^{n/2} \frac{1}{P(\xi)}(1-\phi)(\xi).$$

We have

$$P(D)E = \delta_0 + f$$

where $f \in \mathcal{S}(\mathbb{R}^n)$. Also

$$\operatorname{singsupp} E \subseteq \{0\}.$$

Lemma 4.17. *If $P(D)$ is elliptic and U a neighbourhood of 0 then we can find a tempered distribution E with $\operatorname{supp} E \subseteq U$, $\operatorname{singsupp} E \subseteq \{0\}$, and*

$$P(D)E = \delta_0 + f$$

where $f \in \mathcal{S}(\mathbb{R}^n)$.

We call a distribution like E in Theorem 4.16 or Lemma 4.17 a parametrix.

It will, no doubt, have occured to the reader that if we solve a partial differential equation

$$P(D)u = f$$

in the distributional sense what we get is a distribution u and not necessarily a classical function even if f is a classical function. The next theorem removes this problem in certain cases.

Theorem 4.18 (Theorem 15, special case of Weyl's lemma). *If $u \in \mathcal{D}'$ and $P(D)$ is an elliptic differential operator then u and $P(D)u$ have the same singular support. In particular if $P(D)u$ is smooth so is u.*

4.5 Existence of the fundamental solution

Except when we discussed first order equations, this course has been about linear partial differential equations with constant coefficients. It is an important fact, which we shall not prove, that in this case there is always a fundamental solution.

Theorem 4.19 (Theorem 16). *If $P(\xi_1, \ldots, \xi_n)$ is a polynomial with constant coefficients and $P(D) = P(-i\partial_1, \ldots, -i\partial_n)$ then we can find a distribution E such that*

$$P(D)E = \delta_0.$$

A proof of Theorem 4.19 is given in section 10.4 of Friedlander's book [2]. We shall prove an important special case.

Theorem 4.20. *If $P(\xi_1, \ldots, \xi_n)$ is a polynomial with constant coefficients and there exists an $a \in \mathbb{R}^n$ and an $\eta > 0$ such that*

$$|P(\xi + ia)| \geq \eta$$

for all $\xi \in \mathbb{R}^n$ then we can find a distribution E such that

$$P(D)E = \delta_0.$$

We prove this result by contour pushing so we need a result on analyticity and growthR

Lemma 4.21 (Proposition 5.5, the Paley-Wiener estimate). *If $f \in \mathcal{D}(\mathcal{R}^n)$ then there is a natural extension of $\hat{f}$ to a function on $\mathbb{C}^n$ which is analytic in each variable. If* $\operatorname{supp} f \subseteq B(0,r)$ *then*

$$|\hat{f}(z)| \leq C_N(1+\|z\|)^{-N}e^{r|\Im z|}$$

for all z and integers $N \geq 0$. Here C_N is a constant depending only on N and f.

With the proof of Theorem 4.20 we come to the end of our general account of linear constant coefficient partial differential equations. For the rest of the course we shall study the Laplacian, the heat and the wave operator. We close this section by looking at the heat and wave operator in the context of Theorem 4.20.

Example 4.22 (Example 5.7). *If*

$$P(D) = \frac{\partial}{\partial t} - \sum \frac{\partial^2}{\partial x_j^2}$$

(the heat operator) then if $a < 0$

$$|P(\tau + ia, \xi)| \geq -a$$

but if $a \geq 0$ there exists a $\xi \in \mathbb{R}^n$ such that

$$|P(\tau + ia, \xi)| = 0.$$

Example 4.23 (Example 5.8). *If*

$$P(D) = \frac{\partial^2}{\partial t^2} - \sum \frac{\partial^2}{\partial x_j^2}$$

(the wave operator) then if $a \neq 0$

$$|P(\tau + ia, \xi)| \geq a^4 > 0$$

for $\tau \in \mathbb{R}$, $\xi \in \mathbb{R}^n$.

Note that our shifting contour argument gives one fundamental solution for the heat operator but two for the wave equation.

5 The Laplacian

5.1 A fundamental solution

The reader can probably guess a fundamental solution for the Laplacian. We shall proceed more slowly in order to introduce some useful ideas. Observe first that by Theorem 4.18 any fundamental solution E will have $\operatorname{singsupp} E \subseteq \{0\}$. Since the Laplacian is rotationally invariant an averaging argument shows (as one might fairly confidently guess) that there must be a rotationally invariant fundamental solution.

Exercise 5.1. *We could now proceed as follows. Suppose $f(\mathbf{x}) = g(r)$ with $r = \|\mathbf{x}\|$. Show that*

$$(\Delta f)(\mathbf{x}) = \frac{1}{r^{n-1}} \frac{d}{dr} r^{n-1} f g(r).$$

If f is smooth and $\Delta f = 0$ on $\mathbb{R}^n \setminus \{0\}$ find f.

Instead of following the path set out in Exercise 5.1 we use another nice idea — that of homogeneity. The following remark is due to Euler.

Lemma 5.2. *Let $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ be continuously differentiable and m be an integer. The following statements are equivalent.*

(i) $f(\lambda x) = \lambda^m f(x)$ for all $\lambda > 0$. (In other words, f is homogeneous of degree m.)

(ii) $\left(\sum_{j=1}^{n} x_j \frac{\partial}{\partial x_j} - m \right) f(x) = 0.$

In view of Lemma 5.2 we make the following definition.

Definition 5.3. *If $u \in \mathcal{D}'$ then u is homogeneous of degree $m \in \mathbb{C}$ if*

$$\left(\sum_{j=1}^{n} x_j \frac{\partial}{\partial x_j} - m \right) f(x) = 0.$$

Lemma 5.4. *(i) The delta function δ_0 on $\mathbb{R}^n$ is homogeneous of degree $-n$.*

(ii) If is $u \in \mathcal{D}'$ is homogeneous of degree m then $\partial_{x_j} u$ is homogeneous of degree $m - 1$.

Looking at Lemma 5.4 we guess that one fundamental solution E_n of the Laplacian will be radially symmetric of degree $2 - n$ and we arrive at the guess

$$E_n = C_n \|x\|^{2-n}.$$

If $n \neq 2$ we shall now define $E_n = C_n \|x\|^{2-n}$ and verify that it is indeed a fundamental solution for the appropriate choice of C_n.

Lemma 5.5. *Suppose $n \neq 2$.*

(i) E_n is a well defined distribution according to our usual formula

$$\langle E_n, f\rangle = \int_{\mathbb{R}^n} E_n(x) f(x)\, dx.$$

(ii) $\operatorname{supp} \Delta E_n \subseteq \{0\}$.

(iii) ΔE_n is a multiple of δ_0.

(iv) If $C_n = ((2-n)\omega_{n-1})^{-1}$ where ω_{n-1} is the area of the surface of the unit sphere in $\mathbb{R}^n$ then $\Delta E_n = \delta_0$.

If $n = 2$ we use previous knowledge to guess that

$$E_2(\mathbf{x}) = -\frac{1}{2\pi} \log \|x\|.$$

The verification that this is indeed a fundamental solution follows the pattern of Lemma 5.5.

Theorem 5.6 (Theorem 5.1). *The Laplacian on $\mathbb{R}^n$ has the fundamental solution*

$$E(x) = \frac{1}{(2-n)\omega_{n-1}\|x\|^{2-n}}$$

when $n \neq 2$ and

$$E(x) = -\frac{1}{2\pi} \log \|x\|$$

for $n = 2$.

5.2 Identities and estimates

The reader will be familiar with the Gauss divergence theorem and the associated Green's identities. There are problems (or at least apparent problems) in proving this for complicated domains but the standard mathematical method proof is rigorous when applied to the ball which is all we shall need.

Theorem 5.7 (Theorem 5.2, Gauss divergence theorem). *Let $g : \overline{B(0,r)} \to \mathbb{R}^n$ be continuous on $\overline{B(0,r)}$ and continuously differentiable on $B(0,r)$. Then*

$$\int_{\|x\|<r} \nabla . g \, dV = \int_{\|x\|=r} g . n_x \, dS$$

where $n_x = \|x\|^{-1}x$ is the outward normal.

Green's identities correspond to integrating by parts and are more of a method than a theorem.

Lemma 5.8 (Proposition 5.1, Green's identities). *If $u, v \in C^2(\overline{B(0,r)})$*

(a) $\displaystyle\int_{\|x\|<r} v\Delta u\, dV = -\int_{\|x\|<r} \nabla v.\nabla u\, dV + \int_{\|x\|=r} v\nabla u.n_x\, dS.$

(b) $\displaystyle\int_{\|x\|<r} v.\Delta u - u.\Delta v\, dV = \int_{\|x\|=r} (v\nabla u - u\nabla v).n_x\, dS.$

Much of the early work on the Laplacian used the following observation.

Lemma 5.9 (Proposition 5.2, energy estimate). *If $u \in C^2(\overline{B(0,r)})$*

$$\int_{\|x\|<r} \|\nabla u\|^2\, dV = \int_{\|x\|=r} u\nabla u.n_x\, dS - \int_{\|x\|<r} u\Delta u\, dV.$$

Theorem 5.10. *(i) The Dirichlet problem:- find $u \in C^2(\overline{B(0,r)})$ subject to the conditions*

$$\begin{aligned} \Delta u &= f \qquad && \text{on } B(0,r) \\ u &= g \qquad && \text{on } \{x : \|x\| = r\} \end{aligned}$$

has at most one solution.

(ii) The Neumann problem:- find $u \in C^2(\overline{B(0,r)})$ subject to the conditions

$$\begin{aligned} \Delta u &= f \qquad && \text{on } B(0,r) \\ \nabla u.n_x &= h \qquad && \text{on } \{x : \|x\| = r\} \end{aligned}$$

has at most one solution up to a constant.

We can do much better than Theorem 5.10 (i) by using the weak maximum principle.

Theorem 5.11 (Proposition 5.3, weak maximum principle). *Let Ω be a non-empty bounded open set in $\mathbb{R}^n$ with boundary $\partial\Omega$. If $u \in C(\overline{\Omega})$ and u is twice continuously differentiable in Ω with $\Delta u \geq 0$ in Ω then*

$$\sup_{x\in\overline{\Omega}} u(x) = \sup_{x\in\partial\Omega} u(x).$$

Theorem 5.12 (Uniqueness for the Dirichlet problem). *The Dirichlet problem:- find $u \in C(\overline{\Omega})$ with u is twice continuously differentiable in Ω subject to the conditions*

$$\begin{aligned} \Delta u &= f \qquad && \text{on } \Omega \\ u &= g \qquad && \text{on } \{x : \|x\| = r\} \end{aligned}$$

has at most one solution.

The point here is not so much the generality of Ω as that we only need u differentiable on Ω.

If u is twice continuously differentiable on some open set Ω with $\Delta u =$ we say that u is a harmonic function. The reader is already aware of the connection with analytic functions. As an example we give another proof of a result from 1B.

Lemma 5.13 (Maximum principle for analytic functions). *If Ω is a bounded non-empty open set in $\mathbb{C}$ and $f : \overline{\Omega} \to \mathbb{C}$ is a continuous function, analytic on Ω then*

$$\sup_{z\in\overline{\Omega}} |f(z)| = \sup_{z\in\partial\Omega} |f(z)|.$$

Lemma 5.14 (Gauss mean value theorem). *(a) The value of a harmonic function at a point is equal to its average over any sphere (boundary of a ball) centred at that point.*

(b) The value of a harmonic function at a point is equal to its average over any ball centred at that point.

We shall see in Lemma 6.6 that condition (a) actually characterises harmonic functions.

Lemma 5.15 (Proposition 5.5, the strong maximum principle). *If Ω is a bounded open connected set and $u \in C(\overline{\Omega})$ satisfies*

$$u(\xi) \leq \frac{1}{\textit{Area sphere radius } \rho} \int_{\|x-\xi\|=\rho} u(x)\, dS$$

for all ρ sufficiently small (depending on ξ) and all $\xi \in \Omega$ then either u is constant or $u(\xi) < \sup_{x\in\partial\Omega} u(x)$ for all $\xi \in \Omega$.

We shall probably not have time to develop the matter much further but the associated definition (Definition 5.16) is important in later work.

Definition 5.16. *If Ω is open and $u \in C(\overline{\Omega})$ satisfies*

$$u(\xi) \leq \frac{1}{\textit{Area sphere radius } \rho} \int_{\|x-\xi\|=\rho} u(x)\, dS$$

for all ρ sufficiently small (depending on ξ) and all $\xi \in \Omega$ then we say that u is subharmonic.

Lemma 5.17. *If Ω is open and $u \in C(\overline{\Omega})$ and u is twice differentiable on Ω with $\Delta u \geq 0$ then u is subharmonic.*

5.3 The dual Dirichlet problem for the unit ball

Let Ω be a bounded open set with smooth boundary $\partial\Omega$.

(a) The *Dirichlet problem* asks for a solution of $\Delta f = 0$ in Ω with $f = h$ on $\partial\Omega$.

(b) The *dual Dirichlet problem* asks for a solution of $\Delta f = 0$ in Ω with $f = g$ on $\partial\Omega$.

We have been deliberately vague about the exact nature of f, h and g. Maintaining the same vagueness we see that solving one problem is 'more or less' equivalent to solving the other.

Let us try to solve the dual Dirichlet problem for the unit ball B.

$$\begin{aligned} \Delta f &= g \qquad \text{for } \|x\| < 1 \\ f &= 0 \qquad \text{for } \|x\| = 1 \end{aligned}$$

To make life easier for ourselves we shall initially assume $g \in C(\overline{B})$ and g infinitely differentiable on B. (In Lemma 6.3 we shall see that that the formula obtained can be extended to a much wider class of functions.) We shall work in $\mathbb{R}^n$ with $n \geq 3$.

We start by defining

$$\begin{aligned} \tilde{g} &= g \qquad \text{for } \|x\| \leq 1 \\ \tilde{g} &= 0 \qquad \text{for } \|x\| > 1. \end{aligned}$$

If E is the fundamental solution obtained earlier we start by looking at $f_1 = E * \tilde{g}$. We know that, in a distributional sense,

$$\Delta E * \tilde{g} = \tilde{g}$$

but we wish to have a classical solution.

Lemma 5.18 (Lemma 5.1). *(i) $f_1 \in C^1(\mathbb{R}^n)$ and $\partial_j f_1 = \partial_j E * \tilde{g}$.*

(ii) There is a constant such that $|f_1(x)| \leq A\|x\|^{2-n}$, $|\partial_j f_1(x)| \leq A\|x\|^{1-n}$ for $\|x\|$ large.

Lemma 5.19 (Lemma 5.2). *We have f_1 infinitely differentiable on $\mathbb{R}^n \setminus \{x : \|x\| = 1\}$ and*

$$\begin{aligned} \Delta f_1 &= g \qquad \textit{for } \|x\| < 1 \\ \Delta f_1 &= 0 \qquad \textit{for } \|x\| > 1. \end{aligned}$$

We thus have $\Delta f_1 = g$ in B but the boundary conditions are wrong. To get round this we use Kelvin's method of reflections. (If you have done electrostatics you will recognise this for dimension $n = 2$.) We put

$$(Kf)(x) = \|x\|^{2-n} f\left(\frac{x}{\|x\|^2}\right).$$

Lemma 5.20. *If we set $f_2 = Kf_1$ and $f = f_1 - f_2$ then*
(i) $\Delta f_2 = 0$ in B.
(ii) $f_1 = f_2$ on ∂B.
(iii) f solves our Dirichlet problem.

The proof of Lemma 5.20 (i) involves substantial calculation.

Lemma 5.21 (Lemma 5.3). *(i) If f is twice differentiable*

$$\Delta(Kf) = \|x\|^{-4} K(\Delta f).$$

(ii) If f is once differentiable and we write $r\dfrac{\partial}{\partial r} = \sum x_j \frac{\partial}{\partial x_j}$ we have

$$r\frac{\partial}{\partial r} = (-n+2)Kf - K\left(r\frac{\partial}{\partial r}\right).$$

Lemma 5.21 shows us that f_2 is harmonic in $B \setminus \{0\}$ but a little more work is needed to establish the full result.

Lemma 5.22 (Lemma 4). *The function f_2 is harmonic on B.*

Doing the appropriate simple substitutions we obtain our final result.

Theorem 5.23 (Theorem 5.4). *We work in $\mathbb{R}^n$ with $n \geq 3$. If we set*

$$G(x,y) = \frac{\|x-y\|^{2-n}}{\omega_{n-1}(n-2)} - \frac{\|x\|^{2-n}\|\|x\|^{-2}x - y\|^2}{\omega_{n-1}(n-2)}$$

then

$$f(x) = \int_{\|y\|<1} G(x,y)g(y)\,dy$$

solves the dual Dirichlet problem.

The reader will recognise $G(x,y)$ as a Green's function.

We shall not have time to do the case $n = 2$ but the reader will find no problem in redoing our arguments to obtain.

Theorem 5.24. *We work in $\mathbb{R}^2$. If we set*

$$G(x,y) = \frac{1}{2\pi}(\log\|x-y\| - \log\|\|x\|^{-2}x - y\|)$$

then

$$f(x) = \int_{\|y\|<1} G(x,y)g(y)\,dy$$

solves the dual Dirichlet problem.

6 Dirichlet's problem for the ball and Poisson's formula

We noted earlier that a solution of the dual Dirichlet problem will give us a solution for Dirichlet's problem.

Lemma 6.1. *Consider the Dirichlet problem for the unit ball B.*

$$\begin{aligned} \Delta f &= 0 \qquad \text{for } \|x\| < 1 \\ f &= h \qquad \text{for } \|x\| = 1 \end{aligned}$$

with h infinitely differentiable.

If $n \geq 3$ the solution is given by

$$f(x) = \int_{\|y\|=1} P(x, y) h(y) \, dS(y)$$

where $P(x, y) = \dfrac{d}{dn} G(x, y)$ the directional derivative along the outward normal.

Lemma 6.2. *Consider Let the Dirichlet problem for the unit ball B.*

$$\begin{aligned} \Delta f &= 0 \qquad \text{for } \|x\| < 1 \\ f &= h \qquad \text{for } \|x\| = 1 \end{aligned}$$

with h infinitely differentiable.

If $n \geq 2$ the solution is given by

$$f(x) = \int_{\|y\|=1} P(x, y) h(y) \, dS(y)$$

where

$$P(x, y) = \frac{1}{\omega_{n-1}} \frac{1 - \|x\|^2}{\|x - y\|^n}$$

We call P the Poisson kernel. We can improve Lemma 6.2 and provide a proof of the improvement which is essentially independent of the work already done.

Lemma 6.3. *Consider the Dirichlet problem for the unit ball B.*

$$\begin{aligned} \Delta f &= 0 \qquad \text{for } \|x\| < 1 \\ f &= h \qquad \text{for } \|x\| = 1 \end{aligned}$$

with h continuous.

If $n \geq 2$ the solution is given by

$$f(x) = \int_{\|y\|<1} P(x,y)h(y)\,dy$$

We need a preliminary lemma.

Lemma 6.4. *(i) If y is fixed with $\|y\| < 1$ then $P(x,y)$ is harmonic in x for $\|x\| < 1$.*

(ii) $\displaystyle\int_{\|y\|=1} P(x,y)\,dS(y) = 1$

Dilation and translation gives the following result.

Lemma 6.5. *The Dirichlet problem for any ball with continuous data has a unique solution.*

Lemma 6.5 is a very useful tool as the proof of the converse of Lemma 5.14 (a) shows.

Lemma 6.6 (Proposition 5.7). *Let Ω be open and $u : \Omega \to \mathbb{R}$ continuous. If the value of u at any point $x \in \Omega$ is equal to its average over any sufficiently small sphere centred at that x then u is harmonic.*

Lemma 6.7 (Corollary 5.2). *The uniform limit of harmonic functions is itself harmonic.*

(Compare Morera's theorem in complex variable theory.)

Poisson's solution for the Dirichlet problem for the ball gives a useful inequality.

Theorem 6.8 (Harnack's inequality, Theorem 5.5). *If $f \in C(\overline{B(0,R)})$ is everywhere non-negative and f is harmonic on $B(0,R)$ then if $\|x\| = r < R$ we have*

$$\frac{R^{n-2}(R-r)}{(R+r)^{n-1}} f(0) \leq f(x) \leq \frac{R^{n-2}(R+r)}{(R-r)^{n-1}} f(0).$$

Lemma 6.9 (Harnack's convergence theorem, Theorem 5.6). *Let Ω be an open set in $\mathbb{R}^n$. Suppose that u_m is harmonic on Ω and that*

(a) $u_m(x) \leq u_{m+1}(x)$ for all $m \geq 1$

(b) $u_m(x) \to u(x)$ as $m \to \infty$

for all $x \in \Omega$. Then $u_m \to u$ uniformly on bounded sets and u is harmonic.

7 The wave equation

Recall that the wave operator $\Box$ given by

$$\Box u(t,\mathbf{x}) = \frac{\partial^2 u}{\partial t}(t,\mathbf{x}) - \Delta_{\mathbf{x}} u(t,\mathbf{x})$$

where $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ and $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is a suitable function.

Guided by our knowledge of physics we concentrate on solving two problems. The first is the forcing problem.

$$\Box u = f \ \text{ subject to } u(t,x) = 0 \text{ for } t \leq 0$$

where $f(t,x) = 0$ for $t \leq 0$. (Later we shall replace $t \leq 0$ by the slightly more general t large and negative.) The second is the Cauchy problem

$$\Box u = f \ \text{ subject to } u(x,0) = u_0(x) \text{ and } \frac{\partial u}{\partial t}(x,0) = u_1(x).$$

The two problems are, as one might expect, closely related and we shall solve the second problem by reducing it to the first.

Looking at Theorem 4.20 and Example 4.23 we see that we know two fundamental solutions E_+ and E_- are given by

$$\langle E_\pm, \phi\rangle = \left(\frac{1}{2\pi}\right)^{(n+1)/2} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \frac{\hat{\phi}(-(\tau \pm i\alpha), -\xi)}{\|\xi\|^2 - (\tau \pm i\alpha)^2}\, d\tau\, d\xi,$$

where α is a strictly positive number. We can make a good guess at what E_+ and E_- are though, unless we have a good background in physics, our guess will be wrong when $n = 2$ (and indeed for all $n \geq 4$)! We therefore start with the case $n = 3$ when our guess turns out to be right and look at E_+. Since our guess suggests that E_+ is not a function but a distribution we look at $\langle E_\pm, \phi\rangle$ rather than E_+ by itself.

Lemma 7.1. *For all n have*

$$\langle E_\pm, \phi\rangle = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} \int_0^\infty \hat{\phi}_t(-\xi) \frac{\sin(\|\xi\| t)}{\|\xi\|}\, d\xi\, dt$$

We now see if our guess is appropriate.

Lemma 7.2. *Let $n = 3$. The formula*

$$\langle u, f\rangle = \int_{\|x\|=t} f(x)\, dS(x)$$

defines a distribution of compact support. We have

$$\hat{u}(\xi) = \frac{4\pi t}{\|\xi\|} \sin(\|\xi\| t)$$

Putting everything together we obtain our solution.

Lemma 7.3. *If $n = 3$ the the forward fundamental solution E_+ is given by*

$$\langle E_+, f\rangle = \int_0^\infty \frac{1}{4\pi t}\int_{\|x\|=t} f(t,x)\,dS(x)\,dt.$$

The formula just given can be rewritten (we write $E_{3,+}$ to emphasise that $n = 3$) in various ways such as

$$\langle E_{3,+}, \phi\rangle = \int_0^\infty \frac{t}{4\pi}\int_{\|x\|=1} f(xt,t)\,dS(x)\,dt = \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{f(y,\|y\|)}{\|y\|}\,dy.$$

We now turn our attention to the case $n = 2$. Observe that if $a, b : \mathbb{R}^2 \to \mathbb{R}$ are reasonable functions and we define $\tilde{a}, \tilde{b} : \mathbb{R}^3 \to \mathbb{R}$ by $\tilde{a}(t, x_1, x_2, x_3) = a(t, x_1, x_2)$ and $\tilde{b}(t, x_1, x_2, x_3) = b(t, x_1, x_2)$ then $\Box a = b$ if and only if $\Box\tilde{a} = \Box\tilde{b}$. This strongly suggests the guess that

$$\langle \epsilon_{2,+}, f\rangle = \langle E_{3,+}, \tilde{f}\rangle.$$

The main point at issue is whether this actually defines a distribution since even when $f \in \mathcal{D}$ we will not (unless $f = 0$) have $\tilde{f} \in \mathcal{D}$. However the argument of Lemma 4.3 applies since $\operatorname{supp} E_{3,+} \subseteq \{(t,x) : \|x\| \leq t\}$ so $\epsilon_{2,+}$ is well defined. Now

$$\begin{aligned}
\langle \Box\epsilon_{2,+}, f\rangle &= \langle \epsilon_{2,+}, \Box f\rangle \\
&= \langle E_{3,+}, \widetilde{\Box f}\rangle \\
&= \langle E_{3,+}, \Box\tilde{f}\rangle \\
&= \langle \Box E_{3,+}, \tilde{f}\rangle \\
&= \langle \delta_0, \tilde{f}\rangle \\
&= \langle \delta_0, f\rangle
\end{aligned}$$

for all $f \in \mathcal{D}(\mathbb{R}^2)$ so $\Box\epsilon_{2,+} = \delta_0$.

Thus $\epsilon_{2,+}$ is a fundamental solution. It is easily checked that $\operatorname{supp}\epsilon_{2,+} \subseteq \{(t,x) : \|x\| \leq t\}$. We may thus apply a simple uniqueness result.

Lemma 7.4. *If E and F are fundamental solutions with* $\operatorname{supp} E,\ \operatorname{supp} F \subseteq \{(t,x) : t \geq 0\}$ *then $E = F$.*

Lemma 7.5. *If $n = 2$ the the forward fundamental solution $E_{2,+}$ is given by*

$$\langle E_{2,+}, f\rangle = \frac{1}{2\pi}\int_0^\infty \int_{\|x\|\leq t} \frac{f(t,x)}{(t^2 - \|x\|^2)^{1/2}}\,dx\,dt.$$

When we throw a pebble into a pond we do not get a single ripple but a train of ripples!

It may be shown (see [1]) that in all dimensions the forward fundamental solution has support in the 'light cone' $\{(t, x) : \|x\| \leq t\}$, and that in odd dimensions $n > 1$ the support lies on the surface $\{(t, x) : \|x\| \leq t\}$ (but if $n \geq 5$ the form of $E_{n,+}$ is more complicated than that of $E_{3,+}$, involving analogues of δ_0', δ_0'' and so on). Notice the contrast with 'elliptic regularity' as described in Theorem Weyl. If n is even the support of $E_{n,+}$ is spread out over the whole light cone.

If $n = 1$ the matter follows an idea of D'Alembert familiar from 1B mathematical methods.

Exercise 7.6. *Let $n = 1$.*

(i) Show that if we make the change of coordinates $w = t + x$, $y = t - x$ the wave operator becomes $2\dfrac{\partial^2}{\partial w \partial y}$.

(ii) Show that the fundamental solution with support in $\{(w, y) : w \geq 0, y \geq 0\}$ is $\dfrac{1}{2}H(w)H(y)$.

A little thought shows that we have in fact a very general solution to our forcing problem.

Lemma 7.7. *Let A be a real number. If T is a distribution with $\operatorname{supp} T \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \geq A\}$ then there exists a unique distribution S with $\operatorname{supp} S \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \geq A\}$ and*

$$\Box S = T.$$

We have $S = E_+ * T$. If T is a k times differentiable function then so is S.

Reversing the sign of t we obtain E_- the backwards fundamental solution. We have the analogue of Lemma 7.7.

Lemma 7.8. *Let A be a real number. If T is a distribution with $\operatorname{supp} T \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \leq A\}$ then there exists a unique distribution S with $suppS \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \leq A\}$ and*

$$\Box S = T.$$

We have $S = E_- * T$. If T is a k times differentiable function then so is S.

If $T \in \mathcal{D}'$ we can write $T = T_1 + T_2$ with

$$\operatorname{supp} T_1 \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \geq A\} \text{ and } \operatorname{supp} T_2 \subseteq \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \leq B\}$$

for some A and B.

Lemma 7.9. *If T is a distribution then we can find a distribution S with*

$$\Box S = T.$$

If T is a k times differentiable (respectively smooth) function then we can choose S to be k times differentiable (respectively smooth).

It is pretty obvious that the decomposition $T = T_1 + T_2$ and the resulting solution S in Lemma 7.9 are not unique. A more direct way of seeing this is to observe that if u is a twice differentiable function on $\mathbb{R}$ and $\omega \in \mathbb{R}^n$ is a vector of unit length then $\Box u(t - x.\omega) = 0$.

Exercise 7.10. *(i) Verify this.*

(ii) Why should you have guessed this from physical or other considerations.

(iii) Why does the result not contradict Lemma 7.7.

The nature of the lack of uniqueness can be resolved if we can prove the following theorem.

Theorem 7.11. *Suppose that $u_0,\ u_1 : \mathbb{R}^n \to \mathbb{R}$ are smooth. Then if f is a continuous function there is a unique solution of*

$$\Box u = f$$

with

$$u(0,x) = u_0(x),\ \frac{\partial u}{\partial t}(0,x) = u_1(x).$$

Clearly it is enough to prove a simpler version.

Theorem 7.12. *Suppose that $u_0,\ u_1 : \mathbb{R}^n \to \mathbb{R}$ are smooth. Then there is a unique solution of*

$$\Box u = 0$$

with

$$u(0,x) = u_0(x),\ \frac{\partial u}{\partial t}(0,x) = u_1(x).$$

It should be noted that if

$$\Box u = 0$$

with

$$u(0,x) = u_0(x),\ \frac{\partial u}{\partial t}(0,x) = u_1(x),$$

then by the definition of $\Box$

$$\frac{\partial^2 u}{\partial t^2} = \Delta u$$

so in particular

$$\frac{\partial^2 u}{\partial^2 t}(0,x) = \Delta u_0(x).$$

Again

$$\frac{\partial^3 u}{\partial t^3} = \frac{\partial u}{\partial t}\Delta u$$

so

$$\frac{\partial^3 u}{\partial^3 t}(0,x) = \Delta u_1(x)$$

and so on. Thus $\dfrac{\partial^j u}{\partial^j t}(0,x)$ is specified to all orders. We shall need a converse observation.

Lemma 7.13. *If $u_0,\ u_1 : \mathbb{R}^n \to \mathbb{R}$ are smooth we can find smooth $u_j : \mathbb{R}^n \to \mathbb{R}$ $[j \geq 2]$ such that if $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ is smooth and*

$$\frac{\partial^j u}{\partial^j t}(0,x) = u_j(x)$$

for all $j \geq 0$ then

$$\left(\frac{\partial^k u}{\partial^k t}\Box u\right)(0,x) = 0$$

for all x.

We also need a result of Borel.

Theorem 7.14 (Theorem 6.1, Borel's lemma). *Given smooth $u_j : \mathbb{R}^n \to \mathbb{R}$ we can find a smooth $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ such that*

$$\frac{\partial^j u}{\partial^j t}(0,x) = u_j(x)$$

for all x and all $j \geq 0$.

Borel's lemma is plausible but not trivial. (The obvious Taylor series argument fails because Taylor series need not converge except at the origin.)

The proof of Theorem 7.12 is now easy. By Theorem 7.14 and Lemma 7.13 we can find a smooth function $v : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ such that

$$\left(\frac{\partial^k v}{\partial^k t} \Box u\right)(0, x) = 0$$

and

$$v(0, x) = u_0(x),\ \frac{\partial V}{\partial t}(0, x) = u_1(x).$$

If we set $f = \Box u$ then by our choice of u all the partial derivatives of f with respect to t vanish when $t = 0$ and the functions $f_1,\ f_2 : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$f_1(t, x) = H(t)f(t, x),\ f_2(t, x) = (1 - H(t))F(t, x)$$

are smooth. Setting

$$u = v - E_+ * f_1 - E_- * f_2$$

we have a solution of the required form. (Since $E_+ * f_1$ is infinitely differentiable with support in $t \geq 0$ all its derivatives must vanish at $t = 0$.)

To prove uniqueness, observe that by linearity we need only prove that if $\Box u = 0$ and u and its first partial derivative in t vanish at $t = 0$ then $u = 0$. But if $\Box u = 0$ and u and its first partial derivative in t vanish at $t = 0$ then all the partial derivatives of u with respect to t vanish when $t = 0$. We can thus write

$$u = u_+ + u_0$$

with $u_\pm$ smooth, supported in $\pm t \geq 0$ and satisfying $\Box u_\pm = 0$. By Lemma 7.7 $u_+ = 0$. Similarly $u_- = 0$ and we are done.

8 The heat equation

8.1 Existence

Recall that the heat operator J is given by

$$Ju(t, x) = \left(\frac{\partial}{\partial t} - \Delta\right) u(t, x)$$

where $u : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$.

Lemma 8.1. *A fundamental solution of the heat equation is given by*

$$\begin{aligned} K(t,x) &= \left(\frac{1}{4\pi t}\right)^{n/2} \exp(-\frac{\|x\|^2}{4t}) && \textit{if } t > 0, \\ K(t,x) &= 0 && \textit{if } t > 0. \end{aligned}$$

Note that we do not find a 'backward fundamental solution' – the heat equation has an arrow of time built in.

Since the only singularity of our fundamental solution is at $(0,0)$ we can build a parametrix and apply the argument of Theorem 4.18 to obtain a corresponding result.

Lemma 8.2 (Proposition 7.1). *If $u \in \mathcal{D}'$ and J is the heat operator then u and Ju have the same singular support. In particular if Ju is smooth so is u.*

This property is sometimes called *hypo-ellipticity* — saying that although the operator is not elliptic it behaves like an elliptic operator it behaves similarly in this respect.

We prove the next result by direct verification.

Theorem 8.3. *If $u_0 : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$ are continuous functions then the system*

$$\left(\frac{\partial}{\partial t} - \Delta\right) u(t,x) = g(t,x) \textit{ for } t > 0$$

subject to

$$\lim_{t\to 0+} u(t,x) = u_0(x)$$

has a solution

$$u(t,x) = K_t * u_0 + \int_0^t K_{t-s} * g_s \, ds$$

where $K_t(x) = K(t,x)$ and $g_s(x) = g(s,x)$.

If $g = 0$ we observe that $u(t,x)$ is a smooth function of x for fixed $t > 0$ whatever the choice of u_0 this is another indication of the arrow of time.

8.2 Uniqueness

We did not claim uniqueness in Theorem 8.3 because the infinite propagation speed of heat exhibited by our solution allows other (non-physical) solutions.

Exercise 8.4. *(This is included for interest, it is non-examinable.) We work with $n = 1$ for simplicity.*

(i) Show formally (without worrying about rigour) that if g is smooth

$$u(t, x) = \sum_{k=0}^{\infty} \frac{g^k(t) x^{2k}}{(2k)!}$$

satisfies $Ju = 0$.

(ii) If $g^k(0) = 0$ for all k but g is not identically zero then u is a solution of $Ju = 0$ with $u(t, x) \to 0$ as $t \to 0+$ but u is not the zero function.

(iii) We now need to choose g and make the argument rigorous. It is up to the reader to choose whether to try this. My own preference would be to manufacture g by hand but $g(t) = \exp(-1/t^2)$ $[t \neq 0]$, $g(0) = 0$ will do (see [3], Section 7.1).

Because parabolic equations lie 'on the boundary of elliptic equations' we try and use ideas about maxima of solutions similar to those of Theorems 5.11 and 5.12.

Lemma 8.5. *Consider the bounded open set*

$$\Omega = \{(t, x) : \|x - y\| < r,\ 0 < t < T\}.$$

If $u : \overline{\Omega} \to \mathbb{R}$ is continuous and u is twice differentiable with $Ju \leq 0$ on Ω then the maximum value of u occurs on

$$\{(t, x) : \|x - y\| \leq r,\ t = 0\} \cup \{(t, x) : \|x - y\| = r,\ 0 \leq t \leq T\}.$$

Theorem 8.6. *Consider the open set*

$$\Omega = \{(t, x) : 0 < t < T\} = (0, T) \times \mathbb{R}^n.$$

If $u : \overline{\Omega} \to \mathbb{R}$ is continuous and u is twice differentiable with $Ju \leq 0$ on Ω and further there exists an M such that

$$u(t, x) \leq M \text{ for all } (t, x) \in \overline{\Omega}$$

then

$$u(t, x) \leq \sup_{y \in \mathbb{R}^n} u(0, y)$$

for all $(t, x) \in \Omega$.

We can now provide a uniqueness result to go with the existence result of Theorem 8.3.

Theorem 8.7. *If $u_0 : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ are bounded continuous functions then the system*

$$\left(\frac{\partial}{\partial t} - \Delta\right) u(t,x) = g(t,x) \text{ for } t > 0$$

subject to

$$\lim_{t \to 0+} u(t,x) = u_0(x)$$

has exactly one solution.

We also have a result on continuous dependence on initial data.

Exercise 8.8. *Suppose $u_1,\ u_2 : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ are bounded continuous functions. If $\tilde{u}_j$ is the bounded solution of the system*

$$\left(\frac{\partial}{\partial t} - \Delta\right) \tilde{u}_j(t,x) = g(t,x) \text{ for } t > 0$$

subject to

$$\lim_{t \to 0+} \tilde{u}_j(t,x) = u_j(x)$$

for $j = 1,\ 2$. Then

$$\sup_{t \geq 0,\ x \in \mathbb{R}^n} \|\tilde{u}_2(t,x) - \tilde{u}_1(t,x)\| = \sup_{x \in \mathbb{R}^n} \|u_2(x) - u_1(t,x)\|.$$

9 References

The reader will be in doubt of my admiration for Friedlander's little book [2]. The additions to the second edition do not concern this course but some niggling misprints have been removed and you should tell your college library to buy the second edition even it already has the first. Friedlander's book mainly concerns the 'distributional' side of the course. Specific partial differential equations are dealt with along with much else in Folland's *Introduction to Partial Differential Equations* [1]. The two books [2] and [1] are also excellent further reading. The book [3] is a classic by a major worker in the field.

References

[1] G. B. Folland *Introduction to Partial Differential Equations* 2nd Edition, Princeton University Press, 1995. QA324

[2] F. G. Friedlander and M. Joshi *Introduction to the Theory of Distributions* 2nd Edition, CUP, 1998. QA324

[3] F. John *Partial Differential Equations* 4th Edition, Springer-Verlag, 1982 QA1.A647