LECTURE NOTES FOR IIB PARTIAL DIFFERENTIAL EQUATIONS

M. S. JOSHI AND A. J. WASSERMANN

Contents

IMPORTANT NOTE

These notes are being used with the consent of the authors for the 1999 course. The e-mail of the lecturer for the course is **twk@dpmms** but anyone wishing to get in touch with the authors should contact Dr Wassermann.

1. Introduction

Partial differential equations (PDEs) play a central role in most branches of applied maths, theoretical physics as well as geometry, analysis, probability theory and topology. In fact most analysis was developed to solve the problems posed by differential equations. For example the theory of Hilbert spaces and the spectral theorem for self-adjoint operators was invented to solve the classical Sturm-Liouville problem(see Hilbert Space course). The purpose of this course is to give a taste of many of the techniques invented to solve and analyse PDEs.

A PDE of order k is roughly an expression

$$* \qquad F(x, u, \partial_1 u, \dots, \partial_n u, \dots, \partial_n^k u) = 0 \qquad (x \in \Omega \subset \mathbb{R}^n \text{ open})$$

relating $u(x)$, $(x \in \mathbb{R}^n)$ and its derivatives of order $\leq k$. We shall sometimes say that u is a *classical solution of* $(*)$, this is to distinguish from the notion of a *weak* or *distributional* solution which we shall introduce later.

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$ set $\partial^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} u$ where $|\alpha| = \sum \alpha_i$. (Sometimes we write $\frac{\partial^{|\alpha|}}{\partial x^\alpha} u$ for $\partial^\alpha u$)

A *linear* PDE is one that can be written in the form

$$Pu = f \tag{1.1}$$

$$P = \sum_{|\alpha| \leq k} a_\alpha(x) \partial^\alpha \tag{1.2}$$

where a_α and f are functions of x. The order of the equation is then k. Most of this course is devoted to studying the case where all the a_α are constant - this is called a constant coefficient linear PDE. We shall see that these can be pretty complicated but also that they are now well understood.

A *quasi-linear* equation is one where a_α and f may also depend not only on x but also $\partial^\alpha u$ for $|\alpha| < k$.

There are also non-linear equations — Navier Stokes, Euler's equations for geodesics, K-dV, minimal surface equations — which require special methods of solution beyond this course.

Hadamard introduced the concept of a well-posed problem into PDEs - this is a problem which satisfies three basic criteria - existence, uniqueness and continuous dependence of solution on data given. Note that in a problem coming from physics one would expect all these three to be satisfied - the first two because something will happen and only one thing will happen. The third because data can never be measured exactly and so if two solutions coming from data close together are not similar then one can not make a useful prediction.

So in PDEs the problems we study are variants of the following.

- Existence: Prove that there is u satisfying $(*)$ possibly with prescribed conditions in a neighbourhood of a point or in Ω. We also want a constructive proof - i.e. one that tells us what u is!
- Uniqueness: Prove that u is unique. If u is not unique, what boundary conditions can be imposed to make u unique?
- Continuity: How does u depend on the boundary conditions and on f in the linear case?
- Smoothness: How many times differentiable is u? Does u have points of singularity? How long is the solution valid for? Often in non-linear problems the solution will blow-up (ie tend to infinity) after some finite time.

1.1. **The Basic Constant Coefficient Linear PDEs.** The most important differential operator in mathematics is the Laplacian which is equal to

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$$

and acts on functions on $\mathbb{R}^n$. It can also be defined on manifolds and then becomes a variable coefficient operator. The other linear differential operators we will consider are

$$\begin{array}{ll} \frac{\partial}{\partial t} - \Delta & (\text{ heat operator }) \\ \Box = \frac{\partial^2}{\partial t^2} - \Delta & (\text{ wave operator}) \end{array}$$

More generally, we'll study PDEs with constant coefficients using the Fourier transform as a key tool.

1.2. **Symbols and Definitions.** Associated to any linear, partial differential operator is a polynomial called the *total symbol* obtained by replacing $\frac{\partial}{\partial x_j}$ by $i\xi_j$:

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

The important point is that if $\hat{u}$ is the Fourier transform of u then

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

Proposition 1. *If P is a differential operator of order k and Q is a differential operator of order l then*

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

where $\alpha! = \prod \alpha_i!$.

The principal symbol is the top order part of the symbol:

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

The principal symbol is particularly useful because

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

When the principal symbol is never zero or only ever vanishes to first order - the operator is said to be of *principal type* and the lower order terms have little effect on the qualitative behaviour of the PDE. We define the characteristic set to be the points where the principal symbol vanishes -

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

The operator P is *elliptic* (of order k) at x if only if $\sigma_k(P)(x,\xi) \neq 0$ for $\xi \neq 0$ and elliptic if this is true for all x.

A hypersurface S is said to be *characteristic* for P at x if the normal vector is a characteristic vector for P. S is called *non-characteristic* if it isn't characteristic at any point.

Example 1. *The operator $P = \frac{\partial}{\partial x} + i\frac{\partial}{\partial y}$, on $\mathbb{R}^2$ (or $\mathbb{C}$) has principal symbol*

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

and so is elliptic. The Laplacian has principal symbol $-\sum_{j=1}^n \xi_j^2$ and thus is elliptic.

However the wave and heat operators are not.

$$char(\square) = \{\tau^2 - \xi^2 = 0 \subset \mathbb{R}^{n+1}_{x,t} \times \mathbb{R}^{n+1}_{\xi,\tau}\}$$
$$char(\partial_t - \Delta) = \{\xi = 0 \subset \mathbb{R}^{n+1}_{x,t} \times \mathbb{R}^{n+1}_{\xi,\tau}\}$$

It is traditional to classify second order operator into three main types corresponding to the three main operators. This comes from the geometry of the level sets of the principal symbols. (one has to consider the principal symbol of the heat operator to be $\tau - \xi^2$, rather than $i\tau - \xi^2$ or the actual value ξ^2.)

Elliptic – Laplace
Parabolic – Heat
Hyperbolic – Wave

These of course correspond to the non-degenerate conic sections - this insight is not however particularly useful.

1.3. **Examples of Non-linear PDEs.** These will not be treated in this course but you should be aware they are out there.

(1) **The Navier-Stokes Equation in Fluid Dynamics**

$$\frac{\partial u}{\partial t} - \triangle u + u\cdot \nabla u = f - \nabla p, \qquad \nabla\cdot u = 0$$

$f(t,x)$, $u(t,x) \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}$ (see Constantine and Foias).

(2) **The Euler Equations for Geodesics in an "Ideal Fluid"** Typically (Yang-Mills Instantons) differential equations arise through a minimisation or variational problem. Let $G = GL(n,\mathbb{R}) \subset M_n(\mathbb{R})$ be the group of invertible matrices and let $\langle a,b\rangle$ be an inner product on $M_n(\mathbb{R})$. If $g : [0,1] \to G$ is a differentiable path (C^1), its **length** $e(g) = \int_0^1 \|g'(t)\| dt$ where $g'(t) = g^{-1}\dot{g}$ and its **energy** $E(g) = \int_0^1 \|g'(t)\|^2 dt$. A curve g is called a **geodesic** if it is a critical point for the energy subject to the end points being fixed. It turns out that a C^2 curve is a geodesic if and only if g satisfies the **Euler equation**

$$\langle \frac{dg'}{dt}, X\rangle = \langle g', [g', X]\rangle \qquad (X \in M_n(\mathbb{R})).$$

Locally geodesics also minimise length. (see Arnold).

(3) **Beltrami's Equation and conformal structures** (linear PDE) Let

$$ds^2 = Edx^2 + 2Fdxdy + Gdy^2$$

be a Riemannian metric on a piece of $\mathbb{R}^2$ (or a 2-D manifold). Setting $z = x + iy$, we can write

$$ds^2 = \lambda|dz + \mu d\overline{z}|^2$$

where $\lambda > 0$ and $|\mu| < 1$ is complex. Coordinates (u, v) are called isothermal for ds^2 if $ds^2 = \rho(du^2 + dv^2)$ with $\rho > 0$. Let $w = u + iv$; then

$$\rho|dw|^2 = \rho|w_z|^2|dz + \frac{w_{\overline{z}}}{w_z}d\overline{z}|$$

so $\frac{dw}{d\overline{z}} = \mu\frac{dw}{dz}$. This is **Beltrami's equation** – it can always be solved. The complex structure w is called the **conformal structure** induced by the the metrics ds^2 (see Ahlfors)

(4) **The Korteweg-de Vries equation and solutions** $u_t - 6uu_x + u_{xxx} = 0$ was first discovered in connection with solitary water waves. It is an example of a **completely integrable system** and its solutions form an infinite dimensional subspace (indexed by an infinite-dimensional Grassmannian, ie special subspaces of an infinite-dimensional Hilbert space). (see Drazin, Kac)

(5) **Minimal Surface** $z = u(x, y)$ satisfies $(1 + u_y^2)u_{xx} - 2u_xu_yu_{xy} + (1 + u_x^2)u_{yy} = 0$.

2. Existence and Uniqueness of Solutions for ODEs

2.1. The Method of Successive Approximations.

Theorem 1. *(The Contraction Mapping Theorem) Let (X, d) be a complete, non-empty, metric space and $T : X \to X$ a map such that $d(Ty_1, Ty_2) \leq kd(y_1, y_2)$ with $k \in (0, 1)$. Then T has a unique fixed point in X; in fact if $y_0 \in X$, then $T^m y_0 \to$ fixed point as $m \to \infty$.*

Proof. Using the geometric progression

$$\frac{1}{1-k} = \sum_{m\geq 0} k^m,$$

we check that $T^m y_0$ forms a Cauchy sequence in X. So by completeness of X, $T^m y_0 \to y$ some y. But then $T^{m+1}y_0 \to Ty$, so $Ty = y$ and y is a fixed point. Clearly

$$Ty_i = y_i \ (i = 1, 2) \ \Rightarrow d(y_1, y_2) \leq kd(y_1, y_2)$$

so $d(y_1, y_2) = 0$ and $y_1 = y_2$, Thus T has a unique fixed point. □

Corollary 1. *Suppose that T^n is a contraction mapping for some n. Then the same conclusions hold.*

Proof. By Theorem 1, T^n has a unique fixed point, y. We also have that

$$T^n(Ty) = T^{n+1}y = T(T^n y) = Ty.$$

So Ty is also a fixed point of T^n and fixed points are unique so $Ty = y$. Also

$$T^{mn}y_0 \to y\ T^{mn+1}y_0 \to y,\ \ldots,\ T^{mn+(n-1)}y_0 \to y\ (m \to \infty).$$

Putting these together, we see $T^m y_0 \to y$. □

Note that this result not only says there is a fixed point but also gives a method for finding it.

Let $f(t, x)$ be a vector-valued continuous function $|t - t_0| \leq a$, $\|x - x_0\| \leq b$ where $x \in \mathbb{R}^n$. Suppose f also satisfies the Lipschitz condition

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

Note that a Lipschitz function is automatically continuous. It follows from the Mean Value Theorem that a differentiable is at least locally Lipschitz. Let $M = \sup |f(t, x)|$ and set $h = \min(a, \frac{b}{M})$.

Theorem 2. *The differential equation*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0 \tag{2.1}$$

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

Proof. (Picard-Lindelöf)

We will prove this by using the contraction mapping theorem - to do so we need a suitable metric space and a contraction mapping on it. As differential operators make things less smooth and integral operators make things more smooth, we work with an equivalent problem defined in terms of integrals.

Let

$$(Tx)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds. \tag{2.2}$$

Clearly x solves (2.1) if only if $Tx = x$. (just integrate or differentiate).

Now let

$$X = \{x \in C([t_0 - h, t_0 + h], \mathbb{R}^n) \mid \|x(t) - x_0\| \leq M \cdot h\ \forall t\}.$$

This is a complete metric space for

$$d(x_1, x_2) = \sup_{|t - t_0| \leq h} \|x_1(t) - x_2(t)\|,$$

(this follows from the fact that a uniform limit of continuous functions is continuous.) As $Mh \leq b\, Tx$ is well-defined $x \in X$, and Tx is also in X. We claim

$$\|T^k x_1(t) - T^k x_2(t)\| \leq \frac{c^k}{k!}|t-t_0|^k \,\mathrm{d}(x_1, x_2).$$

For $k = 0$, this is obvious, and in general it follows by induction since

$$\begin{aligned} \|T^k x_1(t) - T^k x_2(t)\| &\leq \int_{t_0}^t \|f(s, T^{k-1}x_1(s)) - f(s, T^{k-1}x_2(s))\| ds \\ &\leq c\int_{t_0}^t \|T^{k-1}x_1(s) - T^{k-1}x_2(s)\| ds \\ &\leq \tfrac{c^k}{(k-1)!}\int_{t_0}^t |s-t_0|^{k-1}\, ds\ d(x_1,x_2) \\ &\leq \tfrac{c^k}{k!}|t-t_0|^k d(x_1,x_2). \end{aligned}$$

But then T^n is a contraction mapping for n sufficiently big and the result follows □

NB why, in the above proof, is the solution x is differentiable?

2.2. **Dependence of ODEs on Initial Conditions.**

Theorem 3. *The solution of* (2.1) *depends continuously on the initial data* x_0.

Proof. The idea in this proof is to solve for all possible initial data simultaneously, thus obtaining a continuous function both of t and the initial data.

Pick $h_1 < h$ and take $\delta > 0$ such that $Mh_1 + \delta \leq b$.

Let

$Y = \{y \in C([t_0 - h_1, t_0 + h_1] \times \overline{B}(x_0, \delta), \mathbb{R}^n) : \|y(t,x) - x\| \leq M \cdot h, y(t_0, x) = x\}$.

Again Y is complete for the supremum metric

$$(y_1, y_2) = \sup \|y_1(t,x) - y_2(t,x)\|.$$

Let

$$(Ty)(t,x) = x + \int_{t_0}^t f(s, y(s,x))ds$$

Since $Mh_1 + \delta \leq b$, T maps Y into Y and as before we can check by induction that

$$\|T^k y_1(t,x) - T^k y_2(t,x)\| \leq \frac{c^k}{k}|t-t_0|^k \,\mathrm{d}(y_1, y_2).$$

So T^n is a contraction mapping for n sufficiently large and T has a unique fixed point y which satisfies

$$\begin{aligned} \tfrac{\partial y}{\partial t} &= f(t,y) \\ y(t_0, x) &= x \end{aligned}$$

Now y is a continuous function of both t and x and if we fix $x = x_0$ then $y(t, x_0)$ solves the initial value problem (2.1) so the solution of (2.1) depends continuously on x_0 which is what we wanted to prove.

Note this works for any $h_1 < h$ so we have continuity everywhere on the open interval. □

2.3. **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 $f(t, x)$

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

so to each time t and point x we have associated a vector. We can take f to be continuous, C^k (continuous derivatives of all orders $\leq k$) or smooth (derivatives of all orders.)

Let $x_0 \in U$. An *integral curve* for f with *initial condition* (or starting point) x_0 is a map

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

such that

$$\frac{\mathrm{d}\,\varphi}{dt} = f(t, \varphi(t)),$$

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

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

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

where $x \in U_0$ open $\subseteq U$, such that

$$\begin{cases} \frac{\mathrm{d}\,\alpha(t,x)}{dt} &= f(t, \alpha(t, x)) \\ \alpha(0, x) &= x \end{cases} .$$

Thus $\varphi_x(t) = \alpha(t, x)$ is an integral curve for f with initial condition x.

2.4. **Time-independent Vector Fields.** Suppose f does not depend on t, it is just a map $f : U \to \mathbb{R}^n$ assigning a vector to each point of U.

Let $\alpha(t, x) = \alpha_t(x)$ be the flow determined by f. It exists for t small enough and is as smooth as f is.(see below) The chain rule shows that

$$t \mapsto \alpha(t, \alpha(t_0, x)), \tag{2.3}$$

$$t \mapsto \alpha(t + t_0, x) \tag{2.4}$$

are integral curves of f with the same initial condition $\alpha_{t_0}(x)$ at $t=0$. But integral curves are unique so they must coincide. Hence $\alpha_{t+t_0} = \alpha_t \circ \alpha_{t_0}$, thus

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

whenever this make sense. This means that we have a local bit of an action of $\mathbb{R}$ on U which may only be partially defined. It is called a *local* 1*-parameter group or dynamical system.*

2.5. **Perturbations of Linear ODEs.**

Theorem 4. *Let $A(t,x)$, $B(t,x)$ be continuous matrix-valued functions of t and x and let*

$$M \geq \sup_{t,x} \|B\|.$$

The solutions of the ODEs

$$\begin{array}{rclcrcll} \frac{\mathrm{d}\,\xi(t,x)}{dt} & = & A(t,x)\xi(t,x), & & \xi(t_0,x) & = & a(x) & (\textit{vector-valued}) \\ \frac{\mathrm{d}\,\eta(t,x)}{dt} & = & B(t,x)\eta(t,x), & & \eta(t_0,x) & = & b(x) & \end{array}$$

satisfy $\sup_x \|\xi(t,x) - \eta(t,x)\| \leq C\|A-B\|\frac{e^{M|t-t_0|}-1}{M} + \|a-b\|e^{M|t-t_0|}$ *where C is a constant depending only on A and a.*

Proof. By the method of successive approximations, we know that the sequences defined by

$$\begin{array}{rclr} \xi_k & = & a + \int_{t_0}^t A\xi_{k-1}ds, & \xi_0 = a \\ \eta_k & = & b + \int_{t_0}^t B\eta_{k-1}ds, & \eta_0 = b \end{array}$$

will be such that $\xi_n \to \xi$, $\eta_k \to \eta$. Let $g_k(t) = \sup_x \|\xi_k(t,x) - \eta_k(t,x)\|$ and $C = \sup_{k,x,t} \|\xi_k\|$. Note that

$$\|\xi_k\|$$

is bounded as ξ_k is convergent which implies that $\|\xi_k\|$ is also.

Then we check that

$$g_n(t) \leq \|a-b\| + C\|A-B\||(t-t_0)| + M\int_{t_0}^t g_{n-1}(s)ds \tag{2.5}$$

Define f_n by $f_0(t) = \|a-b\|$ and then inductively by defining

$$f_n(t) = \|a-b\| + C\|A-B\||(t-t_0)| + M\int_{t_0}^t f_{n-1}(s)ds. \tag{2.6}$$

Comparing (2.5) and (2.6), we see that

$$f_n \geq g_n.$$

As we have a contraction mapping, $f_n \to f$ where f is a solution of

$$f(t) = \|a-b\| + C\|A-B\|(t-t_0) + M\int_{t_0}^{t} f(s)ds$$

Solving the corresponding differential equation we get

$$f(t) = \|a-b\|e^{M|t-t_0|} + C\|A-B\|\frac{e^{M|t-t_0|}-1}{M}$$

As $g_n(t) \leq f_n(t)$,

$$\sup_x \|\xi_n(t,x) - \eta_n(t,x)\| \leq f_n(t).$$

The theorem now follows by passing to the limit as $n \to \infty$. □

2.6. **Smoothness Properties of Flows.** The smoother the vector field f is, the smoother we would expect the associated flow α to be.

Theorem 5. *If f is C^k and*

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

then α is also $C^k, 1 \leq k \leq \infty$.

Proof. This proof is not examinable. The hardest case is $k = 1$, the others follow almost trivially by induction.

So we assume f is C^1, ie $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$ exist and are continuous. We must show that α is also C^1. Note that formally, if we set $\lambda(t,x) = \left(\frac{\partial\alpha(t,x)}{\partial x_i}\right) = D_x\alpha$ (an $n \times n$ matrix). We expect λ to satisfy the linear ODE

$$\frac{d\lambda}{dt} = D_x f(t,\alpha)\lambda \qquad (1)$$

Let λ be the continuous solution of (1). We show $D_\alpha a$ exists and equals λ. Let $F(s) = f(t, a+s(b-a))$. Then

$$\frac{dF}{ds} = D_x f(t, a+s(b-a)) \cdot (b-a)$$

so $f(t,b) - f(t,a) = \int_0^1 D_x f(t, a+s(b-a)) \cdot (b-a)ds$. But then

$$\begin{aligned} & \tfrac{d}{dt}(\alpha(t,x+y) - \alpha(t,x)) \\ = \;& f(t,\alpha(t,x+y)) - f(t,\alpha(t,x)) \\ = \;& \textstyle\int_0^1 D_x f(t,\alpha(t,x)) + s(\alpha(t,x+y) - \alpha(t,x)) \cdot (\alpha(t,x+y) - \alpha(t,x))ds \end{aligned}$$

Let $A(t,x) = D_x f(t,\alpha(t,x)), \xi(t,x) = \lambda(t,x)y$. $B_y(t,x) = \int_0^1 D_x f(t,\alpha(t,x)+ s(\alpha(t,x+y)-\alpha(t,x)))ds$, $\eta_y(t,x) = \alpha(t,x+y)-\alpha(t,x)$. The perturbation theorem for ODEs applies ($\frac{\partial \xi}{\partial t} = A\xi$, $\frac{\partial \eta}{\partial t} = B\eta$) and implies that

$$\sup_{|t|\leq \epsilon} \|\lambda(t,x)y - \{\alpha(t,x+y)-\alpha(x)\}\| = o(\|y\|)$$

So $D_x\alpha = \lambda$; since $\frac{d\alpha}{dt} = f(t,\alpha)$, this means α is C^1. Now $\frac{\mathrm{d}\lambda}{dt} = A\lambda$. Suppose f is C^k and α is known to be C^p. Then A is C^p, so λ is C^p (by induction). So $D_x\alpha$ is C^p. Also $\frac{d\alpha}{dt} = f(t,\alpha)$ is C^p. Hence α is $C^{p+1} = C^k$. □

2.7. **Critical Points.** Let $f : U \to \mathbb{R}^n$ be a vector field. A *critical point* of f is a point x_0 such that $f(x_0) = 0$.

Observation (1) If φ is an integral curve of f passing through a critical point x_0, then φ is constant.

Proof. By uniqueness, since $\varphi(t) = x_0$ is an integral curve.

Observation (2) If $\lim_{t\to\infty} \varphi(t) = x$, then x is a critical point of f.

Proof. By definition

$$\varphi(t_1) - \varphi(t_0) = \int_{t_0}^{t_1} f(\varphi(s))ds.$$

Write

$$f(\varphi(s)) = f(x_1) + g(s).$$

Then $g(s) \to 0$ as($s \to \infty$.) So estimating the integral, we have

$$\|f(x_1)\||t_1 - t_0| \leq \|\varphi(t_1) - \varphi(t_0)\| + |t_1 - t_0| \sup_{t_1\geq s\geq t_0} |g(s)|.$$

Set $t_0 = R$, $t_1 = 2R$ and let $R \to \infty$. We get $f(x) = 0$.

2.8. **First Order Semi-Linear PDEs.** Consider the semi-linear first order PDE

$$Lu \equiv \sum_{j=1}^{n} a_j(x)\frac{\partial u}{\partial x_j} = f(x,u) \tag{2.7}$$

where a_j, b are real C^1 functions of $x \in \mathbb{R}^n$ and f is C^1 but possibly complex-valued. We want to solve this with the value of u on some hypersurface, S, given - this is called a *Cauchy problem.* Let $A(x)$ be the vector field $(a_1(x),\ldots,a_n(x))$. Let γ be an integral curve of A then

$$\frac{d}{dt}(u(\gamma(t)) = \sum_j a_j(\gamma(t))\frac{\partial u}{\partial x_j}(\gamma(t)). \tag{2.8}$$

So differentiating along integral curves of A is equivalent to applying the operator $\sum_j a_j(x)\frac{\partial}{\partial x_j}$. This means that solving (2.7) is equivalent to solving the ODE

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

along each integral curve γ. So a method of solution is now clear. We need to specify data on a hypersurface intersecting each integral curve once and then solve along each integral curve.

We therefore assume the data is given on a non-characteristic hypersurface, S, this means that the normal to S, call it ξ, does not satisfy

$$\sum a_j \xi_j = 0$$

that is that A is never tangent to S.

Theorem 6. *Locally, there is a unique solution of* (2.7) *which takes given data on a non-characteristic hypersurface S.*

We shall not complete the details of the proof of this theorem, as we can deduce it (when f is real) from a more general theorem about solutions of quasi-linear equations which is proved using a similar technique.

Example 2. *Solve the PDE*

$$\frac{\partial u}{\partial x} + 2x\frac{\partial u}{\partial y} = u^2, \quad u(0, y) = f(y).$$

First we find the integral curves of the vector field $(1, 2x)$. *So we have*

$$\frac{dx}{dt} = 1, \frac{dy}{dt} = 2x.$$

We solve this to obtain

$$x(t) = t + c_1, y(t) = t^2 + 2c_1 t + c_2.$$

We want the integral curve γ_{y_0} to start at $(0, y_0)$. So then $c_1 = 0$ and $c_2 = y_0$. We now have to solve

$$\frac{d}{dt}(u \circ \gamma_{y_0})(t) = (u \circ \gamma_{y_0})(t)^2,\ (u \circ \gamma_{y_0})(0) = f(y_0).$$

This has the solution,

$$(u \circ \gamma_{y_0})(t) = -\frac{1}{t - f(y_0)^{-1}}.$$

If $f(y_0) = 0$ the solution is identically zero. We express u as a function of (x, y) instead of (y_0, t). We have $x = t, y = x^2 + y_0$ so

$$u(x, y) = -\frac{f(y - x^2)}{xf(y - x^2) - 1}.$$

(Note this will solve regardless of whether $f(y - x^2)$ is zero.)

Note that the correspondence given here between first order partial differential operators and vector fields is quite an important fact and indeed in the study of differential geometry it is customary to identify the two.

2.9. **First Order Quasi-Linear PDEs.** These are a bit more general than semi-linear equations as we allow the coefficients of the derivatives to vary with the solution. If S is a hypersurface we study the problem

$$\sum_{j=1}^{n} a_j(x, u)\frac{\partial u}{\partial x_j} = b(x, u), \ u_{|S} = \phi \tag{2.9}$$

where all the functions are real. The solution technique for this relies on regarding u as a variable on the same basis as x. Suppose S is parametrised by a function g that is

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

We work with the vector field

$$(a_1, a_2, \dots, a_n, b) \text{ on } \mathbb{R}^{n+1}$$

and solve for the integral curves

$$\frac{dx}{dt} = a(x, y) \tag{2.10}$$

$$\frac{dy}{dt} = b(x, y) \tag{2.11}$$

$$x(0) = g(s) \tag{2.12}$$

$$y(0) = \phi(s). \tag{2.13}$$

Our solution is then 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 invertible derivative at $t = 0$ provided the vector

$$(a_1(g(s), \phi(s)), \dots, a_n(g(s), \phi(s)))$$

is not tangent to S - so we assume this is true - this is our non-characteristic condition. It then follows from the inverse function theorem that the map has an inverse locally. We then define

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

and this is our solution. It clearly satisfies the initial conditions on S we need only check that the differential equation is satisfied. We compute

$$\begin{aligned}
\sum a_j \frac{\partial u}{\partial x_j} &= \sum a_j \Big(\sum \frac{\partial s_k}{\partial x_j}\frac{\partial u}{\partial s_k} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial x_j}\Big) \\
&= \sum \frac{\partial u}{\partial s_k} \sum a_j \frac{\partial s_k}{\partial x_j} + \frac{\partial u}{\partial t}\sum a_j \frac{\partial t}{\partial x_j} \\
&= \sum \frac{\partial u}{\partial s_k} \sum \frac{\partial s_k}{\partial x_j}\frac{\partial x_j}{\partial t} + \frac{\partial u}{\partial t}\sum \frac{\partial t}{\partial x_j}\frac{\partial x_j}{\partial t} \\
&= \sum \frac{\partial u}{\partial s_k}\frac{\partial s_k}{\partial t} + \frac{\partial u}{\partial t}\frac{\partial t}{\partial t} \\
&= 0 + b(x, u).
\end{aligned}$$

So we have proven the existence half of

Theorem 7. *The equation* (2.9) *has a unique solution near S provided the vector field $(a_1(x(s), \phi(s)), a_2(x(s), \phi(s)), \ldots, a_n(x(s), \phi(s)))$ is not tangent to S anywhere.*

The uniqueness comes from the fact that the system used in the existence part has a unique solution and that any u solving the equation will give a solution of the system.

Note if we have a semi-linear equation with real coefficients we can use the technique for quasi-linear equations if we wish but if it has a complex right hand side, we must use the first technique.

The proof was also a technique so we can use it to find the solutions.

Example 3. *Solve $u\frac{\partial du}{\partial dx} + \frac{\partial du}{\partial dy} = 1$ with $u = s/2$ on $x = y = s$. .*

The characteristic condition is satisfied provided $s \neq 2$. We first find integral curves

$$\frac{dx}{dt} = u, \ \frac{dy}{dt} = 1, \ \frac{du}{dt} = 1,$$

with initial data $(x(s,0), y(s,0), u(s,0)) = (s, s, s/2)$. This has solution

$$u = t + s/2, \ y = t + s, \ x = t^2/2 + st/2 + s.$$

After eliminating s and t, we obtain

$$u = \frac{4y - 2x - y^2}{2(2-y)}.$$

Note the singularity is precisely at $y = 2$ where the non-characteristic condition fails.

The rest of section two is optional and not examinable.

2.10. **Formal Power Series Solutions of Holomorphic ODEs (optional).** Consider the complex ODE

$$\frac{dX}{dz} = A(z)X(z), \quad X(0) = X_0 \tag{2.14}$$

where $A(z) = \sum_{m\geq 0} A_m z^m$ is a holomorphic matrix valued function defined for $|z| < r$. We look for a formal power series solution of (2.14) of the form

$$X(z) = \sum_{m\geq 0} X_m z^m.$$

We get the recurrence relation

$$mX_m = \sum_{i+j=m-1} A_i X_j \tag{2.15}$$

which can we use to compute all the X_m. Our problem is to show that the series for $X(z)$ will converge. ¿From (2.15), we have the inequality

$$m\|X_m\| \leq \sum_{i+j=m-1} \|A_i\|\|X_j\|. \tag{2.16}$$

Let $a_i = \|A_i\|$ and define x_m by $x_0 = \|X_0\|$,

$$mx_m = \sum_{i+j=m-1} a_i x_j. \tag{2.17}$$

Clearly $\|X_i\| \leq x_i$. so the radius of convergence of $X(z)$ will be less than that of

$$x(z) = \sum x_i z^i.$$

Now if we set $a(z) = \sum a_i z^i$, then from (2.17), $x(z)$ is a formal solution of

$$\frac{dx}{dz} = a(z)x(z)$$

and $x(0) = x_0$. By construction $a(z)$ is holomorphic for $|z| < r$ and this scalar ODE can be explicitly integrated:

$$x(z) = x_0 \exp \int_0^z a(w)dw.$$

Thus $x(z)$ is analytic for $|z| < r$ (since there is a unique formal solution by (2.17).) Since $\|X_i\| \leq x_i$, $X(z)$ is also analytic for $|z| < r$, as required.

So to summarise

Theorem 8. *If $A(z)$ is a holomorphic matrix-valued function defined in $|z| < r$ then there is a unique holomorphic function $X(z)$ defined in $|z| < r$ solving*

$$\frac{dX}{dz} = A(z)X(z),\ X(0) = X_0$$

with X_0 given.

Hans Lewy's Counter example.

If we try to solve $Lf = g$ when L is a complex vector field, we can split up into real and imaginary parts, giving Z simultaneous real equations of the same form. The previous geometric arguments do not therefore apply and there may be no solutions.

Theorem. *Let $\Omega = \{(x,y,z) : x^2 + y^2 < R, |z| < R\}$ and let $f(z)$ be continuous and real valued. If there is a C^1 function u on Ω satisfying*

$$Lu \equiv \frac{\partial u}{\partial x} + i\frac{\partial u}{\partial y} - z_i(x+iy)\frac{\partial u}{\partial z} = f(z)$$

then f must be real analytic on $|z| < R$.

Proof. Let $v(r,\theta,z) = e^{i\theta}\sqrt{r}u(\sqrt{r}\cos\theta, \sqrt{r}\sin\theta, z)$, a C' function on $0 < r < R$, $\theta \in [0, 2\pi]$, $|z| < R$ with period 2π in θ.

By change of variables

$$Lu = z\frac{\partial v}{\partial r} + \frac{i}{r}\frac{\partial v}{\partial \theta} - z_i\frac{\partial v}{\partial z}(\text{ Check! })$$

Set $V(r,z) = i\int_0^{2\pi} N(r,\theta,z)d\theta$, a C' function on $0 < r < R$, $|z| < R$. Now

$$\begin{aligned} \tfrac{\partial v}{\partial z} + i\tfrac{\partial v}{\partial r} &= i\int_0^{2\pi}(\tfrac{\partial v}{\partial z} - \tfrac{1}{2r}\tfrac{\partial v}{\partial \theta} + i\tfrac{\partial v}{\partial r})d\theta \\ &= i\int_0^{2\pi}\tfrac{i}{2}f(z)d\theta \\ &= -\pi f(z) \end{aligned}$$

Let $F(z) = \int_0^z f(s)ds$ and set $W(z,r) = V(r,z) + \pi f(z)$. Let $\sigma = z + ir$, a complex variable. So $\frac{\partial w}{\partial \xi} = 0$ and hence W is homomorphic on $|z| < R$, $o < r < R$. Moreover W extends to a continuous function on $|z| < R$, $0 < r < R$ with $W(z,0) = \pi F(z)$ real valued (note that $V = 0$ if $r = 0$). By Schwartz's reflection principle, the extension $W(\xi) = \overline{W(\xi)}$ makes W holomorphic on $|z| < R$, $|r| \subset R$. So W is real analytic on $r = 0$. Hence $F(z)$ and $f(z)$ are real analytic.

Thus if f is continuous but not real analytic, there is no solution.

The Cauchy-Kowalewski theorem for linear PDEs

We start by showing that the Cauchy problem always has a solution for 1st order linear PDEs with analytic coefficients. Then we reduce the higher order case to this one. We use the method of majoring power series.

Theorem *The Cauchy problem*

$$\frac{\partial Y}{\partial t} = \sum_{i=1}^{n} A^i(x,t)\frac{\partial Y}{\partial x_i} + B(x,t) \qquad (*)$$
$$Y(x,0) = 0,$$

where Y *and* B *are vector-valued functions and* $A_1, \dots, A_n$ *are matrix-valued with* A_i, B *analytic near* $(0,0)$, *has a unique analytic solution in a neighbourhood of* $(0,0)$.

Proof. Suppose $A^i(x,t) = \sum_{m\geq 0} A^i_m(x)t^m$, $B(x,t) = \sum_{m\geq 0} B_m(x)t^m$ where $A^i_m(x)$ and $B_m(x)$ are power series in x. Let a^i, b, a^i_m, b_m be the power series obtained by replacing all coefficients (matrices or vectors) by their norms.

We look for a formal solution of $(*)$, $Y(x,t) = \sum_{m\geq 1} Y_m(x)t^m$ starting from $m = 1$ to satisfy the b.c.

This gives the recurrence relation for $m \geq 1$

$$mY_m = \sum_i \sum_{p+q=m-1} A^i_p(x)\frac{\partial Y_q}{\partial x_i} + B_{m-1}(x)$$

which uniquely determines $Y(x,t)$. Let $y(x,t) = \sum_{m\geq 1} y_m(x)t^m$ be the solution of

$$my_m = \sum_i \sum_{p+q=m-1} a^i_p(x)\frac{\partial y_q}{\partial x_i} + b_{m-1}(x) \qquad (m \geq 1)$$

Then clearly each coefficient of y dominates the norm of the con.coefficient of Y. So it suffices to show that $y(x,t)$ is analytic at $(0,0)$, ie the formal power series converges. But $y(x,t)$ is a solution of

$$\frac{\partial u}{\partial t} = \sum a^i(x,t)\frac{\partial y}{\partial x_i} + b(x,t), \qquad y(x,0) = 0$$

Note that if $\sum a_\alpha x^\alpha$ is convergent for $|x_i| \leq r$ then $a_\alpha r^{|\alpha|} \to 0$ as $|\alpha| \to \infty$. Hence $\|a_\alpha\| \leq Kr^{-|\alpha|}$ for some $k > 0$ and $\sum a_\alpha x^\alpha$ is majorised coefficient by coefficient
$K\sum(\frac{x}{r})^\alpha = K\prod(1-\frac{x_i}{r})^{-1}$. Since the coefficients in $\prod(1-x_i)^{-1}$ are

all one while those in $(1-\sum x_i)^{-1}$ are greater than 1, we see that $\prod(1-\frac{x_i}{r})^{-1}$ is majorised by $(1-\frac{\sum x_i}{r})^{-1}$.

Since a^i, b are analytic at $(0,0)$ they are majorised by $C\prod(1-\frac{x_i}{r})^{-1}(1-\frac{t}{r})^{-1}$ and hence $C=(1-\frac{t}{r})^{-1}(1-\frac{\sum x_i}{r})^{-1}$ for some C, $r>0$. But then y will be majorised by the solution of

$$(\neq)\ \frac{\partial z}{\partial t}=C(1-\frac{\sum x_i}{r})^{-1}(1-\frac{t}{t})^{-1}(\sum\frac{\partial z}{\partial x_i}+1),\qquad z(x,0)=0.$$

But if w is the solution of $\frac{\partial w}{\partial t}=C(1-\frac{t}{r})^{-1}(1-\frac{s}{r})^{-1}(\frac{\partial w}{\partial s}+1)$,

$w(s,0)=0$ then the solution of $(\neq)$ is $z(x,t)=w(\sum x_i,t)$. But the solution of $(**)$ is given by

$$w(s,t)=r-s-\sqrt{(r-s)^2+zr^2\log(1-t/r)}\qquad\text{(check!)}$$

which is analytic at $(0,0)$ for s, t sufficiently small. Hence z,y and Y are analytic at $(0,0)$.

Theorem *The Cauchy problem*

$$(1)\quad\left[\begin{array}{l}\frac{\partial^k Y}{\partial t^k}=\sum\limits_{|\alpha|+j\le k}A_{\alpha,j}(x,t)\frac{\partial^{|\alpha|}}{\partial x^\alpha}\frac{\partial^j}{\partial t^j}Y+B(x,t)\\ \frac{\partial^j Y}{\partial t^j}(x,0)=\phi_j(x)\quad(j=0,\dots,k-1)\quad\textit{(Cauchy data)}\end{array}\right.$$

where Y and B are vector valued functions, the $A_{\alpha,j}$ are matrix valued and B, $A_{\alpha,j}$ are analytic near $(0,0)$, has a unique analytic solution in a neighbourhood of $(0,0)$.

Proof. We reduce the problem to the 1st order case by introducing derivatives as new variables.

Set $y_{\alpha j}=\frac{\partial^{|\alpha|}}{\partial x^\alpha}\frac{\partial^j}{\partial t^j}Y$ for $j<k$, $|\alpha|+j\le k$. Then the equations become

$$(2)\quad\left[\begin{array}{l}\frac{\partial}{\partial t}y_{\alpha,j}=y_{\alpha,j+1}\text{ for }|\alpha|+j\le k\\ \frac{\partial}{\partial t}y_{\alpha,j}=\frac{\partial}{\partial x_i}y_{\beta,j+1}\text{ if }|\alpha|+j=k,j<k\text{ where }\beta_p=\alpha_p\text{ except }\beta_i=\alpha_i-1\\ \frac{\partial}{\partial t}y_{o,k}=\frac{\partial}{\partial t}\sum\limits_{\substack{j+|\alpha\le k\\ j<k}}A_{\alpha,j}(x,t)\frac{\partial^{|\alpha|}}{\partial x^k}\frac{\partial^j}{\partial t^j}Y+\frac{\partial B}{\partial t}\end{array}\right.$$

with initial conditions

$$\left[\begin{array}{lcl}y_{\alpha j}(x,0)&=&\frac{\partial^{|\alpha|}}{\partial x^\alpha}\phi_j(x)\\ y_{0,k}(x,0)&=&\sum A_{\alpha,j}\frac{\partial^{|\alpha|}}{\partial x^\alpha}\frac{\partial^j}{\partial t^j}Y(x,0)+B(x,0)\\ &=&\sum A_{\alpha,j}y_{\alpha j}(x,0)+B(x,0)\end{array}\right.$$

A moments reflection shows that system of equations (2) is equivalent to the system (1). So the problem is reduced to one of the form

$$\frac{\partial Y}{\partial t} = \sum A^j(x,t)\frac{\partial Y}{\partial x_j} + B(x,t), \qquad Y(x,0) = \phi(x).$$

Setting $Y(x,t) - \phi(x)$ in place of Y we reduce to the case $\phi = 0$, for which we have just proved the result.

Remark

(1) The C-K theorem is also true for quasi-linear equations, when the A^j and B's depend also analytically on Y. The proof is essentially the same but more complicated to write down.

(2) Let L be an mth order differential operator with analytic coefficients and S an analytic hyper-surface non-characteristic for L. Then the Cauchy problem

$$\begin{aligned} Lu &= f \\ D^\alpha u &= \phi \text{ on } s(|\alpha| < m) \end{aligned}$$

has a unique analytic solution locally near S for any f, ϕ analytic. In fact we make an analytic change of coordinate so that S is given by $t = 0$ in $\mathbb{R}^{n+1} = \{(x,t)\}$. Since S is non-characteristic the coefficient of $(\frac{\partial}{\partial t})^m$ must be invertible, so we are in the situation of the C-K theorem.

3. The Fourier Transform and PDEs with Constant Coefficients

3.1. The Fourier Transform on Schwartz Functions. Let

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) \mid \sup_x |x^\beta \partial^\alpha f(x)| < \infty, \forall \alpha, \beta\}$$

this is called the space of *Schwartz functions.* It is easy to see that $p(x)e^{-\alpha\|x\|^2}$ lies in $\mathcal{S}(\mathbb{R}^n)$ for any polynomial p and $\alpha > 0$.

We will be interested in maps on $\mathcal{S}(\mathbb{R}^n)$ - for example the Fourier transform - and we therefore want a notion of continuity. We therefore define,

$$\|f\|_{\alpha,\beta} = \sup_x |x^\beta \partial^\alpha f(x)|,$$

for $f \in \mathcal{S}(\mathbb{R}^n)$. A linear map T on $\mathcal{S}$ is then continuous if for all α, β there exists $k_{\alpha,\beta}$ such that

$$\|Tf\|_{\alpha,\beta}\| \leq C_{\alpha,\beta} \sum_{|\gamma|,|\delta|} \leq k_{\alpha,\beta}\|f\|_{\gamma,\delta}.$$

Roughly this says we can control the size and decay of the derivatives of Tf by those of f.

Any smooth function of compact support lies in $\mathcal{S}(\mathbb{R}^n)$ - it is not obvious that such functions exist. These can be constructed in the following way.

Lemma 1. *(Bump functions) There is a smooth function, f, on $\mathbb{R}$ such that*
$f(t) = 1$ *for* $|t| \leq 1$, $f(t) = 0$ *for* $|t| \geq 1 + \delta$ *and* $0 \leq f(t) \leq 1$ *all* t.

Proof. Let

$$g(x) = \begin{cases} \exp\left(\frac{-1}{1-x^2}\right) & |x| < 1 \\ 0 & |x| \geq 1. \end{cases}$$

We have a constructed a smooth function of compact support. Now let

$$h(x) = \int_{-\infty}^x g(t)dt \Big/ \int g.$$

Then $h(x) = 0$ for $x \leq -1$ and $h(x) = 1$ for $x \geq 1$. Moreover $0 \leq h(x) \leq 1$ for all x. Taking

$$k(x) = h(\alpha x + \beta)$$

for suitable α and β, we get a function such that $k(x) = 0$ for $x \leq -1 - \delta$, $k(x) = 1$ for $x \geq -1$ and $0 \leq k(x) \leq 1$. Now set

$$f(x) = k(x)k(-x).$$

□

For $f \in L^1(\mathbb{R}^n)$, (see appendix) define the *Fourier transform* $\hat{f}$ by

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int e^{-ix\cdot\xi} f(x)dx \tag{3.1}$$

Clearly $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, since

$$(1 + \|x\|^2)^{-n} \in L^1(\mathbb{R})$$

for example or

$$\prod (1 + x_i^2)^{-1} \in L^1(\mathbb{R}^n).$$

We set $D_j = -i\frac{\partial}{\partial x_j}$ - this turns out to be very useful in studying the Fourier transform. Clearly

$$D_j : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \tag{3.2}$$

$$x_j : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n) \tag{3.3}$$

Lemma 2. *The Fourier transform $f \mapsto \hat{f}$ maps $\mathcal{S}$ into $\mathcal{S}$ and*

$$\widehat{D_j f} = \xi_j \hat{f},$$
$$\widehat{x_j f} = -D_j \hat{f}$$

Proof. If we differentiate (3.1), we get

$$D^\alpha \hat{f}(\xi) = \int e^{-ix\cdot\xi}(-x)^\alpha f(x)dx$$

which is valid since $x^\alpha f(x)$ is integrable. So $\hat{f}(\xi)$ is smooth and $D^\alpha \hat{f} = \widehat{(-x)^\alpha f}$. Integrating,

$$\int e^{-ix.\xi} \xi_j f(x)dx$$

by parts we obtain the first statement.

So we get

$$\xi^\beta D^\alpha \hat{f}(\xi) = \int e^{-ix\cdot\xi} D^\beta((-x)^\alpha f(x))dx = \int e^{-ix\cdot\xi} g(x)dx$$

with g Schwartz.

Hence

$$\sup |\xi^\beta D^\alpha \hat{f}(\xi)| \leq C \sup \prod (1 + |x_i|^2)|g(x)|$$

where $C = \int \frac{1}{\prod(1+x_i^2)} dx < \infty$. So $f \mapsto \hat{f}$ takes $\mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}(\mathbb{R}^n)$. □

These facts are useful as they show that the Fourier transform converts constant coefficient linear operators into multiplication by a polynomial and it is also the key to one method of inverting the Fourier transform. Before proving the Fourier inversion theorem we introduce a couple of lemmas.

Lemma 3. *If* $f \in \mathcal{S}(\mathbb{R}^n)$ *and* $f(0) = 0$ *then* $f(x) = \sum_{i=1}^{n} f_i(x)x_i$ *with* $f_i \in \mathcal{S}(\mathbb{R}^n)$.

Proof. An n dimensional version of Taylor's theorem says that

$$f(x) = \sum F_i(x)x_i \tag{3.4}$$

with F_i smooth but not necessarily Schwartz. (Iterate the result of Theorem 28 (ii).) On the other hand for $x \neq 0$

$$f(x) = \sum G_i(x)x_i \tag{3.5}$$

where $G_i(x) = f(x)x_i/\|x_i\|^2$. The functions G decay correctly but need not be smooth at $x = 0$.

We construct our function by taking a mixture of these two; let ψ be a bump function equal to 1 for t small and 0 for $t \geq 1$ and set

$$f_i(x) = \psi(\|x\|^2)F_i(x) + (1 - \psi(\|x\|^2))G_i(x).$$

Both these summands lie in $\mathcal{S}(\mathbb{R}^n)$. As $\psi + (1 - \psi) = 1$, the result follows. □

Corollary 2. *If* $f \in \mathcal{S}(\mathbb{R}^n)$ *and* $f(a) = 0$ *then* $f(x) = \sum(x_i - a_i)f_i(x)$ *with* $f_i \in \mathcal{S}(\mathbb{R}^n)$.

Proof. By the lemma $f(x + a) = \sum x_i g_i(x)$ with $g_i \in \mathcal{S}$. So

$$f(x) = \sum(x_i - a_i)g_i(x - a)$$

where $f_i(x) = g_i(x - a) \in \mathcal{S}$. □

Lemma 4. *Let* $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ *be a linear map commuting with* x_j *and* D_j *for all* j. *Then*

$$Tf = cf$$

for some $c \in \mathbb{C}$.

Proof. If $f(a) = 0$ then from the corollary above we have that

$$Tf(x) = \sum(x_i - a_i)Tf_i$$

so $Tf(a) = 0$. i

So observing that $f(x) - f(a)e^{-\|x-a\|^2}$ is zero when $x = a$, we have that $T(f - f(a)e^{-\|x-a\|^2})(a) = 0$ and hence

$$T(f)(a) = f(a)T(e^{-|x-a|^2})(a) = c(a)f(a)$$

for some function c which is independent of f.

Now take some particular $g \in \mathcal{S}(\mathbb{R}^n)$ with $g > 0$ (for example $g(x) = \exp(-x^2)$). We observe that

$$c = Tg/g$$

is smooth. But then

$$cD_j fTD_j g = D_j Tg = D_j(cg) = (D_j c)g + c(D_j g).$$

Hence $(D_j c)g = 0$, so $D_j c \equiv 0$ and hence c is a constant function. □

Theorem 9. *The Fourier transform*

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

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

$$f(x) = \left(\frac{1}{2\pi}\right)^{n/2} \int e^{ix\cdot\xi} \hat{f}(\xi) d\xi.$$

Proof. Let $F(f) = \hat{f}$. Then F^2 is a linear map on $\mathcal{S}$ and

$$F^2 x_j = -x_j F^2, \; F^2 D_j = -D_j F^2.$$

Now let $Rf(x) = f(-x)$. Then $Rx_j = -x_j R$, $RD_j = -D_j R$ and so $T = RF^2$ commutes with x_j and D_j. So, applying the lemma, $T = RF^2 = c$ for some constant c.

We are thus reduced the problem to computing the constant c. Let

$$f_0(x) = e^{-\|x\|^2/2}.$$

Then

$$(x_j + iD_j)f_0 \equiv 0.$$

Hence

$$(-D_j + i\xi_j)\hat{f}_0 \equiv 0.$$

So

$$\hat{f}_0(\xi) = c_1 e^{-\|\xi\|^2/2}$$

for some c_1 by the uniqueness of solutions to ODEs. Setting $\xi = 0$, we get

$$c_1 = \frac{1}{(2\pi)^{n/2}} \int e^{-\|x\|^2/2} dx = 1$$

since $\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$. So $F^2 f_0 = f_0$, so $Tf_0 = f_0$ and hence $c = 1$. □

3.2. **Properties of the Fourier Transform.** For $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$ define the convolution $\phi * \psi$ of ϕ and ψ by

$$\phi * \psi(x) = \frac{1}{(2\pi)^{n/2}} \int \phi(x-y)\psi(y)dy.$$

Clearly $\phi * \psi = \psi * \phi$.

Theorem 10. *For* $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$

(*a*) $\int \hat{\varphi}\psi = \int \varphi\hat{\psi}$
(*b*) $\int \varphi\overline{\psi} = \int \hat{\varphi}\overline{\hat{\psi}}$ *(Parseval's formula)*
(*c*) $\widehat{\varphi * \psi} = \hat{\varphi}\hat{\psi}$
(*d*) $\widehat{\varphi\psi} = \hat{\varphi} * \hat{\psi}$.

Proof. (*a*) Both sides are given by $\frac{1}{(2\pi)^{n/2}} \iint \varphi(x)\psi(\xi)e^{-ix\cdot\xi}dxd\xi$.

(*b*) Set $\chi = \overline{\hat{\psi}}$. Then $\overline{\hat{\chi}(\xi)} = (2\pi)^{-n/2}\int \hat{\psi}(x)e^{ix\cdot\xi}dx = \psi(\xi)$ so result follows from (a), replacing ψ by χ.

(*c*) Both sides are given by $\frac{1}{(2\pi)^n} \iint \varphi(x)\psi(y)e^{-i(x+y)\cdot\xi}dxdy$.

(*d*) The Fourier transform of $\widehat{\varphi\psi}$ is $\varphi(-x)\psi(-x)$ while

$$\hat{\varphi} * \hat{\psi} = \hat{\hat{\varphi}}\hat{\hat{\psi}} = \varphi(-x)\psi(-x).$$

So the result follows from (c).

□

3.3. **The Paley-Wiener Theorem (optional).** In general, the Fourier transform exchanges growth at infinity with smoothness properties. The Paley-Wiener says that if a function is of compact support then its Fourier transform is analytic and vice-versa. One can provide more general statements that relate the boundedness of the support in certain directions with analyticity in certain sectors. The idea is to realise that f extends to the whole of $\mathbb{C}^n$ in this case, the same idea will be used to prove that any non-zero differential operator $p(D)$ with constant coefficients has a fundamental solution.

Theorem 11. *If* $f \in C_0^\infty(\mathbb{R}^n)$ *has support in* $B(0,r) = \{x : \|x\| \leq r\}$ *and if*

$$\hat{f}(z) = \frac{1}{(2\pi)^{n/2}} \int f(x)e^{-iz\cdot x}dx \qquad (z \in \mathbb{C}^n) \tag{1}$$

then $\hat{f}$ *is entire and there are constants* C_N *s.t.*

$$|\hat{f}(z)| \leq C_N(1+|z|)^{-N}e^{r|\mathrm{Im}(z)|} \qquad (N = 0,1,2,\dots) \tag{2}$$

Conversely every entire function satisfying (2) is the Fourier-Laplace transform of a smooth function in $C_0^\infty(\mathbb{R}^n)$.

Proof. We start by recalling that $g(z)$ is *entire* if $g : \mathbb{C}^n \to \mathbb{C}$ is continuous and separately holomorphic in each coordinate.

If g vanishes on $\mathbb{R}^n$ then g is identically zero; to see this let $a_1, \dots, a_n$ be real variables. Then $g(a_1, \dots, a_n) = 0\ \forall a_i$. So $g(a_1, \dots, a_{n-1}, z_n) \equiv 0$ by the one dimensional result. Continuing in this way we get $g \equiv 0$.

Note that if $\|x\| \le r$, then $|e^{-iz\cdot x}| \le e^{r\,\mathrm{Im}(z)}$. So $\hat{f}$ exists and is continuous in z, since the integration need only be performed over $B(0, r)$. By Morera's theorem applied to each coordinate of z separately, $\hat{f}$ is holomorphic so entire. Moreover

$$z^\alpha \hat{f}(z) = \frac{1}{(2\pi)^{n/a}} \int (D^\alpha f)(x) e^{-ix\cdot z} dx$$

Hence $|z^\alpha||\hat{f}(z)| \le \|D^\alpha f\|_1 e^{r|\,\mathrm{Im}\,z|}$ which immediately gives (2).

Now suppose that $g(z)$ is an entire function on $\mathbb{C}^n$ satisfying (2) for all N. Set

$$f(x) = \frac{1}{(2\pi)^{n/2}} \int g(\xi) e^{ix\cdot\xi} d\xi$$

Since $(1 + |\xi|)^N g(\xi)$ is in L^1 for all N, f is C^∞ with

$$D^\alpha f(x) = \frac{1}{(2\pi)^{n/2}} \int \xi^\alpha g(\xi) e^{ix\cdot\xi} d\xi$$

Next we claim that for any $\eta \in \mathbb{R}^n$

$$f(x) = \frac{1}{(2\pi)^{n/2}} \int g(\xi + i\eta) e^{ix\cdot(\xi+i\eta)} d\xi \qquad (3)$$

It clearly suffices to check this when $\eta = (0, \dots, 0, \lambda, 0, \dots, 0)$ in which case it follows from Cauchy's theorem in one variable by taking a rectangular contour and noting that $|g| \to 0$ on the vertical part of the contour. Set $\eta = \alpha x$ with $\alpha > 0$ for $x \ne 0$. Then

$$|g(\xi + i\eta) e^{ix\cdot(\xi+i\eta)}| \le C_N (1 + |\xi|)^{-N} e^{\alpha\|x\|(r-\|x\|)}$$

and hence

$$|f(x)| \le \frac{C_N}{(2\pi)^{n/2}} e^{\alpha\|x\|(r-\|x\|)} \int (1 + |\xi|)^{-N} d\xi \qquad (4).$$

where N is chosen large enough for the latter integral to converge. Now suppose $\|x\| > r$. Let $\alpha \to \infty$ in (4); we get $f(x) = 0$.

(1) follows now from the inversion formula since $g(z)$ and $\hat{f}$ agree on $\mathbb{R}^n$ and hence $\mathbb{C}^n$, since they're both entire. □

Corollary 3. *If $f \in C_0^\infty(\mathbb{R}^n)$ then $P(D)u = f$ has a solution u in $C_0^\infty(\mathbb{R}^n)$ if and only if $\hat{f}(\xi)/P(\xi)$ is entire. The solution is then uniquely determined and if f is supported in $B(0,r)$ then u is also.*

Proof. Taking Fourier-Laplace transforms gives $P(z)\hat{u}(z) = \hat{f}(z)$ so $\hat{f}(z)/P(z)$ is the entire function $\hat{u}(z)$. To prove the converse we need the following 1 variable result.

Lemma 5. *Let $h(z)$ be holomorphic and $p(z)$ a polynomial with leading coefficient a ($z \in \mathbb{C}$). Then*

$$|ah(0)| \leq \max_{|z|=1} |h(z)p(z)|.$$

Proof. Set $q(z) = z^m \overline{p}\frac{(1}{z)}$ where $m = \deg p$. Then $q(0) = \overline{a}$ and $|ah(0)| = |q(0)h(0)| \leq \max_{|z|\leq 1} |q(z)h(z)| = \max_{|z|=1} |p(z)h(z)|$. □

Lemma 6. *Let $p(z) = \sum_{|\alpha|\leq m} a_\alpha z^\alpha$ be a polynomial of degree m. Then there is a real orthogonal change of coordinates such that the coefficient of z_1^m is non-zero.*

Proof. We may assume $P(z) = \sum_{|\alpha|=m} a_\alpha z^\alpha$ is homogeneous of degree m and proceed by induction on the number of coordinates. Write $p(z) = \sum_{a\leq k} z_1^a p_a(z_2, \ldots, z_n)$. Make an orthogonal change of coordinates in $z_2, \ldots, z_n$ s.t. the coefficients b_0 of z_2^{m-a} in p_k is non-zero. Now replace z_1 by $cz_1 - sz_2$ and z_2 by $sz_1 + cz_2$ where $c = \cos\theta$, $s = \sin\theta$. The coefficient of z_1^m is then clearly

$$b_0 s^{m-a} c^a + b_1 c^{a-1} s^{m-a+1} + \cdots$$

with $b_i \in \mathbb{C}$, $b_0 \neq 0$. This is essentially a polynomial in $\frac{s}{c}$ so is non-zero for all but finitely many values of θ as required. □

Remark 1. *This also follows from the fact that O_n acts irreducibly on homogeneous polynomials of degree m.*

End of Proof of Corollary After an orthogonal change of coordinates, we may assume that the coefficient of z_1^m in P is non-zero. Suppose $g(z) = \hat{f}(z)/p(z)$ is entire. Note that the coordinate change doesn't affect the bounds on f since $|z|$ and $|\operatorname{Im} z|$ are not altered. Set $p(\xi) =$

$p(\xi + z_1, \dots, z_n)$ and $h(z) = g(\xi + z_1, \dots, z_n)$. By Lemma 1 $|ah(0)| \leq \sup_{|\xi|=1} |h(\xi)p(\xi)|$, so

$$\begin{aligned} |g(z)| &\leq \tfrac{1}{|a|} \sup_{|\xi|=1} |\hat{f}(\xi + z_1, z_2, \dots, z_n)| \\ &\leq \tfrac{CN}{|a|} \sup_{|\xi|=1} (1 + |z + (\xi, 0, \dots, 0)|)^{-N} e^{r|\operatorname{Im}(z+\xi)|} \end{aligned}$$

Now $|z| \leq 1 + |z + \xi|$ since $|\xi| = 1$ so $(1 + |z|) \leq z(1 + |z + \xi|)$. Hence

$$|g(z)| \leq \frac{C_N}{|a|} 2^{-N} (1 + |z|)^{-N} e^{r|\operatorname{Im} z|} e^r$$

So g is the Fourier transform of a C^∞ function supported in $\|x\| \leq r$. □

3.4. **Smooth Partitions of Unity (optional).** Partitions of unity are an important idea - they allow us to *localise.* We divide a compact set up into lots of small balls where we can prove our result and then glue together with a partition of unity. It is important to realize this is not possible in the analytic theory where local and global are irrevocably intertwined.

Theorem 12. *(Smooth Partitions of Unity) Let K be a compact subset of $\mathbb{R}^n$ and $U_1, \dots, U_n$ open sets in $\mathbb{R}^n$ such that $K \subset \cup U_i$. Then we can find $f_i \in C_0^\infty(\mathbb{R}^n)$ with $0 \leq f_i(x) \leq 1$, supp $f_i \subseteq U_i$ with $\sum f_i(x) = 1$ on K and $\sum f_i(x) \leq 1$ all x.*

Proof. Recall that on any open ball in $\mathbb{R}^n$, we can find a bump function $g \in C_0^\infty(\mathbb{R}^n)$ with $0 \leq g \leq 1$ and $g > 0$precisely on the given open open ball. We shall use these functions in stead of the distance functions.

Since K is compact, we can cover K by open balls $B_1, \dots, B_l$ with each $\overline{B_j}$ contained wholly in some U_i. Then $\cup B_j$ is a bounded set in $\mathbb{R}^n$ so contained in some closed ball $\overline{B}(0, R)$. Then $\overline{B}(0, R) \backslash \cup B_j$ is compact and disjoint from K so can be covered by finitely many open balls $C_1, \dots, C_q$ all disjoint from K.

For each ball B_i pick a bump function g_i and each ball C_j pick a bump function h_j. Finally for the ball $\|x\| > R$ "at ∞" pick a C^∞ function $k \geq 0$ with supp $k =$ this ball, e.g. $k(x) = \psi(\frac{x}{\|x\|^2} R)$ where ψ is a bump function for $\|x\| \leq 1$.

Thus $\sum h_i + \sum g_j + k > 0$ on $\mathbb{R}^n$. Now set $F_i = g_i / \sum h_i + \sum g_i + k$. So supp $F_j = B_j$ and $\sum F_j(x) \leq 1$ on $\mathbb{R}^n$ with equality on K (since

$h_i \circ k = 0$ on K). Finally match up the B_j's with U_i's in which they wholly lie and set $f_i(x) = \sum_{B_j \subset U_j} F_j(x)$. □

3.5. **The Schur Test(optional).** If K is a locally integrable function on $\mathbb{R}^n \times \mathbb{R}^n$ then one can define

$$Pu(x) = \int K(x,y)u(y)dy$$

this will be a map from smooth functions of compact support to smooth functions (not necessarily of compact support.) We call K the Schwarz kernel of P (it has nothing to do with the null space.) There are important connections between the properties of K and P.

Let $K(x,y)$ be a continuous function in $\mathbb{R}^n \times \mathbb{R}^n$ such that

$$\sup_y \int |K(x,y)|dx \leq C, \quad \sup_x \int |K(x,y)|dy \leq C.$$

Then the integral operator defined by the kernel K has norm $\leq C$ in $L^2(\mathbb{R}^n)$, ie

$$(\int | \int K(x,y)f(y)dy|^2 dx)^{1/2} \leq C(\int |f(x)|^2 dx)^{1/2}.$$

Proof. By the Cauchy-Schwartz inequality

$$\begin{aligned} |\tilde{K}f(x)|^2 &= |\textstyle\int K(x,y)f(y)dy|^2 \\ &\leq \textstyle\int |K(x,y)||f(y)|^2 dy \int |K(x,y)|dy \\ &\leq C \textstyle\int |K(x,y)||f(y)|^2 dy \end{aligned}$$

Hence $\int |\tilde{K}f(x)|^2 dx \leq C \int\int |K(x,y)||f(y)|^2 dydx \leq C^2 \int |f(y)^2|dy$. □

Corollary 4. *If $f \in L^1$, $g \in L^2$ then $\|f * g\|_2 \leq \|f\|_1 \|g\|_2$ where $\|f\|_1 = \frac{1}{(2\pi)^{n/2}} \int |f(x)|dx$.*

Proof.

$$f * g(x) = \frac{1}{(2\pi)^{n/2}} \int f(x-y)g(y)dy$$

so $K(x,y) = \frac{f(x-y)}{(2\pi)^{n/2}}$ so the conditions of the Schur test are satisfied with $C = \|f\|$. □

Remark If $\|f\|_p$ is the norm $\|f\|_p = (\int |f(x)|^p dx)^{1/p}$ and $L^p(\mathbb{R}^n) = \{f \mid \|f\|_p \leq \infty\}$, then the Schur test is also true with L^2 replaced by L^p. The proof needs the Hölder inequality $\int |f(x)g(x)|dx \leq (\int |f(x)|^p dx)^{1/p} (\int |g(x)|^q dx)^{1/q}$ where $\frac{1}{p} + \frac{1}{q} = 1$. In particular this gives $\|f * g\|_p \leq \|f\|_1 \|g\|_p$ (see Ex sheet 2).

4. Definitions of Test Functions and Distributions

4.1. **Test Functions.** The space of test functions on $\mathbb{R}^n$ is the space $C_0^\infty(\mathbb{R}^n)$, ie smooth functions of compact support, with the definition that $f_n \to f$ as test functions if f_n, f are all supported in some fixed ball $\overline{B}(0,R)$ and

$$\sup |D^\alpha(f_n - f)| \to 0 \ (n \to \infty) \ \forall \alpha.$$

Example 4. *Let $f \in C_0^\infty(\mathbb{R}^n)$ be non-zero. Let $f_n = \frac{1/n}{f}$ then f_n converges to zero in $C_0^\infty(\mathbb{R}^n)$. But if we let $g_n(x) = f(x-n)$ then g_n does not converge to zero in $C_0^\infty(\mathbb{R}^n)$ although it converges to zero pointwise.*

Distributions are motivated by the fact we have, provided one of f and g has compact support, that

$$\int D^\alpha f(x) g(x) dx = (-1)^{|\alpha|} \int f(x) D^\alpha g(x) dx \qquad (*)$$

which we can write $\langle D^\alpha f, g\rangle = (-1)^{|\alpha|}\langle f, D^\alpha g\rangle$. This means that the derivative of a non-differentiable function can be defined in terms of how it pairs with a smooth function of compact support. With this in mind, we define a *distribution*, T (a generalised function) to be a linear map $f \mapsto \langle T, f\rangle$ from the space of test functions to $\mathbb{C}$ which satisfies the continuity condition

$$f_n \to f \Rightarrow \langle T, f_n\rangle \to \langle T, f\rangle.$$

It is traditional to denote this class $\mathcal{D}'(\mathbb{R}^n)$ or $C^{-\infty}(\mathbb{R}^n)$. The important thing to note it that any locally integrable function, u, will define a distribution by

$$< u, f >= \int u(x) f(x) dx \ f \in C_0^\infty.$$

Linearity is obvious but we need to check continuity. If $f_n \to f$ in C_c^∞ then there exists K compact such that $\mathrm{supp}(f_n)$ is contained in K for all n. So we have

$$|\langle u, f_n - f\rangle| \leq \int_K |u(x)||(f - f_n)(x)| dx$$

which will tend to zero as

$$\sup |f_n - f| \to 0.$$

Note that we will generally regard locally integrable functions as a subset of distributions. (a locally integrable function is a function which is in L^1 when multiplied by the characteristic function of any compact ball) The most important distribution that is not a function is the Dirac delta function

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

We define $D^\alpha T$ by

$$\langle D^\alpha T, f \rangle = (-1)^{|\alpha|} \langle T, D^\alpha f \rangle$$

so $\langle D^\alpha \delta_a, f \rangle = (-1)^{|\alpha|} D^\alpha f(a)$.

Example 5. *The Heaviside function $H(x)$ on $\mathbb{R}$ is zero on $x < 0$ and one otherwise. As it is locally integrable it defines a distribution. So DH is given by*

$$\langle DH, f \rangle = -\langle H, DF \rangle = -\int_0^\infty Df(x)dx$$

which by the fundamental theorem of calculus is just $\frac{1}{i} f(0)$. So the derivative of the Heaviside function is the delta function.

The main aim now is to extend as many operations as possible from functions to distributions - convolution, Fourier Transform, change of coordinates. We must be careful since, for example, it will not be possible to multiply distributions nor can one define the Fourier transform even for all smooth functions.

A *fundamental solution* of $p(D)$, where $p(\xi_1, \dots, \xi_n)$ is a polynomial and $p(D) = p(-i\frac{\partial}{\partial x_1}, \cdots, -i\frac{\partial}{\partial x_n})$, is a distributional solution, T, of

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

So the Heaviside function is a fundamental solution for $\frac{\partial}{\partial x}$ on $\mathbb{R}$.

General fact: every constant coefficient operator $P(D)$ has a fundamental solution. It is not necessarily unique as we can add any solution of $p(D)T = 0$. We'll find fundamental solutions for

$$\sum \frac{\partial^2}{\partial x_i^2} (\text{ Laplacian}), \quad \frac{\partial}{\partial t} - \sum \frac{\partial^2}{\partial x_i^2} (\text{ heat operator })$$

and

$$\frac{\partial^2}{\partial t^2} - \sum \frac{\partial^2}{\partial x_i^2}$$

(wave operator or D'Alembertian).

The reason we are interested in fundamental solutions is that we can use them to solve PDEs as they are effective inverses to differential operators. Then to solve $P(D)f = g$, we set $f = T * g$, then

$$\begin{aligned} P(D)(T * g) &= (P(D)T) * g \\ &= (\delta_0) * g = g. \end{aligned}$$

For the three operators above, T will be given very explicitly so this gives a complete solution to the problem. We do have to be careful though about when do the convolutions exist and we have to understand what convolution means for distributions.

4.2. **Linear Operations on Distributions.** Suppose $A : C_c^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^n)$ is a linear map which is continuous, i.e. $f_n \to f$ implies $A(f_n) \to A(f)$. The dual or transpose of A, if it exists is a map

$$A' : C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^n)$$

which is linear and continuous with

$$\langle Af, g\rangle = \langle f, A'g\rangle \qquad \forall f, g \in C_0^\infty(\mathbb{R}^n).$$

If the transpose exists then we can then extend A to distributions u by the formula

$$\langle Au, f\rangle = \langle u, A'f\rangle \qquad (f \in C_0^\infty(\mathbb{R}^n)).$$

It is important to realize that transposes do not always exists as $C_0^\infty(\mathbb{R}^n)$ is **not** a Hilbert space with respect to the pairing

$$\langle f, g\rangle = \int f(x)g(x)dx.$$

Examples

(*a*) *Multiplication* $Af = \psi f$ where $\psi \in C^\infty(\mathbb{R}^n)$. Clearly $A' = A$ so any distribution u can be multiplied by ψ

$$\langle \psi u, f\rangle = \langle u, \psi f\rangle.$$

NB multiplication by a distribution is another matter as it is not a map on $C_c^\infty(\mathbb{R}^n)$ and in general is not possible.

(*b*) *Differentiation* $Af = \partial^\alpha f$. Then $A'f = (-1)^{|\alpha|}\partial^\alpha f$ so we get

$$\langle \partial^\alpha u, f\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha f\rangle$$

(*c*) *Reflection* $Af = \tilde f$ where $\tilde f(x) = f(-x)$. Thus $A' = A$. So we set

$$\langle \tilde u, f\rangle = \langle u, \tilde f\rangle.$$

(*d*) *Translation* $T_a f = f(x-a)$ then $T_a' = T_{-a}'$. So we define

$$\langle T_a u, f\rangle = \langle u, T_{-a} f\rangle.$$

Abusing notation $T_a u$ is often written $u(x-a)$.

An important concept when dealing with distributions is that of support. This is the set of points where the distribution is not identically zero. As is usual for distributions, we first define the concept for functions and then use a backwards definition to do so for distributions.

So if $f \in C^\infty(\mathbb{R}^n)$, we define

$$\mathrm{supp}(f) = \overline{f^{-1}(\mathbb{C} - \{0\})}.$$

And if $u \in \mathcal{D}'(\mathbb{R}^n)$ then $x \notin \mathrm{supp}(f)$ if and only if there exists an open set U such that $x \in U$ and if $\mathrm{supp}(f) \subset U$ then $\langle u, f\rangle = 0$. It can be shown, using partitions of unity, that the definition implies the apparently stronger statement, if U is an open set with $U \cap \mathrm{supp}\, u = \emptyset$ and $\mathrm{supp}\, f \subset U$ then $\langle u, f\rangle = 0$.

For example if a function is supported in $\mathbb{R}^n - \{0\}$ then it will always pair with the delta function to give zero. We therefore conclude that

$$\mathrm{supp}(\delta_0) = \{0\}.$$

Clearly the only distribution whose support is empty is the zero function.

A related concept is that of singular support, this is the set of points where the distribution is not smooth. We say $x \notin \mathrm{singsupp}(u)$ if there exists a smooth function f such that $x \notin \mathrm{supp}(u-f)$. An equivalent definition is that there exists a smooth ψ such that $\psi(x) \neq 0$ and $\psi u \in C^\infty$. If the singular support is empty then the distribution is given by a smooth function.

However the operation we really want to extend to distributions, the Fourier transform, can not be extended to the full class as the space of compactly supported smooth functions is not invariant under it. Indeed, the only smooth function of compact support whose Fourier transform is compactly supported is the zero function - to see this observe that the Fourier transform will be analytic. We therefore need to work within a larger class of functions - the Schwartz functions.

4.3. **The Fourier Transform and Tempered Distributions.** We define the space of tempered distributions, $\mathcal{S}'(\mathbb{R}^n)$, to be the space of linear maps, u, from $\mathcal{S}(\mathbb{R}^n)$ to $\mathbb{C}$ such that if $f_n \to f$ in $\mathcal{S}(\mathbb{R}^n)$ then

$$\langle u, f_n\rangle \to \langle u, f\rangle.$$

An important but non-obvious fact is that the space of tempered distributions form a natural subspace of the space of distributions. To see this, one needs to check two things. The first is that it is continuous on $C_c^\infty(\mathbb{R}^n)$, to see this we need to show that if $f_n \to f$ in $C_c^\infty(\mathbb{R}^n)$ then it does in $\mathcal{S}(\mathbb{R}^n)$ - as we then know that

$$\langle u, f_n \rangle \to \langle u, f \rangle.$$

The second is that if a tempered distribution vanishes on all functions of compact support then it vanishes everywhere - this means that a tempered distribution is determined by its value on $C_c^\infty(\mathbb{R}^n)$. This will follow if given any Schwartz function f there exists a sequence of $f_n \in C_c^\infty(\mathbb{R}^n)$ converging to it in $\mathcal{S}(\mathbb{R}^n)$ (but not of course in $C_c^\infty(\mathbb{R}^n)$.)

First let's show a tempered distribution is continuous on $C_c^\infty(\mathbb{R}^n)$. If f_n converges to f in $C_c^\infty(\mathbb{R}^n)$ then $\operatorname{supp}(f_n)$ is contained in some fixed ball $B(0, R)$ for some R. So,

$$\|f_n - f\|_{\alpha,\beta} = \sup |x^\alpha D_\beta (f_n - f)(x)| \leq R^\alpha \sup |f_n - f|,$$

which converges to zero.

Theorem 13. *Given $f \in \mathcal{S}(\mathbb{R}^n)$ there exists a sequence of functions f_n in $C_c^\infty(\mathbb{R}^n)$ such that*

$$\sup |x^\alpha D^\beta (f_n - f)| \to 0, \ \forall \alpha, \beta$$

as $n \to \infty$.

Proof. Let $\psi \geq 0$ be a bump function with support in $\|x\| \leq 1$ with $\psi(x) = 1$, $\|x\| \leq \frac{1}{2}$. So D_ψ^α is bounded for each α. Let

$$\psi_R(x) = \psi\left(\frac{x}{R}\right)$$

for $R > 0$. Then $\sup |D^\alpha \psi_R| = \frac{1}{R^{|\alpha|}} \sup |D^\alpha \psi|$; so for $R \geq 1$ the $D^\alpha \psi_R$'s are uniformly bounded. We set

$$f_n = \psi_n f \in C_0^\infty(\mathbb{R}^n).$$

Then

$$\begin{aligned} \sup |x^\alpha D^\beta (f_n - f)| &= \sup |x^\alpha D^\beta (\psi_n - 1) f| \\ &= \sup_{\|x\| \geq \frac{n}{2}} |x^\alpha D^\beta (\psi_n - 1) f| \\ &\leq C \sum_{\alpha_1 \leq \alpha, \beta_1 \leq \beta} \sup_{\|x\| \geq n/2} |x^{\alpha_1} D^{\beta_1} f| \end{aligned}$$

But $x^{\alpha_1} D^{\beta_1} f$ is a Schwartz function and so is rapidly decaying at infinity. □

We can now define the Fourier transform on tempered distributions. We have seen that the Fourier transform is a bijection on $\mathcal{S}(\mathbb{R}^n)$. Moreover we showed that the estimate

$$\sup|\xi^\beta D^\alpha \hat{f}(\xi)| \leq C \sup \prod(1+|x_i|^2)|(D^\beta(-x)^\alpha f(x))|.$$

This shows that the Fourier transform is continuous on $\mathcal{S}$. That is if

$$f_n \to f \text{ in } \mathcal{S}$$

then

$$\hat{f}_n \to \hat{f} \text{ in } \mathcal{S}.$$

So if u is a tempered distribution then so is $f \mapsto \langle u, \hat{f}\rangle$; on the other hand if $f,\, g \in \mathcal{S}(\mathbb{R}^n)$ we have

$$\langle \hat{f}, g\rangle = \langle f, \hat{g}\rangle.$$

So for a tempered distribution u we define its Fourier transform by

$$\langle \hat{u}, f\rangle = \langle u, \hat{f}\rangle.$$

Example 6. *What is the Fourier transform of the delta function δ_0?*

$$\begin{aligned}\langle \hat{\delta}_0, f\rangle &= \langle \delta_0, \hat{f}\rangle \\ &= \hat{f}(0) \\ &= \left(\frac{1}{2\pi}\right)^{n/2} \int f(x)dx\end{aligned}$$

So $\hat{\delta}_0$ is the constant function $\left(\frac{1}{2\pi}\right)^{n/2}$.

We can extend the relations we proved for the Fourier transform on Schwartz space to cover the Fourier transform on distributions too. If u is a tempered distribution and f is Schwartz then

$$\begin{aligned}\langle \widehat{D_j u}, f\rangle &= \langle D_j u, \hat{f}\rangle \\ &= -\langle u, D_j \hat{f}\rangle \\ &= \langle u, \widehat{\xi_j f}\rangle \\ &= \langle \hat{u}, \xi_j f\rangle \\ &= \langle \xi_j \hat{u}, f\rangle.\end{aligned}$$

So

$$\widehat{D_j u} = \xi_j \hat{u}.$$

Note here that as f is being paired with $\hat{u}$ we regard it as being a function of ξ rather than x. A similar argument shows that

$$\widehat{x_j u} = -D_j \hat{u}.$$

We can now swiftly deduce that the Fourier transform of a derivative of the delta function is a polynomial and vice-versa.

We have already made some progress towards solving PDEs - if $P(D)$ is such that $p(\xi)$ is never zero - for example $P(D) = -\Delta + 1$. Then to solve
$$P(D)u = f$$
we let
$$\hat{u} = p(\xi)^{-1}\hat{f}$$
and the fundamental solution is just given by
$$\hat{u} = p(\xi)^{-1}\left(\frac{1}{2\pi}\right)^{n/2}.$$
A large part of the theory of PDEs is different ways of dealing with the zeros of $p(\xi)$.

It is important to realize that when u is an L^1 function that one can compute the Fourier transform in the "traditional" way.

Proposition 2. *If $u \in L^1$ then*

$$\hat{u} = \left(\frac{1}{2\pi}\right)^{n/2} \int e^{-ix.\xi} u(x)dx$$

and $\hat{u}$ is a continuous function.

Proof. To see the first statement, just observe that if f is Schwartz then

$$\langle \hat{u}, f\rangle = \langle u, \hat{f}\rangle = \left(\frac{1}{2\pi}\right)^{n/2} \iint e^{-ix.\xi} u(x) f(\xi) dx d\xi$$

as all integrals involved are absolutely convergent there is no problem with interchanging orders and so the result is immediate.

The second statement follows from the Dominated Convergence theorem as

$$\int |e^{-ix.\xi} f(x)| dx$$

exists. □

In fact, the Fourier transform will tend to zero at infinity - this is called the Riemann-Lebesgue lemma. An important corollary to this for us is

Corollary 5. *If $q(\xi)$ is smooth and*

$$|q(\xi)| \le C\langle\xi\rangle^{-n-\epsilon-l}$$

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

NB $< \xi >= (1 + |\xi|^2)^{1/2}$.

Proof. Just observe that the $D_x^\alpha \hat{q}$ is the Fourier transform of an L^1 function. □

If P is an elliptic operator then $p(\xi)$ is non-zero for ξ large enough as

$$|p(\xi)| \ge C|\xi|^m - C'|\xi|^{m-1}.$$

So, we can find something almost good as a fundamental solution by putting

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

where $\psi \in C_0^\infty(\mathbb{R}^n)$ is identically one in a neighbourhood of the zeroes of p. We then have that

$$\widehat{P(D)E} = \left(\frac{1}{2\pi}\right)^{n/2} - \left(\frac{1}{2\pi}\right)^{n/2} \psi(\xi)$$

which means that

$$P(D)E = \delta_0 - \hat{\psi}(-x)$$

so E is a fundamental solution up to a Schwartz error. In fact, we can make this Schwartz error be supported in arbitrarily small set about the origin, once we have proven

Theorem 14. *If $P(D)$ is elliptic of order k, ψ is a smooth function of compact support identically one on a neighbourhood of the zero set of p and*

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

then

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

with $f \in \mathcal{S}(\mathbb{R}^n)$ and E is singular only at the origin.

We say E is a *parametrix* for P.

Proof. To check the singularity property, we show that if ϕ is a smooth function of compact support, support in $x_j \neq 0$ for some j then ϕE is smooth.

We have that

$$\phi E = (x_j^{-l}\phi)x_j^l E$$

so its enough to show that $x_j^l E$ is C^α for any α for l sufficiently large. (l depending on α of course!) Now $x_j^l E$ is the inverse Fourier transform of

$$D_\xi^l \left(\frac{1-\psi(\xi)}{p(\xi)}\right).$$

So need to show that differentiating increases the decay. Any derivative falling on χ will yield something of compact support so we need only consider

$$(1-\chi)D_\xi^l p^{-1}.$$

By using an inductive argument, we can show that $D_\xi^l p^{-1}$ is a sum of terms of the form q_r/p^r with q_r a polynomial of order s_r where $s_r - lr = -k - l$ and the result follows. □

If we now multiply the parametrix by a bump function identically one near zero and supported in the set $||x|| < \rho$. We have a parametrix supported in $||x|| < \rho$. The existence of such a parametrix will allow us to show that the singular support of $P(D)u$ is equal to that of u for any distribution u but first we have to understand what convolution means for distributions.

4.4. **More Operations on Distributions including Convolution.** We use a method going back to Schwartz to define convolutions of distributions. Convolutions are essential to using fundamental solutions to invert linear partial differential operators and have many other uses besides. Recall that if f, g are L^1 functions of which one is compactly supported then one can define

$$f * g = \int f(x-y)g(y)dy.$$

We can regard this as three operations.

- Exterior Product
$$\mathcal{D}'(\mathbb{R}^n) \otimes \mathcal{D}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^{2n})$$
$$(f, g) \mapsto f \otimes g,$$
this is the extension of the map taking (f, g) to $f(x)g(y)$ to distributions.
- A Linear Change of Coordinates We wish to do the linear change of coordinates,
$$(x, y) \mapsto (x-y, y).$$
- Push-forward - we map the distribution $h(x, y) = f(x-y)g(y)$ to $\int h(x, y)dy$. This will only be defined given certain conditions on the support of h.

We will consider each of these operations separately. Note that it is a quite useful standard technique to divide a complicated operation into a sequence of quite simple basic operations.

4.4.1. *Exterior Product.* If f and g are L^1 functions then we could define for $h \in C_c^\infty$

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

So if $u \in \mathcal{D}'(\mathbb{R}^n), v \in \mathcal{D}'(\mathbb{R}^m)$, we put $h_x(y) = h(x, y)$ and define

$$\langle u \otimes v, h \rangle = \langle u, \langle v, h_x \rangle \rangle.$$

For this to be well-defined we need to know that $\alpha(x) = \langle v, h_x \rangle$ is a smooth function of x. Now we have that

$$h(x + se_j, y) - h(x, y) - s\frac{\partial h}{\partial x_j}(x, y) = s^2\psi(x, s, y)$$

by Taylor's theorem with ψ smooth. So

$$\alpha(x + se_j) - \alpha(x) - s\langle v, \frac{\partial h}{\partial x_j}\rangle = s^2\langle v, \psi_{s,x}\rangle.$$

But as $s \to 0$, $\psi_{s,x}$ tends uniformly to $\frac{\partial \psi}{\partial x_j}(x, y)$ and the same is true for the derivatives so we conclude that

$$\frac{\partial \alpha}{\partial x_j}(x) = \langle v, \frac{\partial h}{\partial x_j}_x\rangle.$$

The smoothness then follows by induction and so our operation is well-defined. We need to check that $u \otimes v$ is continuous but if $h_n \to h$ then clearly $h_{nx} \to h_x$ so $\alpha_n \to \alpha$ but the argument above shows the derivatives of α will converge also and so we get a distribution.

4.4.2. *Linear Changes of Coordinates.* If A is a linear map on $\mathbb{R}^n$ then it induces a change of coordinate map on smooth functions:

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

The transpose of this map is

$$(A^*)^t g = f(A^{-1}) \det(A)^{-1}$$

as

$$\int f(Ax)g(x)dx = \int f(y)g(A^{-1}y) \det(A^{-1})dy.$$

The transpose is clearly continuous so we can extend to distributions in the standard way.

4.4.3. *Pairings and Push-forwards.* If $f \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^m)$ and $g \in C_c^\infty(\mathbb{R}^n)$ then we have

$$\langle \int f(x, y)dy, g(x)\rangle = \int f(x, y)g(x)dydx.$$

So the transpose operation is the pull-back

$$\pi^* : C_c^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n \times \mathbb{R}^m)$$

given by

$$\pi^*g(x, y) = g(x).$$

This is a pretty simple operation but has the problem that it maps compactly supported functions to functions which are not compactly supported! We have defined distributions to be maps from the space of compactly supported functions to the complex numbers so we have a problem. However, the pairing between a function and a distribution can in fact always be defined provided the intersection of their supports is compact.

If K is a subset of $\mathbb{R}^n$, we denote $\mathcal{D}'(K)$ to be the distributions supported in K and $C^\infty(K)$ to be the smooth functions supported there.

Proposition 3. *If $K, L \subset \mathbb{R}^n$ are such that $K \cap L$ is compact there is a well-defined pairing*

$$\mathcal{D}'(K) \times C^\infty(L) \to \mathbb{C}.$$

Proof. Let $M = K \cap L$ and let $\phi = 1$ on an open neighbourhood of M and be compactly supported. We define

$$\langle u, f \rangle_\phi = \langle u, \phi f \rangle.$$

We need to check that this is independent of the choice of ϕ but if ϕ_1 is another choice we have that the difference is

$$< \langle u, (\phi - \phi_1) f \rangle.$$

The support of $(\phi - \phi_1) f$ is disjoint from that of u so this equal to zero and the pairing is well-defined. □

We leave to the interested reader to formulate the continuity properties of this pairing.

Let π be the projection, $\pi(x, y) = x$. We say that a set, $L \subset \mathbb{R}^n \times \mathbb{R}^m$ is proper with respect to the projection if

$$\pi^{-1}(K) \cap L$$

is compact for every K compact in $\mathbb{R}^n$.

We now have

Proposition 4. *If $u \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^m)$ is proper with respect to π then there is a well-defined distribution $\pi_*(\phi)$ on $\mathbb{R}^n$.*

Proof. We just define for $f \in C_0^\infty(\mathbb{R}^n)$

$$\langle \pi_* u, f \rangle = \langle u, \pi^* f \rangle.$$

□

4.4.4. *Convolution.* So with all that done we can now define convolution to be

$$u * v = \pi_* A^*(u \otimes v)$$

where A is the map $A(x,y) = (x-y, y)$ and π is the projection $(x,y) \mapsto y$. This of course requires the push-forward to be well-defined, which is implied by requiring

$$\{(x,y) : \exists y, (x-y) \in \operatorname{supp}(u), y \in \operatorname{supp}(v)\}$$

to be compact for x in each compact set K.

The first important point here is that convolutions are always defined provided one of the distributions is of compact support. The delta function is the identity for this operation:

$$\begin{aligned}\langle \delta_0 * v, \phi \rangle &= \langle \delta_0(x) v(y), \phi(x+y) \rangle \\ &= \langle \delta_0(x), \langle v(y), \phi(x+y) \rangle_x \rangle \\ &= \langle v, \phi \rangle.\end{aligned}$$

As

$$\langle u * v, \phi \rangle = \langle u \otimes v, \phi(x+y) \rangle$$

we have that $u * v = v * u$ so δ_0 is both left and right inverse. As

$$\langle D^\alpha(u * v), \phi \rangle = (-1)^{|\alpha|} \langle u \otimes v, (D^\alpha \phi)(x+y) \rangle$$

we have that

$$D^\alpha(u * v) = (D^\alpha u) * v = u * D^\alpha v$$

and that

$$D^\alpha u = D^\alpha \delta_0 * u.$$

This last statement implies that applying any differential operator is equivalent to convolving with a distribution.

Another important property of convolutions is associativity - without associativity one can do nothing! This follows from observing that

$$\langle u * (v * w), \phi \rangle = \langle u \otimes v \otimes w, \phi(x+y+z) \rangle,$$

provided the right hand pairing make sense i.e. that

$$\{(x,y,x) : x+y+z \in K, x \in \operatorname{supp}(u), y \in \operatorname{supp}(v), z \in \operatorname{supp}(z)\}$$

is compact for all K compact - so *provided this condition is satisfied, we have associativity.* (Of course, this requires exterior product to be associative but it is.) Note that if E is fundamental solution then regarding the operator $P(D)$ as being convolution with $P(D)\delta_0$ we have that

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

and that if $P(D)u = f$ then $u = E * f$ provided the convolution exists! Note that while a fundamental solution is a left and right inverse - the standard proof that left and right inverses are equal only works when the fundamental solutions are convolvable. So a differential operator can have many fundamental solutions. One can think of these as being inverses on different classes of functions.

For example, if $\frac{\partial u}{\partial x_1} = f$ then if $u, f = 0$ for x_1 large negative, we have that

$$u = \int_{-\infty}^{x_1} f(s, x'')ds$$

and if u, f are zero for x_1 large positive we have that

$$u = -\int_{x_1}^{\infty} f(s, x'')ds.$$

We can regard these solutions as being convolution with the fundamental solutions $H(x_1)$ and $H(x_1) - 1$.

Another important fact about convolving is that convolving a distribution with a smooth function always yields a smooth function. (given the condition on supports.) The philosophical reason this is true is that if you differentiate you can put all the differentiations onto the smooth part which remains smooth so differentiating never makes the convolution worse so it must be smooth. How do we prove it though? Let u be a distribution and f a smooth function which are convolvable and consider the map

$$x \mapsto \langle u, g_x \rangle$$

where $g_x = f(x - y)$. This is a smooth as a function of x - same argument we used when defining exterior products. This is morally the convolution - we check

$$\begin{aligned}\langle \langle f(x-y), u(y) \rangle_x, \phi(x) \rangle &= \langle u(y), f(x-y)\phi(y) \rangle \\ &= \langle f(x-y)u(y), \phi(x) \rangle \\ &= \langle f * u, \phi \rangle.\end{aligned}$$

It is quite easy to establish an upper bound for the support of a convolution in terms of the supports of the original distributions:

$$\operatorname{supp}(f * g) \subset \{x : \exists y, y \in \operatorname{supp}(f), x - y \in \operatorname{supp}(g)\}.$$

This is easy to prove - one just computes the action on supports of each of the three basic operations.

With all this done, we can now prove an important theorem - elliptic regularity. This is sometimes expressed as "weak implies strong" as it says that every weak or distributional solution is also a smooth or strong solution. This is historically very important as one can construct distributional solutions of elliptic PDEs using the Hahn-Banach theorem and it is then useful to know that they are in fact ordinary solutions too.

Theorem 15. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $P(D)$ is an elliptic differential operator then u and $P(D)u$ have the same singular support. In particular,*

$$P(D)u \in C^\infty \implies u \in C^\infty.$$

Proof. Given $\rho > 0$ there is a parametrix E supported in $||x|| < \rho$. So

$$P(D)E = \delta_0 + f, \quad EP(D) = \delta_0 + g$$

with f, g smooth and supported in $||x|| < \rho$. So

$$u + g * u = E * (P(D)u).$$

Now $g * u$ is smooth as g is. So u and $E * (P(D)u)$ have the same singularities. Thus if $P(D)u$ is smooth, so is u.

If $P(D)u = h$ is not smooth then we can write it as $h_1 + h_2$ with h_2 smooth and h_1 supported within ρ of the singular support of h. So u has the same singularities as $E * h_1$ but $E * h_1$ is supported within 2ρ of the singular support of h.

This is true for any $\rho > 0$ and singular support is contained in support so the result follows. □

Remark 2. *While our proof required $P(D)$ to be constant coefficient, this result holds in the variable coefficient case too and this is an important part of Hodge theory which relates the cohomology of a manifold to its differential geometry. Note this theorem is often called "Weyl's lemma."*

A very important fact is that every constant coefficient operator has a fundamental solution.

Theorem 16. *If $P(D)$ is a constant coefficient operator on $\mathbb{R}^n$ then there exists a distribution E such that*

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

We will prove this only in a special (but important case), see for example [1] for the general case.

We assume that there exists an $a \in \mathbb{R}^n$ such that

$$|p(\xi + ia)| \geq \alpha > 0 \ \ \forall \xi \in \mathbb{R}^n.$$

Example 7. *If* $P(D) = \frac{\partial}{\partial t} - \sum_i \frac{\partial^2}{\partial x_i^2}$ *then* $p(\xi, \tau) = i\tau + \xi^2$. *So,*

$$\begin{aligned} p(\tau + ia, \xi) =& \xi^2 + i(\tau + ia) \\ =& \xi^2 + i\tau - a. \end{aligned}$$

So taking $a < 0$, *we have*

$$|p(\tau + ia, \xi)| \geq -a > 0.$$

Note the asymmetry here.

Example 8. *If* $P = \frac{\partial^2}{\partial t^2} - \sum_i \frac{\partial^2}{\partial x_i^2}$ *then* $p(\xi, \tau) = \xi^2 - \tau^2$. *For* $a \in \mathbb{R}$, *we have*

$$\begin{aligned} \left|\xi^2 - (\tau + ia)^2\right|^2 &= \left(\xi^2 - \tau^2 + a^2\right)^2 + 4a^2\tau^2 \\ &= ||\xi||^4 + \tau^4 + a^4 + 4a^2\tau^2 - 2||\xi||^2\tau^2 - 2a^2\tau^2 + 2a^2||\xi||^2 \\ &= (\xi^2 - \tau^2)^2 + a^4 + 2a^2\tau^2 + 2a^2\xi^2 \\ &\geq a^4. \end{aligned}$$

Before we can prove our result we need some results on the analyticity and growth of Fourier transforms of compactly supported smooth functions so we can move contours around - this will allow us to move a towards 0.

Proposition 5. *(Paley-Wiener Estimate) If* $f \in C_0^\infty(\mathbb{R}^n)$ *then* $\hat{f}(\xi)$ *has an analytic extension to* $\mathbb{C}^N$, $\hat{f}(z)$ *and*

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

Note that this implies that $\hat{f}(x + iy)$ is uniformly Schwartz in x for y in a compact set - that is the norms $\|f\|_{\alpha,beta}$ regarded as functions of y are bounded for y in compact sets.

Proof. As f is compactly supported, the integral

$$\hat{f}(z) = \int e^{-ix.z} f(x) dx$$

will converge for any $z \in \mathbb{C}^N$. This will be the extension. As everything is smooth and compactly supported, we can differentiate under the integral sign so $\hat{f}(s + it)$ is smooth and

$$z^\alpha \hat{f} = \widehat{D^\alpha f}.$$

But $e^{-ix.z}$ is analytic in z and so satisfies the Cauchy-Riemann equations in (s,t) and so $\hat{f}$ does also. Thus $\hat{f}$ is analytic. To get the estimate on growth, just estimate $\widehat{D^\alpha f}$. □

Now, if $|p(\xi + a)| \geq \alpha > 0 \ \forall \xi$, for some fixed $a \in i\mathbb{R}^n$. We define

$$\langle E, \phi \rangle = \left(\frac{1}{2\pi}\right)^{n/2} \int \frac{\hat{\phi}(-\xi - a)}{p(\xi + a)} d\xi.$$

We compute

$$\begin{aligned} \langle P(D)E, \phi \rangle &= \langle E, P(-D)\phi \rangle \\ &= \left(\frac{1}{2\pi}\right)^{n/2} \int \frac{\hat{\phi}(-\xi - a)}{p(\xi + a)} p(\xi + a) d\xi \\ &= \left(\frac{1}{2\pi}\right)^{n/2} \int \hat{\phi}(-\xi - a) d\xi \end{aligned}$$

We can now shift the contour using the Paley-Wiener estimate and Cauchy's theorem to conclude that this is equal to

$$\left(\frac{1}{2\pi}\right)^{n/2} \int \hat{\phi}(-\xi) d\xi = \phi(0).$$

So E is indeed a fundamental solution.

Note that, although initially, it seems that we have a different fundamental solution for each a, if we have $|p(\xi + a)| \geq \alpha > 0$ for a in a connected, compact set then the same shifting contour argument as above yields that the fundamental solutions obtained are the same. So for the wave equation, we obtain two different fundamental solutions according to the sign of $\Im a$. For the heat equation we only obtain one - from a such that $\Im a > 0$.

5. The Laplacian

5.1. **Finding the Fundamental Solution.** We have already proven some things about elliptic operators of which the Laplacian is a special case - they always have a compactly supported parametrix which is singular only at the origin and any distributional solution is always a smooth function. This implies that any fundamental solution is singular only at the origin as if E is fundamental solution and P is a parametrix then

$$\Delta(E-P)\in C^\infty$$

and so by elliptic regularity

$$E-P\in C^\infty.$$

Now an important fact about the Laplacian is that is rotationally invariant. The highbrow way to see this is to observe that the Laplacian is defined by the metric and that rotations are an isometry. The low brow way is just to compute

$$\Delta(f(Ax)),(\Delta f)(Ax)$$

for an orthogonal matrix and observe that they are equal. (Indeed as two-dimensional rotations generate, it is enough to do this for two-dimensional ones only.) One therefore expects there to exist a fundamental solution which is rotationally invariant - indeed given that there exists some fundamental solution, E, one can construct a rotationally invariant one by averaging:

$$E'=\int_{SO(n)}A^*E dA.$$

But this requires an understanding of integration over Lie groups which is beyond our scope.

Another important guide to finding fundamental solutions is homogeneity. A differentiable function, f, on $\mathbb{R}^n-\{0\}$ is (positively) homogeneous of degree m if

$$f(\lambda x)=\lambda^m f(x)\ \ \forall\lambda>0.$$

We do not require differentiability or continuity at the origin as the class would then be very small!

Now, if we differentiate with respect to λ and set $\lambda=1$ then we deduce that

$$\left(x\frac{\partial}{\partial x}-m\right)f=0,$$

where $x\frac{\partial}{\partial x} = \sum_{j=1}^{n} x_j \frac{\partial}{\partial x_j}$. This is called Euler's relation and is in fact equivalent to homogeneity. To prove this observe that if f satisfies the equation then

$$||x||^m f(x/||x||)$$

agrees with f on $||x|| = 1$ and satisfies the same equation so using our uniqueness theorems for first order, real PDEs they must be equal.

Definition 1. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ then u is homogeneous of degree $m \in \mathbb{C}$ if*

$$\left(x\frac{\partial}{\partial x} - m\right) u = 0.$$

For the delta function we compute, for a test function ϕ,

$$\langle x_j \frac{\partial}{\partial x_j}\delta, \phi\rangle = -\langle \delta, \frac{\partial}{\partial x_j}(x_j\phi)\rangle \tag{5.1}$$

$$= -\langle \delta, \phi + x_j \frac{\partial \phi}{\partial x_j}\rangle \tag{5.2}$$

$$= -\langle \delta, \phi\rangle. \tag{5.3}$$

We thus deduce that the delta function is homogeneous of degree $-n$.

Differentiating a distribution reduces its order of homogeneity by one; to see this

$$\frac{\partial}{\partial x_j}\left(x\frac{\partial}{\partial x} - m\right) u = \left(x\frac{\partial}{\partial x} - (m-1)\right)\frac{\partial u}{\partial x_j}.$$

So since the Laplacian is of order 2 and has no lower order parts, we expect its fundamental solution to be homogeneous of order $2-n$ and to be rotationally invariant. It also must be smooth away the origin. Thus the obvious candidate is

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

this is an L^1 function near 0 and so defines a distribution. (In fact, this is the only distribution with these properties.) Now, away from the $x = 0$ we can compute in the usual way to find that

$$\Delta ||x||^{2-n} = 0$$

so we have that $\Delta ||x||^{2-n}$ is both supported at the origin and singular there, (using elliptic regularity). It is also homogeneous of degree $-n$. (We could deduce from this that it is a multiple of the delta function but this would require too much theory.)

Let ψ be a bump function, identically 1 near 0, and $\psi_\epsilon(x) = \psi(x/\epsilon)$. The support of ψ yields that $\langle \Delta E_n, \psi_\epsilon \phi \rangle$ is independent of ϵ. Now using Taylor's theorem, we have that

$$\langle \Delta E_n, \phi \rangle = \phi(0)\langle u, \psi_\epsilon \rangle + \sum_{|\alpha| \leq N} \langle \Delta E_n, x^\alpha f_\alpha \rangle + \langle \Delta E_n, \psi_\epsilon h \rangle,$$

where h is smooth and vanishes to N^{th} order at $x = 0$. The first term is independent of ϵ from the support properties and the second is zero from the homogeneity of ΔE_N - its pairing with any function of the form $x\frac{\partial}{\partial x} g(x)$ is zero and $x^\alpha = \frac{1}{\alpha!} x \frac{\partial}{\partial x} x^\alpha$. The last term equals

$$\int ||x||^{2-n} \Delta(\psi(x/\epsilon)h) dx$$

and this will go to zero as $\epsilon \to 0$ provided N is sufficiently big as h vanishes to order N.

So we know we have a multiple of the delta function - we need to know which one! Let ψ be smooth, radial and identically one near 0, we compute, using polar coordinates,

$$\begin{aligned}
\langle \Delta ||x||^{2-n}, \psi(|x|) \rangle &= \langle ||x||^{2-n}, \Delta \psi(|x|) \rangle \\
&= \omega_{n-1} \int_0^\infty r^{2-n} r^{n-1} \left(\frac{\partial^2 \psi}{\partial r^2} + \frac{n-1}{r} \frac{\partial \psi}{\partial r} \right) dr \\
&= \omega_{n-1} \int_0^\infty r \frac{\partial^2 \psi}{\partial r^2} + (n-1) \frac{\partial \psi}{\partial r} dr \\
&= (2-n)\omega_{n-1}
\end{aligned}$$

where ω_{n-1} is the volume of the unit sphere. We have thus solved the problem when $n \geq 3$. The argument does not quite work when $n = 2$ as in that case $||x||^{2-n}$ is constant and so smooth. We therefore use $\log(||x||)$ instead. This is not homogeneous but is *almost homogeneous* in that

$$x \frac{\partial}{\partial x} \log(||x||) \in C^\infty.$$

We then have that $\Delta \log(||x||)$ is almost homogeneous and supported at the origin which implies that it is homogeneous as a smooth function supported at the origin is zero. The arguments then go through as before and we have proven:

Theorem 17. *The Laplacian on $\mathbb{R}^n$ has the fundamental solution*

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

for $n \geq 3$ and for $n = 2$

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

5.2. **Identities and Estimates.**

Theorem 18. *Gauss Divergence Theorem If $f \in C(\overline{B}(0,1))$, C^1 in $B(0,1)$ then*

$$\int_{\|x\|\leq 1} \frac{\partial f}{\partial x_j} dx = \int_{\|x\|=1} f(x)\frac{x_j}{\|x\|} dx \tag{5.4}$$

Hence

$$\int_{\|x\|\leq 1} \nabla \cdot g dx = \int_{\|x\|=1} g \cdot n_x dx \tag{5.5}$$

where $n_x = \frac{x}{\|x\|}$ is the outward normal at $x \in S^{n-1}$ and g is vector valued.

Proof. The second statement is just a summation of the first one. Take $\phi_\epsilon(r)$ s.t. $\phi_\epsilon = 1$ if $r < 1-\epsilon$, $\phi = 0$ if $r \geq 1$. Then $\int_{\|x\|\leq 1} f = \lim_{\epsilon\to 0}\langle f, \phi_\epsilon\rangle$ But

$$\begin{aligned}
\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_j}\phi_\epsilon dx &= -\int f \frac{\partial \phi_\epsilon}{\partial x_j} dx \\
&= -\int f \frac{\partial r}{\partial x_j}\frac{\partial \phi_\epsilon}{\partial r} dx \\
&= -\int f \frac{x_j}{\|x\|}\frac{\partial \phi_\epsilon}{\partial r} dx \\
&= -\int \int_{||x||=r} \left(f\frac{x_j}{||x||}\right)(r\omega) d\omega r^{n-1}\frac{\partial \phi_\epsilon}{\partial r} dr \\
&= \int \frac{\partial}{\partial r}\left(r^{n-1}\left(f\frac{x_j}{||x||}\right)(r\omega) d\omega\right)\phi_\epsilon dr
\end{aligned}$$

Letting $\epsilon \to 0$ this is equal to

$$\int_0^1 \frac{\partial}{\partial r}\left(\int r^{n-1}\left(f\frac{x_j}{||x||}\right)(r\omega) d\omega\right) dr$$

which of course just

$$\int \left(f\frac{x_j}{||x||}\right)(\omega) d\omega.$$

□

Proposition 6. *Green's Identities: If u, v are C^2 functions on $\overline{B}(0,r)$ then*

$$(a) \qquad \int_B v\Delta u = -\int_B \nabla v\cdot \nabla u + \int_S v.\frac{du}{d\nu}$$

where $\frac{du}{d\nu} = \sum x_i \frac{du}{dx_i} = r\frac{du}{dr}$ *and* $S = \partial B$*.*

$$(b) \qquad \int_B v\Delta u - u\Delta v = \int_S v\frac{du}{d\nu} - u\frac{dv}{d\nu}$$

Proof. (a) follows from Gauss' divergence theorem applied to $v\nabla u = (v\frac{du}{d\nu})$

(b) follows by swapping u and v in (a) and subtracting. ☐

Remark 3. *Setting $v = 1$, we obtain $\int_B \Delta u = \int_S \frac{\mathrm{d}\,u}{\mathrm{d}\,v}$. There are obvious results for spherical shells obtained by subtracting identities (a) & (b) for different values of r.*

Proposition 7. *Energy Estimate:* $\int_B |\nabla u|^2 = \int_S u\frac{du}{d\nu} - \int_B u\Delta u$*.*

To prove this just set $u = v$ in (a).

Thus if $\Delta u = 0$ in B and either u or $\frac{\mathrm{d}\,u}{\mathrm{d}\,v} = 0$ on S, we have $\nabla u = 0$ on B, ie u is a constant, necessarily 0 in first case. Hence we can deduce the uniqueness of solutions to some boundary value problems.

Dirichlet Problem Solve for u with Δu prescribed in B and u prescribed on S.

Neumann Problem Solve for u with Δu prescribed in B and $\frac{du}{d\nu}$ prescribed on S. and so to summarise we have proven:

Theorem 19. *The Dirichlet Problem has at most one solution and the Neumann Problem has at most one solution up to a constant. (for the unit ball)*

Proposition 8. *Weak Maximum Principle: Let Ω be a bounded open domain $\mathbb{R}^n$ with smooth boundary $\partial\Omega$. If $u \in C(\overline{\Omega})$, $u \in C^2(\Omega)$ with $\Delta u \geq 0$ in Ω then* $\sup_{x\in\overline{\Omega}} u(x) = \sup_{x\in\partial\Omega} u(x)$

Proof. If $\Delta u > 0$ in Ω, the result is easy. For at a local maximum in Ω we must have $\frac{\partial u}{\partial x_i} = 0\ \forall i$. But $\Delta u > 0$ implies $\frac{\partial^2 u}{\partial x_j^2} > 0$ some j. So in the x_j direction we can increase u.

Now take $v(x) = \|x\|^2$. Then $\Delta v > 0$ and hence $\Delta(u + \epsilon v) > 0$ on Ω. So from the first part

$$\sup_{\overline{\Omega}}(u + \epsilon v) = \sup_{\partial\Omega}(u + \epsilon v)$$

So $\sup_{\overline{\Omega}} u \leq \sup_{\partial\Omega}(u + \epsilon v) \leq \sup_{\partial\Omega} u + \epsilon \sup_{\partial\Omega} v$. Now let $\epsilon \to 0$ to get the result. □

A function, f, such that Δf is zero is often said to be harmonic.

Corollary 6. *If u is complex valued and harmonic*

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

Proof. Apply the maximum principle to $Re(e^{i\theta}u)$ where θ is chosen, so that

$$\sup_{\overline{\Omega}} Re(e^{i\theta}u) = \sup_{\overline{\Omega}} |u|.$$

□

Proposition 9. *Gauss Mean Value Property:*

(a) The value of a harmonic function at a point is equal to its average over any sphere 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 will see that property (a) characterises harmonic functions.

Proof. In fact, we prove that if $\Delta u \geq 0$ then these are in fact true with inequalities - the result will then follow by applying the inequality for u and $-u$.

By translation invariance, it suffices to consider the case where the point is the origin.

First, putting $v = 1$ in Green's identity we observe that

$$\int_{S(r)} \frac{\partial u}{\partial \nu} dS = \int_{B(r)} \Delta u dx \geq 0.$$

Now, if v is the function $\|x\|^2$ then $\Delta v = 2n$ and applying Green's identity we have

$$2n \int_{B(r)} u dx \leq 2r \int_{S(r)} u(x) dS$$

so putting

$$\phi(r) = \int_{B(r)} u \, dx$$

we have that

$$\phi'(r) = \int\limits_{S(r)} u\, dS \geq \frac{n}{r} \int\limits_{B(r)} u(x)dx = \frac{n}{r}\phi(r).$$

This implies that

$$\frac{d}{dr}\left(r^{-n}\phi(r)\right) \geq 0$$

or that $r^{-n}\phi(r)$ is increasing so we conclude that

$$\lim_{r\to 0} r^{-n}\phi(r) \leq R^{-n}\phi(R) \leq \frac{R^{1-n}}{n}\int_{S(R)} u(x)dS.$$

As the surface area of a sphere of radius R is n/R times the volume the result now follows. ☐

Proposition 10. *Strong Maximum Principle: Suppose*

$$u(\xi) \leq \frac{1}{A(\rho)} \int\limits_{\|x-\xi\|=\rho} u(x) \tag{5.6}$$

for all ξ and ρ sufficiently small in Ω. If Ω is a bounded open connected *domain, then either u is constant or $u(\xi) < \sup_{\partial\Omega} u(x)$ for all $\xi \in \Omega$.*

Definition 2. *Any $u \in C(\overline{\Omega})$ satisfying (5.6) for sufficiently small balls is said to be sub-harmonic. So any $u \in C(\overline{\Omega})$, C^2 in Ω such that $\Delta u \geq 0$ is subharmonic.*

Proof of Strong Maximum Principle. Let $M = \sup u$ and define

$$\Omega_1 = \{x|u(x) = M\} \;\; \Omega_2 = \{x|u(x) < M\}.$$

Then $\Omega = \Omega_1 \sqcup \Omega_2$ (disjoint union). The continuity of u implies that Ω_2 is open. If we can show Ω_1 is also open, the result follows immediately from connectivity since either Ω_1 or Ω_2 is then empty.

Say $u(\xi) = M$ for $\xi \in \Omega$. Then

$$0 \leq \operatorname*{Av}_{\|x-\xi\|\leq p}(M - u(x)) \leq M - u(\xi) = 0$$

Since $M - u(x) \geq 0$ is continuous, this force $u(x) = M$ for $\|x - \xi\| \leq p$. So Ω_1 is open, as required. ☐

5.3. **The Dirichlet and Dual Dirichlet Problems.** Let Ω be a bounded open set in $\mathbb{R}^n$ with smooth boundary $\partial\Omega$.

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

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

These problems are more or less equivalent:

(a) $\Rightarrow$ (b): Define $\tilde{g}$ to be the distribution obtained by setting $\tilde{g} = g$ on Ω and 0 on $\mathbb{R}^n \backslash \Omega$. Set $f_1 = E * \tilde{g}$ where $E =$ find solution of Δ. Then $\Delta f_1 = \tilde{g}$. Solve $\Delta h = 0$, $h = f_1$ on $\partial\Omega$ using (a). Then $\Delta(f_1 - h) = g$ and $f_1 - h = 0$ on $\partial\Omega$.

(b) $\Rightarrow$ (a): Extend h to a function $\tilde{h}$ on $\overline{\Omega}$ which is C^2 on Ω. Solve $\Delta k = \Delta \tilde{h}$ in Ω, $k = 0$ on $\partial\Omega$. Then $\Delta(\tilde{h} - k) = 0$ in Ω and $\tilde{h} - k = h$ on $\partial\Omega$.

5.4. **The Dual Dirichlet Problem for the Unit Ball.** We have to solve

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

We shall assume that $g \in C(\overline{B})$ with $g \in C^\infty$ in B. (Actually once a formula for the solution has been obtained, it will be clear that far less restrictive conditions on g are needed for a solution.)

We start by looking at $f_1 = E * \tilde{g}$ where $\tilde{g}$ is the distribution $= g$ on $\overline{B}$ and 0 on $\mathbb{R}^n \backslash \overline{B}$. (Here E is the fundamental solution of Δ obtained before.) Clearly $\Delta(E * \tilde{g}) = \tilde{g}$ as distributions. We assume $n \geq 3$. (we will discuss $n = 2$ later.)

Lemma 7. *$f_1 \in C^1(\mathbb{R}^n)$ and $\frac{\partial f_1}{\partial x_j} = \frac{\partial E}{\partial x_j} * \tilde{g}$. We also have that $f_1, \frac{\partial f_1}{\partial x_j}$ are of order $\|x\|^{2-n}, \|x\|^{1-n}$ for x large.*

Proof. We show all the derivatives are continuous functions. We compute $\frac{\partial E}{\partial x_j} * \tilde{g}$. This is a convolution of a locally integrable function and a function of compact support and so can be computed directly. Suppose $\|x - x'\| \leq \epsilon$, then in some fixed ball

$$\begin{aligned} \left| \frac{\partial f_1}{\partial x_j}(x) - \frac{\partial f_1}{\partial x_j}(x') \right| &\leq \int \left| \frac{\partial E}{\partial x_j}(x-y) - \frac{\partial E}{\partial x_j}(x'-y) \right| |g(y)|\, dy \\ &\leq \int_{\delta \leq \|x-y\| \leq 1} \left| \frac{\partial E}{\partial x_j}(x-y) - \frac{\partial E}{\partial x_j}(x'-y) \right| |g(y)|\, dy \\ &\quad + 2 \sup |g| \int_{\|x-y\| \leq \delta+\epsilon} \left| \frac{\partial E}{\partial x_j}(x-y) \right| dy \end{aligned}$$

The second integral is small if $\delta+\epsilon$ is small. The first integral is small will go to zero as $x-x'\to 0$ as it is smoothly dependent on them.

Hence $\frac{\partial f}{\partial x_j}$ is continuous. To see the bounds, observe that if $|x|>2$ then $|x-y|\geq |x|/2$ for $|y|<1$ and then just estimate

$$\int_{|y|\leq 1} E(x-y)g(y)dy$$

and

$$\int_{|y|\leq 1} \frac{\partial E}{\partial x_j}(x-y)g(y)dy.$$

□

Lemma 8. *f_1 is C^∞ in $\|x\|<1$ and $\|x\|>1$; in fact f_1 is harmonic in $\|x\|>1$.*

Proof. This is an immediate consequence of elliptic regularity. □

At the moment we have $\Delta f_1 = g$ but f_1 has the wrong boundary value on $\|x\|=1$. The way round this, due to Kelvin, is through the method of *reflection*. We put

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

We then have that Kf_1 and f_1 are equal on the unit sphere and so $f_1 - Kf_1$ is zero there. We will show that K almost commutes with the Laplacian and so will be able to deduce that Kf_1 is harmonic in $\|x\|<1$ and so

$$\Delta(f_1 - Kf_1) = \Delta f_1 = g$$

in the ball - thus solving the problem!

Lemma 9.

$$\Delta(Kf) = \|x\|^{-4}K(\Delta f), \quad r\frac{\partial}{\partial r}Kf = (-n+2)Kf - K(r\frac{\partial}{\partial r}f)$$

Here $r\frac{\partial}{\partial r} = \sum x_i \frac{\partial}{\partial x_i}$.

Proof. Observe that the Laplacian and the Kelvin transform are rotationally invariant and thus so is the statement of the lemma. This means it is enough to show it is true at the point

$$x = (\|x\|, 0, \dots, 0).$$

At such a point, for $j > 1$,

$$\frac{\partial^2}{\partial x_j^2}(f(x/\|x\|^2)) = \|x\|^{-4}\frac{\partial^2 f}{\partial y_j^2} - 2\|x\|^{-3}\frac{\partial f}{\partial y_1}$$

and

$$\frac{\partial^2}{\partial x_1^2}(f(x/\|x\|^2)) = \|x\|^{-4}\frac{\partial^2 f}{\partial y_1^2} + 2\|x\|^{-3}\frac{\partial f}{\partial y_1}.$$

So

$$\Delta(f(x/\|x\|^2)) = \|x\|^{-4}(\Delta f)(x/\|x\|^2) + 2(2-n)\|x\|^{-3}\frac{\partial f}{\partial y_1}.$$

Using,

$$\Delta(ab) = b\Delta a + 2\nabla a.\nabla b + a\Delta b$$

the result then follows.

For the second equality, reduce to the same case as above and then compute as for one dimension. □

Now let $f_2 = Kf_1$. By the lemma, f_2 is harmonic near each $x \neq 0$ with $\|x\| < 1$. However we have to check it does satisfy the equation at $x = 0$.

Lemma 10. *f_2 is harmonic near 0.*

Proof. We must show that if $\varphi \in C_c^\infty(\mathbb{R}^n)$ with $\overline{\text{supp}}\varphi \subseteq \overline{B}(0,p)$ then $\langle \Delta f_2, \varphi\rangle = 0$, i.e. $\langle f_2, \Delta\varphi\rangle = 0$. But

$$\begin{aligned}
\langle f_2, \Delta\varphi\rangle &= \int f_2\Delta\varphi \\
&= \int_{\|x\|\le\epsilon} f_2\Delta\varphi + \int_{\|x\|\ge\epsilon} f_2\Delta\varphi \\
&= \int_{\|x\|\le\epsilon} f_2\Delta\varphi + \int_{\|x\|\ge\epsilon} \Delta f_2\varphi - \int_{\|x\|=\epsilon} f_2\frac{d\varphi}{d\nu} + \int_{\|x\|=\epsilon}\frac{df_2}{d\nu}\varphi \\
&= \int_{\|x\|\le\epsilon} f_2\Delta\varphi - \int_{\|x\|=\epsilon} f_2\frac{d\varphi}{d\nu} + \int_{\|x\|=\epsilon}\frac{df_2}{d\nu}\varphi
\end{aligned}$$

¿From the lemmas, f_2 and $\frac{df_2}{d\nu}$ are $O(\|x\|^0), O(\|x\|^{-1})$ for $\|x\|$ small. So the integrals tend to 0 as $\epsilon \to 0$. Hence $\langle f_2, \Delta\varphi\rangle = 0$ as required. □

So tracing through the arguments above, we have

Theorem 20. *For $n \geq 3$, let*

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

Then $f(x) = \int_{\|y\|<1} G(x,y)g(y)dy$ *solves the dual Dirichlet problem.* $G(x,y)$ *is called the* **Green's function**. *When* $n = 2$, *we obtain*

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

Note that while the Kelvin transform may seem a little mysterious, it can be given a geometric interpretation: map Euclidean space to the sphere by stereographic projection and then reflect in the equator and then map back to the plane.

The energetic can check that the appropriately modified arguments also work in the two dimensional case.

5.5. **Deduction of the Poisson formula for the Dirichlet problem.** We now wish to solve $\Delta f = 0$ in B, $f = h$ on S. We compute f near a point a such that $\|a\| < 1$.

According to the previous prescription, we extend h to a continuous function $\tilde{h}$ on $\overline{B}$, C^2 on B.

Set $\tilde{h}(x) = h(\frac{x}{\|x\|})\psi(\|x\|)$ where ψ is a bump function identically 1 near 1 and supported in $[1-\epsilon, 1+\epsilon]$ where $\|a\| < 1-\epsilon$. Set $k(x) = \int G(x,y)\Delta\tilde{h}(y)dy$. Then $k = 0$ on S and $\Delta k = \Delta\tilde{h}$ in B. So $\Delta(\tilde{h}-k) = 0$ and $\tilde{h} - k = h$ on S. So we set $f(x) = \tilde{h}(x) - k(x)$.

Thus we get

$$f(x) = h(\frac{x}{\|x\|})\psi(\|x\|) - \int_B G(x,y)\Delta(h(\frac{y}{\|y\|})\psi(\|y\|))dy$$

Now let $x = a$. Using Green's formula and the fact that $G(x,y)$ is harmonic in y for $y \neq x$, we get

$$\begin{aligned} f(a) &= \int_S \frac{d}{dn_y}G(x,y)h(\frac{y}{\|y\|})\psi(\|y\|)dy \\ &\quad - \int_S G(a,y)h(\frac{y}{\|y\|})\frac{d}{dn_y}\psi(\|y\|)dy \\ &= \int_S \frac{dG(a,y)}{dn_y}h(y)dy \end{aligned}$$

since $\psi(1) = 1$ and $\psi(r) = 1$ for r near 1, so that $r\frac{\partial}{\partial r}\psi = 0$ for $r = 1$. Thus we get for $\|x\| < 1$

$$f(x) = \int_S \frac{dG(x,y)}{dn_y}h(y)dy \equiv \int_S P(x,y)h(y)dy$$

Since $\sum_{i=1}^n y_i \frac{\partial}{\partial y_i}\frac{1}{\|y-a\|^{n-2}} = (2-n)\frac{\|y\|^2-\langle a,y\rangle}{\|y-a\|^n}$, we get

$$\begin{aligned} P(x,y) &= \frac{dG(x,y)}{dn_y} \\ &= \frac{1}{\omega_{n-1}}\frac{1-\langle x,y\rangle}{\|x-y\|^n} - \frac{\|x\|^{n-2}}{\omega_{n-1}}\frac{(1-\langle \frac{x}{\|x\|^2}, y\rangle)}{\|\frac{x}{\|x\|^2}-y\|^n} \\ &= \frac{1}{\omega_{n-1}}\frac{1-\langle x,y\rangle}{\|x-y\|^n} - \frac{\|x\|^2-\langle x,y\rangle}{\omega_{n-1}\|x-y\|^n} \\ &= \frac{1-\|x\|^2}{\omega_{n-1}\|x-y\|^n} \end{aligned}$$

Thus the *Poisson formula* $P(x,y) = \frac{1}{\omega_{n-1}}\frac{1-\|x\|^2}{\|x-y\|^n}$. (This is also valid for $n=2$)

5.6. **The Dirichlet Problem for the Unit Ball with non-smooth data.** The preceding deduction requires the data to be smooth and so we now verify the formula directly in the case where the data is continuous.

Let $P(x,y) = \frac{1}{\omega_{n-1}}\frac{1-\|x\|^2}{\|x-y\|^n}$ be the Poisson kernel in $\mathbb{R}^n$ for $\|x\| < 1$ and $\|y\| = 1$, where ω_{n-1} = area of unit sphere in $\mathbb{R}^n$. Note that if $z = x - y$

$$\begin{aligned} P(x,y) &= -(2\langle y,z\rangle\|z\|^{-n} + \|z\|^{2-n})/\omega_{n-1} \\ &= -(2\sum\frac{y_i z_i}{\|z\|^n} + \frac{1}{\|z\|^{n-2}})/\omega_{n-1} \end{aligned}$$

Thus, for fixed y, $P(x,y)$ is a harmonic function in $\|x\| < 1$ since $\frac{1}{\|x\|^{n-2}}$ and its derivation $\frac{\partial}{\partial x_i}(\frac{1}{\|x\|^{n-2}}) = \frac{-(n-2)x_i}{\|x\|^n}$ are harmonic in $x \neq 0$.

The constant function 1 is the unique solution to the Dirichlet problem with boundary value 1 so we conclude from the Poisson formula in the case we have proven it that

$$\int_{\|w\|=1} P(x,w)dw = 1. \tag{5.7}$$

(One could also prove this directly thus avoiding the reliance on the previously proved case.)

Proposition 11. *Poisson's formula can be used to solve for any f continuous on the unit sphere.*

Proof. Near any point in the interior, the integrand is smooth in x and so we can commute the integration with the Laplacian. As $P(x, y)$ is harmonic in x we conclude that

$$F(x) = \int P(x, y) f(y) dy$$

is too. It therefore suffices to show that $F(ry) \to f(y)$ uniformly in y as $r \uparrow 1$. But, using (5.7)

$$\begin{aligned} |F(ry) - f(y)| &\leq \int P(ry, w)|f(w) - f(y)|dw \\ &\leq \sup_{\|y-w\| \leq \epsilon} |f(w) - f(y)| + 2 \sup |f| \int_{\|w-y\| \geq \epsilon} P(ry, w) dw \end{aligned}$$

The first term is small because f is uniformly continuous on S. To estimate the second time we estimate the denominator by

$$\|w - ry\| \geq \|w - y\| - \|y - ry\| \geq \epsilon - (1 - r) \geq \frac{\epsilon}{2},$$

for $r > 1 - \frac{\epsilon}{2}$. □

Remark 4. *A simple rescaling argument shows that if f is harmonic in $\|x\| < r$ and continuous in $\|x\| \leq r$ then f can be recovered from its boundary values by*

$$f(x) = \int_{\|y\|=r} P_r(x, y) f(y) dy$$

where $P_r(x, y) = \frac{1}{r\omega_{n-1}} \frac{r^2 - \|x\|^2}{\|x-y\|^n}$. There is an analogous result for balls not centred at the origin.

We now know that the Dirichlet problem for the unit ball has a unique solution for continuous data on the unit ball.

An immediate consequence is

Proposition 12. *If u is continuous and satisfies the Mean Value Property for small balls then u is harmonic.*

Proof. On each small ball, we can pick a v agreeing with u on the boundary and harmonic. Applying the strong maximum principle to $u - v$ we deduce that $u = v$ so u is harmonic too. □

Corollary 7. *If f_n is a sequence of harmonic functions which converge uniformly to a function f then f is harmonic.*

Proof. The mean value property is preserved by uniform convergence. □

5.7. **Harnack's Convergence Theorem.** We start by proving

Theorem 21. *Harnack's Inequality: Let f be a non-negative continuous function on $\|x\| \leq R$ harmonic for $\|x\| < R$.*

Then if $\|x\| = r < R$ we have

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

Proof. Poisson's kernel $P(x,y) = \frac{1}{R\sigma^{n-1}} \frac{R^2 - \|x\|^2}{\|x-y\|^n}$. If $\|y\| = R$, $\|x\| = r$ we have

$$R - r \leq \|x - y\| \leq R + r$$

So

$$\frac{1}{R\sigma^{n-1}} \frac{R^2 - r^2}{(R+r)^n} \leq P(x,y) \leq \frac{1}{R\sigma^{n-1}} \frac{R^2 - r^2}{(R-r)^n}$$

Hence, since $f(x) = \int_{\|y\|=R} P(x,y) f(y) dy$

$$\frac{1}{R\sigma^{n-1}} \frac{R^2 - r^2}{(R+r)^n} \int_{\|y\|=R} f(y)dy \leq f(x) \leq \frac{1}{R\sigma^{n-1}} \frac{R^2 - r^2}{(R-r)^n} \int_{\|y\|=R} f(y)dy$$

But, since f is harmonic, $f(0) = \frac{1}{\sigma_{n-1} R^{n-1}} \int_{\|y\|=R} f(y)dy$ So

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

□

Theorem 22. *Harnack's Convergence Theorem: If $u_n \to u$ pointwise and monotonically, with u_n harmonic, then $u_n \to u$ uniformly on bounded sets and u is harmonic.*

Proof. For $\|x\| \leq p < R$

$$0 \leq u_n(x) - u_m(x) \leq c(u_n(0) - u_m(0)) \quad n \geq m.$$

So $\sup_{\|x\| \leq p} |u_n(x) - u_m(x)| \to 0$. So $u_n \to u$ uniformly on $\|x\| \leq p$, all $p > 0$. Hence u is harmonic from the Mean Value Property. □

5.8. **Solution of the Dirichlet Problem by Perron's Method.** Let Ω be a bounded open domain in $\mathbb{R}^n$ with boundary $\partial\Omega$. Consider the Dirichlet Problem $\Delta f = 0$ in Ω, $f = h$ on $\partial\Omega$. Let $u \in C(\overline{\Omega})$ be any sub-harmonic function, ie $u(a) \leq \frac{1}{A} \int\limits_{\|x-a\|=\rho} u(x)$ then if $u \leq f$ on

$\partial\Omega$ we have $u \leq f$ in Ω by the maximum principle. So the solution should be given by Perron's formula

$$f_h(x) = \sup_{\substack{u \text{ sub-harmonic} \\ u \leq h \text{ on } \partial\Omega}} u(x) \tag{5.8}$$

We now show that this yields a solution.

Theorem 23. *Let Ω be a bounded open domain in $\mathbb{R}^n$ with the property that each point $x \in \partial\Omega$ has a* barrier function, *i.e. a sub-harmonic function g such that $g(x) = 0$ and $g(y) < 0$ for $y \in \overline{\Omega}\backslash\{x\}$ then (5.8) solves the Dirichlet problem.*

Remark 5. *$x \in \partial\Omega$ has a barrier function if there is a ball, centred at a, touching $\overline{\Omega}$ only at x. For then $E(y-a) - E(x-a)$ (or its negative) will provide a barrier function.*

Lemma 11. *The maximum of a finite number of subharmonic functions is subharmonic.*

Proof. If $v = \max\{v_1, v_2\}$ then $\int v \geq \int v_1, \int v_2$ over any set and so $v(x) \leq \frac{1}{A}\int v$ on small spheres. □

Lemma 12. *If u is sub-harmonic in Ω and $x \in \Omega$ is such that $B_\delta(x) \subset \Omega$ then there exists a subharmonic function v which is harmonic on $B_\delta(x) \subset \Omega$ and is equal to u outside $B_\delta(x)$. We also have $u \leq v$.*

Proof. We obtain v by solving the Dirichlet problem on $B_\delta x$ with data equal to u on the boundary. We have that $v \geq u$ from the maximum principle. It is trivial that v is sub-harmonic. □

Proof of theorem. Let $\mathcal{E}$ be the set of subharmonic functions which are less than h on the boundary of Ω.

First we note that any element of $\mathcal{E}$ will be bounded by the maximum of h on the boundary and so the supremum does exist.

We first show that f_h is continuous in Ω. If $x \in \Omega$ and $x_n \to x$ then let

$$u_j^n(x_n) \to f_h(x_n)$$

with $u_j^n \in \mathcal{E}$.

We let v_k be the maximum of u_j^n with $j, n \leq k$ then v_k is still in $\mathcal{E}$ and

$$v_k(x_n) \to f_h(x_n),$$

for all n. This all remains true if we replace v_k by w_k with w_k harmonic in a ball about x. We then have w_k converges uniformly to a function

w which is harmonic (and so continuous) on the ball and equal to f_h on the sequence x_n. So

$$f_h(x_n) \to f_h(x)$$

and u is continuous.

We next show f_h is sub-harmonic. Fix a point x. Given $\epsilon > 0$, there exists $u \in \mathcal{E}$ such that $u(x) > f_h(x) - \epsilon$. We then replace u by v with v harmonic in a ball of radius δ with δ fixed and independent of ϵ. We then have

$$f_h(x) < v(x) + \epsilon \leq \epsilon + \frac{1}{A} \int\limits_{\|y-x\|=\delta'} v(y)dy \leq \epsilon + \frac{1}{A} \int\limits_{\|y-x\|=\delta'} f_h(y)dy.$$

This is for any $\epsilon > 0$ and any $\delta' < 1$ so we conclude that f_h is sub-harmonic.

It remains to show that f_h is continuous on the closure of Ω and that it has the right boundary values. Let y be a boundary point. Take g sub-harmonic with $g(y) = 0$ and $g < 0$ on $\overline{\Omega}\backslash\{y\}$. Consider

$$u(x) = h(y) - \epsilon + Kg(x) \quad \text{where } \epsilon > 0$$

Then u is sub-harmonic if $K > 0$. Clearly $u \leq h$ near y on $\partial\Omega$ for $\|z - y\| \leq r$ independent of K. For $\partial\Omega$ $\|z - y\| \geq r$ we can choose K so large that $u \leq h$ everywhere. Hence $h(y) - \epsilon + Kg(x) \leq f_h(x)$. So,

$$\liminf_{x \to y} f_h(x) \geq h(y).$$

That is f is lower semi-continuous.

We need to establish the corresponding estimate from above. We put

$$f_{-h}(x) = \sup u(x),$$

with u subharmonic and less than or equal to $-h$ on the boundary. So $-f_{-h}$, is the infimum of U with $-U$ sub-harmonic and $U \geq h$ on the boundary. (the infimum of superharmonic functions bigger than h on the boundary.) So if $u \in \mathcal{E}$ then $u - U \leq 0$ on the boundary and hence everywhere for any such U. So,

$$f_h \leq -f_{-h}.$$

The same argument as above shows the lower semi-continuity of f_{-h} and so the upper semi-continuity of $-f_{-h}$ and f_h.

We thus have that $f_h \in \mathcal{E}$.

We now use the mean value property to see that f_h is actually harmonic. If it does not satisfy the mean value property at a point p we can replace by a function v which is harmonic in a small ball around

p and is bigger than f_h. But $v \in \mathcal{E}$ so must be smaller than f_h so we have a contradiction. So f_h is harmonic in Ω. $\square$

5.9. **Remarks on the two dimensional Dirichlet Problem.**

5.9.1. *Cauchy-Riemann equations.* Note that $(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y})(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}) = \Delta$. So if E is a fundamental solution of $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ then $(\frac{\partial}{\partial x} \pm i\frac{\partial}{\partial y})E$ is a fundamental solution of the Cauchy-Riemann operator $\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}$. (Recall that f is holomorphic $\Leftrightarrow \frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} = 0$). This gives

$$(\frac{\partial}{\partial x} \pm i\frac{\partial}{\partial y})\frac{1}{4\pi}\log(x^2+y^2) = \frac{1}{4\pi}\frac{2(x\pm iy)}{x^2+y^2} = \frac{1}{2\pi}\frac{1}{x\mp iy}.$$

5.9.2. *Separation of Variables.* The above analysis of the dual Dirichlet problem extends to $n=2$, but we omit it. There is another interesting method involving *separation of variables.*

Recall that in polar coordinates $\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}$. If we look for a solution of the form

$$u(r,\theta) = f(r)\Phi(\theta)$$

and separate variables, we are led to the equations

$$\frac{\partial^2\Phi}{\partial\theta^2} + \lambda\Phi = 0 \tag{5.9}$$

$$r^2\frac{\partial^2 f}{\partial r^2} + r\frac{\partial f}{\partial r} - \lambda f = 0. \tag{5.10}$$

Since Φ has to be 2π periodic, there are solutions only when $\lambda = n^2$ and these are $\cos(n\theta), \sin(n\theta)$. For such λ, the solutions of (5.10) are r^{-n}, r^n and since we want a smooth solution we discard the first one. This leads us to solutions of the form

$$r^n(a_n\cos(n\theta) + b_n\sin(n\theta))$$

which of course tend to $a_n\cos(n\theta) + b_n\sin(n\theta)$ as $r \to 1$. So now given an f on the boundary, we can expand it as a Fourier series

$$a_0 + \sum a_n\cos(n\theta) + b_n\sin(n\theta)$$

and then we have a putative solution

$$a_o + \sum r^n\left(a_n\cos(n\theta) + b_n\sin(n\theta)\right).$$

This will converge to a harmonic function in the interior provided there is a uniform bound for a_n, b_n. It will be continuous up to the boundary giving the right boundary value provided $\sum |a_n| + |b_n|$ converges - this will happen, for example, if f is piecewise smooth. For a general f, one can always compute the coefficients and see if the sum converges!

Remark 6. *One could deduce the Poisson formula in the plane from the Fourier expansion - see for example Petrovsky p173.*

6. The Wave Operator

6.1. **The Problems.** The wave operator or d'Alembertian is the archetypal hyperbolic operator and exists very different behaviour from the Laplacian reflecting the fact that it describes a system evolving in time rather than a steady state system:

$$\Box = \frac{\partial^2}{\partial t^2} - \Delta,$$

where $\Delta = \sum \frac{\partial^2}{\partial x_i^2}$.

We principally want to solve two problems. The first is the forcing problem

$$\Box u = f, \tag{6.1}$$

$$u = 0, \; t << 0 \tag{6.2}$$

$$f = 0, \; t << 0 \tag{6.3}$$

and the second is the Cauchy problem

$$\Box u = 0, \tag{6.4}$$

$$u_{|t=0} = u_0, \tag{6.5}$$

$$\frac{\partial u}{\partial t} = u_1 \tag{6.6}$$

with u_0, u_1 given functions. We will actually solve the second problem by reducing it to the first one.

One way in which the wave operator is very different from the Laplacian is that it is not elliptic and we shall see that it can have singular solutions.

6.2. **Finding the Fundamental Solution.** We have already shown that two fundamental solutions exist and are given by

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

with ϵ a positive number. However we would like to have more explicit knowledge of what they are - that is we need to invert the Fourier transform. First, we invert in t. Note that

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-at-it\tau} dt = \frac{1}{\sqrt{2\pi}} \frac{1}{i\tau + a}$$

if $\Re a > 0$. So the inverse Fourier transform of $\frac{1}{i\tau+a}$ is $\sqrt{2\pi}e^{-at}H(t)$. Now,

$$\langle E_+, f\rangle = \left(\frac{1}{2\pi}\right)^{\frac{n+1}{2}} \iint \frac{\hat{f}(-\tau+i\epsilon,-\xi)}{||\xi||^2-(\tau-i\epsilon)^2}d\tau d\xi \tag{6.7}$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n+1}{2}} \iint \frac{\widehat{e^{\epsilon t}f}(-\tau,-\xi)}{||\xi||^2-(\tau-i\epsilon)^2}d\tau d\xi \tag{6.8}$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n+1}{2}} \iint \frac{\widehat{e^{\epsilon t}f}(-\tau,-\xi)}{2i||\xi||}\left(\frac{1}{i\tau+\epsilon-i||\xi||}-\frac{1}{i\tau+\epsilon+i||\xi||}\right)d\tau d\xi. \tag{6.9}$$

(using partial fractions.) So writing, $f_t(x) = f(t,x)$ and using the fact that the Fourier transform is equal to the inverse Fourier transform up to a reflection, we have

$$\langle E_+, f\rangle = \left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n}\int_0^\infty \frac{e^{\epsilon t}\hat{f}_t(-\xi)}{2i||\xi||}\left(e^{-(\epsilon-i||\xi||)t}-e^{-(\epsilon+i||\xi||)t}\right)dtd\xi,$$

the $e^{\epsilon t}$ terms cancel and we have

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

So our next task is to compute the inverse Fourier transforms of these distributions. Note that as the functions are independent of rotation in $\|\xi\|$ so are the inverse Fourier transforms. The computation of this Fourier transform turns out to be very dimension dependent but there is a technique for calculating the Fourier transform in $n-1$ from the one in n dimensions. So we start with $n = 3$.

So we compute the inverse Fourier transform of $\sin(t\|\xi\|)/\|\xi\|$ in $\mathbb{R}^3$. Actually computing it is hard so we start from the answer and check it works. Let u be defined by

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

This is a compactly supported distribution and

$$\begin{aligned}\langle \hat{u}, f\rangle &= \langle u, \hat{f}\rangle,\\ &= \int_{\|x\|=t}\int e^{-ix.\xi}f(\xi)d\xi dx\\ &= \int\int_{\|x\|=t} e^{-ix.\xi}dxf(\xi)d\xi.\end{aligned}$$

So we have to compute

$$\int\limits_{\|x\|=t} e^{-ix.\xi}dx$$

and by rotational invariance it is enough to take $\xi = (|\xi|, 0, 0)$. So taking spherical polar coordinates θ, ϕ this is equal to

$$t^2 \int\limits_0^{\pi} \int\limits_0^{2\pi} e^{-i\cos(\theta)|\xi|t} \sin(\theta)\, d\phi d\theta.$$

Integrating in ϕ gives us a factor of 2π and integrating in θ we get

$$2\pi t^2 \left[\frac{1}{-i|\xi|t} e^{-i\cos(\theta)|\xi|t}\right]_0^{\pi}$$

which is equal to

$$\frac{2\pi t}{-i|\xi|t}\left(e^{it|\xi|} - e^{-it|\xi|}\right) = \frac{4\pi t}{|\xi|}\sin(|\xi|t).$$

So we deduce that the forward fundamental solution in three (space) dimensions is

$$\langle E_{3,+}, f\rangle = \int\limits_0^{\infty} \frac{1}{4\pi t} \int\limits_{\|x\|=t} f(x,t)dxdt.$$

Rescaling, we can rewrite this as

$$\langle E_{3,+}, f\rangle = \int\limits_0^{\infty} \frac{t}{4\pi} \int\limits_{\|x\|=1} f(xt,t)dxdt.$$

Put $y = xt$ regarding (x,t) as polar coordinates, we also have

$$\langle E_{3,+}, f\rangle = \frac{1}{4\pi}\int\limits_{\mathbb{R}^3} \frac{f(\|x\|, x)}{\|x\|}dx.$$

This is the *forward* fundamental solution as it is supported in $t \geq 0$. If we had taken $\epsilon < 0$, we would have got the backwards solution, supported in $t \leq 0$. Note that the forward fundamental solution in three space dimension is supported on $t = \|x\|$ - this reflects Huygens principle which is that waves travel precisely at speed 1 and will therefore only be detectable at precisely distance t at time t.

We also want to compute the fundamental solution in two dimensions. Now we can regard $\frac{\sin(t|\xi|)}{|\xi|}$ in two dimensions as the restriction of the corresponding function in three dimensions to the hyperplane $\xi_3 = 0$. Restricting a Fourier transform of a Schwartz function to ξ_3 is equivalent to integrating the function with respect to x_3. The equivalent statement for distributions is that provided the push-forward in x_3 exists, its Fourier transform is the restriction of the higher dimensional Fourier transform to $\xi_3 = 0$. Now as $E_{3.+}$ is supported in $\|x\| \leq t$ the push-forward is proper and is defined via pull-backs so we deduce that

$$\langle E_{2,+}, \phi(t, x_1, x_2)\rangle = \langle E_{3,+}, \phi(t, x_1, x_2)\rangle.$$

So we need to compute what

$$\int\limits_{\|x\|=t} \psi(x_1, x_2)$$

is. But we can regard x_1, x_2 as parametrising the sphere by

$$x_3 = \pm(t^2 - x_1^2 - x_2^2)^{\frac{1}{2}}.$$

And so if we cut into the two hemispheres we obtain,

$$\int\limits_{\|x_1,x_2\|\leq t} \frac{\psi(x_1, x_2)}{(t^2 - \|x\|^2)^{\frac{1}{2}}} dx_1 dx_2$$

and thus that

$$\langle E_{2,+}, f\rangle = \frac{1}{2\pi} \int\limits_0^\infty \int_{\|x_1,x_2\|\leq t} \frac{(t, x)}{(t^2 - \|x\|^2)^{\frac{1}{2}}} dx dt.$$

Now in one dimension there is a much easier approach, if one performs a change of coordinates $w = t + x, y = t - x$ then the wave operator becomes

$$c\frac{\partial^2}{\partial w \partial y}$$

which has the fundamental solution

$$H(w)H(y)$$

and so we deduce that (taking care with constants)

$$\langle E_{1,+}, f\rangle = \frac{1}{2} \int\limits_0^\infty \int\limits_{|x|<t} f(x, t) dx dt.$$

In fact, in all dimensions the forward fundamental solution is supported inside the light cone and in any odd dimension bigger than 1 it is supported on the light cone. These support conditions guarantee that these are the only fundamental solutions supported in $t \geq 0$ as if L is another such fundamental solution, the convolution $K * L$ exists and so

$$E_+ * (\Box\delta_0 * L) = (E_+ * \Box\delta_0) * L$$

and thus $E_+ = L$.

Alternatively,

$$\Box(E_+ * L) = (\Box E_+) * L = E_+(\Box L).$$

Reversing the sign of t, we obtain E_- the backwards fundamental solution and the Feynmann fundamental solution is just the average of the two

$$\frac{1}{2}(E_+ + E_-).$$

6.3. **The Method of Descent.** Our argument above to evaluate the Fourier transforms is really the *method of descent* which is that if K is a fundamental solution for the wave operator in n dimensions then pushing forward in x_n will give the fundamental solution in one lower dimension provided the push-forward exists. We compute

$$\begin{aligned}\langle \Box_{n-1}\pi_* K_n, \phi\rangle &= \langle \pi_* K_n, \Box_{n-1}\phi\rangle \\ &= \langle K_n, \pi^*\Box_n \phi\rangle\end{aligned}$$

as $\pi^*\phi$ is independent of x_n

$$\begin{aligned}&= \langle K_n, \Box_n \pi^*\phi\rangle \\ &= \langle \pi_*\Box_n K_n, \phi\rangle \\ &= \langle \pi_*\delta_0, \phi\rangle \\ &= \langle \delta_0, \phi\rangle.\end{aligned}$$

6.4. **Solving the forcing problem.** So we now have the forward fundamental solution, we can solve the forcing problem

$$\Box u = f,$$

with $u, f = 0$ in $t << 0$ simply by

$$u = K_+ * f,$$

even for distributional f - the convolution will exist because K_+ is supported inside the light cone. So, *given the vanishing in the past*, u will be smooth if and only if f is and u is uniquely determined by f. However, if we drop the past vanishing condition life is more complicated: if u is a twice-differentiable function on the real line and $|\omega| = 1$ then

$$\Box u(t - x.\omega) = 0.$$

(in fact, for any distribution this is true) So, there is no version of elliptic regularity and solutions are not unique.

Using the backwards solution, a similar approach works for data vanishing in the future. So given any $f \in \mathcal{D}'$ we can solve

$$\Box u = f$$

by putting

$$f_1 = \psi(t)f, f_2 = f - f_1$$

with ψ identically one in $t > 1$ and zero in $t < -1$. The solution will not be unique though.

6.5. **The Cauchy Problem.** The uncertainty in solutions of the wave operator is expressible in terms of Cauchy data - there is a unique solution of

$$\Box u = f$$

with

$$u_{t=t_0} = u_0,\ \frac{\partial u}{\partial t}_{|t=t_1} = u_1$$

for given $u_0, u_1 \in C^\infty$. (one could consider more general data) We can solve the forcing problem so after subtracting a solution of it, it is enough to consider the case when $f = 0$.

Now as

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

we deduce that

$$\frac{\partial^2 u}{\partial t^2}_{t=t_0} = \Delta u_0$$

and iterating we can recover all the derivatives i.e. the Taylor series of u on $t = t_0$. So we certainly have not specified too little data. One approach is to then sum the Taylor series, however there is no reason this should converge except on $t = t_0$. (if the data were analytic then it would)

However the sum does exist in a more generalised sense:

Theorem 24. *The Borel Lemma - given any sequence of smooth functions $f_j(x)$ there exists a function $f(t, x)$ such that*

$$f^{(j)}(0, x) = f_j(x).$$

So using the Borel lemma, one can pick a function u such that $\Box u$ vanishes to infinite order on $t = t_1$ and $u, \frac{\partial u}{\partial t}$ have the right value on $t = t_0$. Now if $\Box u = f$ then $f_1 = H(t - t_0)f$ and $f_2 = H(t_0 - t)f$ are smooth functions because of the infinite order vanishing. So, we deduce that

$$u' = u - K_+ * f_1 - K_- * f_2$$

is a solution as $K_+ * f_1$ is supported in $t \geq t_0$ and so vanishes to infinite order on $t = 0$ and similarly for f_2 using K_-.

So there is a solution. We can use the same technique to show uniqueness - if u_1, u_2 solve then let $u = u_1 - u_2$, we have that u and its first derivative in t vanish on $t = t_0$ and thus arguing as above that all its derivatives in t vanish on $t = t_0$. This means

$$u = u_+ + u_-$$

with $u_\pm$ supported in $\pm(t - t_0) \geq 0$ and smooth. We then have that

$$\Box u_\pm = 0$$

and thus since the forcing problem in a half space has a unique solution, we conclude that $u_\pm = 0$.

We now prove the Borel lemma:

Proof. The idea is to use cut off functions, cutting off closer and closer to the origin so that the sum is always finite so if ϕ is a bump function we put

$$f(t, x) = \sum_j \frac{t^j}{j!} f_j(x) \phi(e_j t)$$

where $e_j \to +\infty$. As $\phi(e_j t)$ is supported in C/e_j the sum is finite for each t and so converges. We want it to converge uniformly and we also want all its derivatives to converge uniformly. (for x in compact sets)

Let $x \in K$, a fixed compact set, then the supremum of $\frac{t^j}{j!} f_j(x) \phi(e_j t)$ is less than

$$\sup_{x \in K}(|f_j|) \left(\frac{C}{e_j}\right)^j.$$

So picking $e_{0,j}$ sufficiently large this will be less than $1/j^2$ and we have uniform convergence and thus f is continuous.

To get differentiability in x we need

$$\sum \frac{t^j}{j!} D_x^\alpha f_j(x) \phi(e_j t)$$

to converge uniformly. We therefore pick $e_{\alpha,j}$ as above to get uniform convergence.

Of course, we want the same e_j for all α. We do this by diagonalisation. Just let

$$e_j = \max_{|\alpha| \leq j} e_{\alpha,j}$$

then for any α, we eventually have

$$e_{j,\alpha} \leq e_j$$

and the sum

$$f(t,x) = \sum_j \frac{t^j}{j!} f_j(x)\phi(e_j t)$$

converges uniformly as do all its x derivatives to a function which is smooth in x but only continuous in t.

However, we can now play the same game again. If we differentiate in t, k times and take $D^\alpha)_x$, we obtain a sum of terms

$$\sum_{l=0}^{k} \frac{t^{j-l}}{(j-l)!} e_j^{k-l} \phi^{(k-l)}(t) D_x^\alpha f_j(x).$$

Letting $D = \max_{j \leq k} |\phi^{(j)}(t)|$ and taking $j > k$ (which we can do as we are interested in behaviour for large j) this is less than or equal to

$$(C/e_j)^{j-k} D \sup_{x\in K} |D_x^\alpha f_j(x)| e_j^k.$$

As before, picking $e_j = e_{j,\alpha,k}$ sufficiently large this will converge.

So we now do a diagonalisation - thus we let

$$e_j = \max_{l,|\alpha|,k\leq j} e_{l,\alpha,k}.$$

Then for any α, k once j is sufficiently large, we have

$$e_j \geq e_{j,\alpha,k}.$$

So the sum

$$\sum \frac{t^j}{j!} \phi(e_j t) f_j(x)$$

converges uniformly in all derivatives and the result follows.

(note the stage where we did the x derivatives alone was not necessary but has been added as a "warm-up" proof) □

We have specified initial data on a very special hypersurface - $t = t_0$. In fact, we can specify data on a large class of hypersurfaces but not all. There are three basic classifications - depending on the geometry of the normals - space-like, time-like and characteristic. Recall that the symbol of the wave operator is $-(\tau^2 - \xi^2)$.

Definition 3. *A hypersurface H in $\mathbb{R}^n \times \mathbb{R}$ is space-like if $\tau^2 - \xi^2$ is greater than 0 where (τ, ξ) is the normal vector. H is characteristic if $\tau^2 - \xi^2 = 0$ and time-like if $\tau^2 - \xi^2 < 0$.*

We recall that a hypersurface is the zero set of a real-valued, smooth function which has non-zero derivative at every point of its zero set. A typical example is the set $\{t = g(x)\}$ with g smooth and real.

So, a cone is characteristic and the plane $t = t_0$ is space-like. We can, in general, solve the Cauchy problem for space-like hypersurfaces - on a Lorentzian manifold this would be the only notion as there is no special time function.

We proceed similarly to before. The hypersurface can be written in the form

$$H = \{t = \phi(x)\}$$

and so we can take a Taylor series expansion of the solution u in t about $t = \phi(x)$

$$u \sim \sum_{j=0}^{\infty} \frac{(t - \phi(x))^j}{j!} f_j(x).$$

And so computing $\Box u$ has the Taylor expansion

$$\sum_{j=2}^{\infty} \frac{(t-\phi(x))^{j-2}}{(j-2)!} f_j(x) - \sum_{j=1}^{\infty} \frac{(t-\phi(x))^{j-1}}{(j-1)!} (\nabla\phi.\nabla f_j(x) + \Delta\phi f_j) \\ + \sum_{j=0}^{\infty} \frac{(t-\phi(x))^{j}}{j!} \Delta f_j(x) - |\nabla\phi|^2 \sum_{j=2}^{\infty} \frac{(t-\phi(x))^{j-2}}{(j-2)!} f_j(x).$$

Now if we equate coefficients of $(t - \phi(x))^j$ to zero, we get for $j > 1$ that $(1 - |\nabla\phi|^2)f_j$ is equal to a function of the lower order coefficients. We can solve this for f_j as long as $(1 - |\nabla\phi|^2)$ in non-zero, that is as long as the hypersurface is non-characteristic. So then applying the Borel lemma, we have a function u with the correct Cauchy data such that $\Box u$ vanishes to infinite order on H.

As before, we can break up $\Box u$ into two pieces supported, $f_\pm$, on either side of H and then apply the two fundamental solutions to the two pieces. This will work as before except that we need to know that the $K_\pm * f_\pm$ is still supported on one side of H - this will happen provided the surface is space-like. The point being that the support of $K_+ * f_+$ is contained in the union of forward cones with tips on H - the fact that the normal points inside the cone means that the tangent plane to H (which is precisely the orthogonal to the normal to H) will meet the cone at the tip only. (this only works locally , wave hands to get globally.)

6.6. Domains of Dependence. Suppose for f, u_0, u_1, smooth we know

$$\begin{aligned}\Box u =& f,\\ u_{|t=0} =& u_0,\\ \frac{\partial u}{\partial t}_{t=0} =& u_1,\end{aligned}$$

has a unique solution which is smooth. If we just want to know the value of f at a point (t, x) where do we need to know f, u_0, u_1? For $t > 0$ we know from our construction of the solution that the value of f in $t < 0$ is irrelevant. So consider $u_+ = H(t)u$ then

$$\Box u_+ = \frac{\partial^2}{\partial t^2}(H(t)u) - \Delta u.$$

So we compute that

$$\langle \Box u_+, \psi \rangle = \int_0^\infty \int u(x) \frac{\partial^2 \psi}{\partial t^2} dx dt - \langle u_+, \Delta \phi \rangle$$

which on integrating by parts is equal to

$$-\int u(x,0)\frac{\partial \psi}{\partial t}(x,0)dx + \int \frac{\partial u}{\partial t}(x,0)\psi(x,0)dx + \int_0^\infty \int \frac{\partial^2 u}{\partial t^2}(x,t)\psi(x,t)dxdt.$$

Now we have that $\frac{\partial u}{\partial t}(x, 0) = u_1$ and $u(x, 0) = u_0$ and in $t > 0$ that $u(x, t) = u_+(x, t)$ which is equal to $K_+ * \Box u_+$ so we have a formula for $u(x, t)$ in terms of the data. We can rewrite this as

$$u_+ = K_+ * (H(t)f) + K_+ * (\delta(t)u_1) - K_+ * (\delta'(t)u_0).$$

The first term evaluated at (x, t) is just a weighted integral of f on the set $|y - x| \leq t - s$ in dimensions $1, 2$ and on the set $|y - x| = t - s$ in dimension 3. Now we can write

$$\langle K_+, \phi \rangle = \int_0^\infty \langle K_t, \phi_t \rangle$$

where $\langle K_t, \phi_t \rangle$ is a weighted integral over $|y| <\leq t$ or $|y| = t$. Unravelling the convolutions, we deduce

$$u(x,t) = E_t * u_1(x) + \frac{\partial}{\partial t} E_t * u_0 + \int_0^t E_{t-s} * f_s ds$$

where $*$ is now convolution in x. This is called Kirchhoff's formula.

An immediate consequence of this, in all dimensions, is that the value of $u(x, t)$ is determined by the value of $f(y, s)$ on and in the cone $|x - y| \leq t - s$ and by $u_0(y), u_1(y)$ on the set $|x - y| \leq t$. This expresses the property of finite speed of propagation - one need only know the data within distance t of the point at time t.

In three dimensions, E_+ is supported on the cone so we have that the value at $u(x,t)$ is determined by f on the surface of the cone $|x-y| = t-s$ and by u_0, u_1 on $|x-y| = t$. This expresses Huygen's principle in three dimensions that light travels exactly at one speed with no back-wash.

6.7. **Energy Estimates.** If $\Box u = 0$ then integrating over all space in x and over $[0,T]$ in t we have

$$\begin{aligned} 0 &= \iint u_t(u_{tt} - u_{xx})dxdt \\ &= \iint (u_t u_{tt} + u_{tx}u_x)dxdt \\ &= \iint \frac{1}{2}(u_t^2 + u_x^2)_t dxdt \\ &= \left[\frac{1}{2}\int (u_t^2 + u_x^2)dx\right]_0^T \end{aligned}$$

which proves the statement. This is of limited utility as it does not apply if the functions are not compactly supported however there is a local version which also gives a proof of finite speed propagation of information which does work. (to get this integrate over a cone $|x - x_0| < R - t, 0 \le t \le T$.)

The local version is that the energy in the $|x-x_0| < R$ at time T is less than or equal to the energy in the ball $|x - x_0| < R+T$ at time T. This expresses the idea that the total amount of energy is preserved and that it can move around at a maximum rate of 1.

7. The Heat Equation

7.1. **Symmetries.** The heat operator is given by

$$\frac{\partial}{\partial t} - \Delta$$

and so it invariant under isometries of $\mathbb{R}^n$ in x and translations in t. It is *not* invariant under reflections in t unlike the wave equation.

Note that

$$\left(\frac{\partial}{\partial t} - \Delta\right)(u(ax, a^2t)) = a^2\left(\left(\frac{\partial}{\partial t} - \Delta\right)u\right)(ax, a^2t)$$

reflecting the different homogeneities of the operator in x and t. We will find that the factor $|x|^2/t$ will appear often in the study of this operator.

7.2. **The Fundamental Solution.** As usual we start by constructing the fundamental solution; we know that a fundamental solution is given by

$$\langle E, \phi \rangle = \left(\frac{1}{2\pi}\right)^{n+1} \int \frac{\hat{\phi}(-\tau + i\epsilon, -\xi)}{\|\xi\|^2 + i(\tau - i\epsilon)} d\tau d\xi$$

where $\epsilon > 0$. Of course, we want to compute what this is on the non-Fourier transform side.

$$\begin{aligned}\langle E, \phi \rangle &= \left(\frac{1}{2\pi}\right)^{n+1} \iint \frac{\widehat{e^{\epsilon t}\phi}(-\tau, -\xi)}{\|\xi\|^2 + \epsilon + i\tau} d\tau d\xi \\ &= \left(\frac{1}{2\pi}\right)^{n+1} \int_0^\infty \int e^{\epsilon t} \hat{\phi}_t(-\xi) e^{-(\|\xi\|^2+\epsilon)t} dt d\xi \\ &= \left(\frac{1}{2\pi}\right)^{n+1} \int_0^\infty \int e^{\epsilon t} \hat{\phi}_t(\xi) e^{-(\|\xi\|^2+\epsilon)t} dt d\xi \\ &= \left(\frac{1}{2\pi}\right)^{n} \int_0^\infty \int \phi_t(x) \frac{e^{-\frac{\|x\|^2}{4t}}}{(2t)^{n/2}} dx dt.\end{aligned}$$

So the fundamental solution is

$$K(t,x) = \begin{cases} \left(\frac{1}{4\pi t}\right)^{n/2} e^{-\frac{\|x\|^2}{4t}} & t \geq 0 \\ 0 & t < 0 \end{cases}.$$

Note that we do not get a corresponding fundamental solution if we reverse the sign of t - the heat equation expresses a definite direction of time.

The only singularity of the fundamental solution of the heat operator is at

$$x = 0, t = 0.$$

We can therefore construct a parametrix for the heat operator supported in any small ball B_ϵ and so the arguments in the elliptic case applied here imply

Proposition 13. *The singularities of u and $\left(\frac{\partial}{\partial t} - \Delta\right) u$ are the same for any $u \in \mathcal{D}'$.*

This property is sometimes called *hypo-ellipticity* - expressing that while the operator is not elliptic it behaves similarly in this regard.

The standard problem we want to solve for the heat operator is

$$\left(\frac{\partial}{\partial t} - \Delta\right) u = g, \quad t > 0$$
$$\lim_{t \to 0+} u(x,t) = u_0$$

with u_0, g given. Writing

$$K_t(x) = K(t,x)$$

the solution of this is given by

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

The second term here is just the application of the forward fundamental solution to g - this will make sense for any g continuous and bounded on $t = 0$ as the exponential decay of K_{t-s} will ensure that the convolution converges. The fact that the integral of $K(t,x)$ with respect to x is 1 for any t tells us that $K_{t-s} * g_s$ is bounded by $\sup g$ and so the ds integral will converge too. This also shows that

$$\int_0^t K_{t-s} * g_s ds \to 0$$

at $t \to 0+$.

This reduces us to studying the problem for $g = 0$. We check directly that $K_t * u_0$ solves. That it is a solution of the heat equation is clear as

$$\left(\frac{\partial}{\partial t} - \Delta\right) K_t = 0.$$

It is also smooth as K_t is. We check that it is continuous up to $t = 0$ with the correct boundary value. Let $\epsilon > 0$, fix $z \in \mathbb{R}^n$ and take δ such that $|x - z| < \delta$ implies that

$$|u_0(x) - u_0(z)| < \epsilon$$

then

$$
\begin{aligned}
|K_t * u_0(x) - u_0(z)| &= \left|\int K(x-y,t)(u_0(y)-u_0(z))dy\right| \\
&\leq \int_{|y-x|<\delta/2} K(x-y,t)|u_0(y)-u_0(z)|dy \\
&\quad + \int_{|y-x|>\delta/2} K(x-y,t)|u_0(y)-u_0(z)|dy \\
&\leq \int_{|y-z|<\delta} K(x-y,t)|u_0(y)-u_0(z)|dy \\
&\quad + 2\sup|g| \int_{|y-z|>\delta/2} K(x-y,t)dy \\
&\leq \epsilon \int K(x-y,t)dy \\
&\quad + 2\sup|g| \int_{|y-z|>\delta/2} K(x-y,t)dy.
\end{aligned}
$$

The first term equals ϵ so we need only check that for any fixed δ that

$$\int_{|x|>\delta} K(x,t) \to 0$$

as $t \to 0+$. Performing the change of variables $\tilde{x} = t^{-\frac{1}{2}}x$ the integral is equal to

$$C_n \int_{|\tilde{x}|>\delta t^{-\frac{1}{2}}} e^{-\tilde{x}^2/4} d\tilde{x}$$

and the result follows. So we have the existence of a solution and one could deduce from the representation that there is continuous dependence on initial data. We still need to check uniqueness - in fact this only holds if we make constraints on the solution and we look at this in the next section using maximum principles which also give a method of proving continuous dependence on initial data.

If we study the homogeneous problem with $g = 0$ and we take u_0 to be a positive compactly supported smooth function then as $K_t > 0$, we have that the solution $u(t,x) > 0$ for all $t > 0$, and all x so heat propagates at infinite speed - this is in contrast to the wave equation where waves propagate at finite speed. This does not appear very physical! But the solution is decaying exponentially fast so these extra effects are very small.

7.3. **Maximum Principles.** As usual, we wish to have some control over the solution - it should be unique and depend continuously on in initial data. Surprisingly, there are smooth functions u which satisfy the heat equation and tend to zero at $t = 0+$ which are not identically zero. We therefore need to make some restrictions on our class of solutions to avoid this. One approach is to make the requirement that u be positive which is sufficient to guarantee uniqueness and is the physical case. The analysis however is rather long so instead we use a boundedness condition to get a maximum principle.

First, we consider a bounded domain,

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

Proposition 14. *If* $(\frac{\partial}{\partial t} - \Delta)u \leq 0$ *in* Ω *and* u *is continuous on* $\bar{\Omega}$ *then the maximum value of* u *in* Ω *is attained on* $\{|x-y| \leq r, t = 0\} \cup \{|x-y| = r, 0 \leq t \leq T\}$.

Proof. First suppose that $(\frac{\partial}{\partial t} - \Delta)u < 0$ in Ω. We let

$$\Omega_\epsilon = \{|x-y| < r, 0 < t < T - \epsilon\}.$$

Suppose that u has a local maximum inside Ω_ϵ then at such a point, we have

$$u_t = 0, \Delta u \geq 0$$

which contradicts our supposition so u must attain its maximum on the boundary of Ω_ϵ. Now if the maximum is attained at a point in $\{|x-y| < r, t = T-\epsilon\}$ then at such a point we have $u_t \geq 0$ and $\Delta u \leq 0$ which contradicts the assumption. So we have that the maximum is attained on $\{|x-y| \leq r, t = 0\} \cup \{|x-y| = r, 0 \leq t \leq T - \epsilon$. Taking a union over all positive ϵ the result follows in this case.

One can reduce the general case to the already proved case by subtracting kt from u, applying the result and then letting $k \to 0$. □

Uniqueness and continuous dependence on initial data is now immediate if we consider the problem of specifying data on the set where the maximum is attained.

A general maximum principle with x unbounded also applies, *provided* we make an assumption on the function's growth.

Theorem 25. *Let* u *be continuous on* $\Omega = \mathbb{R}^n \times [0, T)$ *and smooth in the interior with*

$$\left(\frac{\partial}{\partial t} - \Delta\right) u \leq 0$$

and with u bounded above by M and

$$u(x,0) = f(x)$$

then for $(x,t) \in \Omega$,

$$u(x,t) \leq \sup u(x,0)$$

in Ω.

Proof. We start by fixing a point $(y,s) \in \Omega$, with $s > 0$ where we shall estimate the value of u. As the problem is translation invariant in y we shall take $y = 0$.

Let

$$v_\mu(x,t) = u(x,t) - \mu(x^2 + 2nt)$$

then applying the maximum principle to the cylinder $|x| \leq \rho, 0 \leq t \leq T$ we have

$$v_\mu(x,t) \leq \max\left\{ \sup_{|x| \leq \rho (v_\mu(x,0))}, \sup_{|x|=\rho, t\in[0,T]} (v_\mu(x,0)) \right\}.$$

Now,

$$\sup_{|x|=\rho, t \in [0,T]} (v_\mu(x,0)) \leq M - \mu\rho^2$$

and

$$\sup_{|x|\leq\rho} (v_\mu(x,0)) \leq \sup(v_\mu(x,0)) \leq \sup u(x,0).$$

So we have

$$u(0,s) - \mu s \leq \max\{\sup u(x,0), M - \mu\rho^2\}.$$

So given ϵ we pick μ so that $\mu s < \epsilon$ and then ρ such that $M - \mu\rho^2 \leq \sup u(x,0)$ and conclude that

$$u(0,s) \leq \epsilon + \sup u(x,0)$$

and as ϵ is arbitrary the result follows. □

We remark that our condition on u is unnecessarily stringent and that this result could be proved under the assumption that

$$|u(x,t)| \leq Ce^{a||x||^2}.$$

(see for example [2].)

Uniqueness in the class of bounded solutions and continuous dependence on initial data are now clear. (use the maximum principle for u and $-u$.)

7.4. **Group Law.** If we are given initial data u_0 at time 0 for the heat equation, flow for time s to get u_s and then using this as initial data flow of time t to get $(u_s)_t$, then this this will be equal to u_{s+t} from the uniqueness of solutions. As the problem is translation invariant in time, if write the solution operator at time t as $e^{\Delta t}$ this says that

$$e^{\Delta(s+t)} = e^{\Delta s}e^{\Delta t} \quad s,t > 0$$

and

$$e^{\Delta t} \to \mathrm{Id},\ t \to 0+ .$$

This says that $e^{\Delta t}$ is an operator semi-group and there is a large theory of such semi-groups.

One can do something similar with the wave equation - working with pairs (u_0, u_1) rather than with u_0. As we can flow backwards in time, we get a group rather than a semi-group.

7.5. **Arrow of time.** We have only solved the forward problem for the heat operator - we specify initial data at time 0 and then compute what the solution is in positive time. Unlike the wave equation, the backwards problem is not solvable for the heat equation - one can not recover the initial value of a distribution from its future behaviour. For example, whatever initial data we start with, we always have a smooth function in all positive time.

7.6. **Brownian Motion.** One can regard the heat kernel $K(x,t)$ for a fixed t as a probability density function as it is positive and has integral equal to 1. This reflects the fact that the heat equation can be interpreted as a limit coming from Brownian motion. If we have a particle at y as time 0 then it will be distributed at time t according to the density function $K(x-y,t)$. So if one imagines there being a particle of heat at a point y at $t=0$ then it is smeared out according to the density K at time t - this is a rather antiquated notion from the point of view of physics but can useful mathematically.

7.7. **Finite Difference Methods.** In this section, we look at how the one dimensional heat equation can be realized as a limit of finite difference equations. This gives a numerical method of computing the solution. It should be noted, although we shall not do this here, that this approach actually gives a method of proving the existence of solutions for variable coefficient operators by showing that the approximations converge to a solution which is not *a priori* given.

What is the finite difference analogue of the heat equation? We assume $h, k > 0$ are given and fixed (Later, we will vary them.) and work on a lattice

$$\Sigma_{h,k} = \{(lh, mk), l, m \in \mathbb{Z}, l \geq 0\}.$$

If u is a function on the lattice, we define

$$\Lambda u = \frac{u(x, t+h) - u(x, t)}{h} - \frac{u(x+k, t) - 2u(x, t) + u(x-k, t)}{k^2}.$$

Note that this will converge to the heat operator as $k, h \to 0+$ (use L'Hopital's rule.)

We can regard this as a recipe for computing $u(x, t+h)$ from $u(y, t)$; if

$$\Lambda u = d$$

then

$$u(x, t+h) = u(x, t) + \frac{h}{k^2}(u(x+k, t) - 2u(x, t) + u(x-k, t)) + hd.$$

If we let $\|.\|$ be the supremum norm in x then we have that

$$\|u(., t+h) \leq |1 - 2hk^{-2}|\|u(., t)\| + 2hk^{-2}\|u(., t)\| + h\|d(., t)\|.$$

If we assume that assume that $\left|\frac{h}{k^2}\right| \leq \frac{1}{2}$ then we have

$$\|u(., t+h)\| \leq \|u(., t)\| + h\|d(., t)\|.$$

So if u has initial data f on $t = 0$ then iterating, we deduce that

$$\|u(., lh)\| \leq \|f\| + lh\|d(., t)\|$$

or putting $t = lh$ that

$$\|u(., t)\| \leq \|f\| + t \sup_{s \leq t} \|d(., s)\|. \tag{7.1}$$

Now suppose we have a solution u of

$$\left(\frac{\partial}{\partial t} - \Delta\right) u = d(x, t) \tag{7.2}$$

$$u(0, t) = f_0. \tag{7.3}$$

We want to compare it to the solutions obtained from the finite difference methods by iteration. Suppose we have a nested sequence of lattices Σ_ν given as above with hk^{-2} fixed independent of ν and $h, k \to 0$ as $\nu \to \infty$. We have then for each ν a solution u_ν. Putting U equal to the union of Σ_ν over all ν, we expect the limiting values of $u_\nu(x, t)$ to converge to $u(x,t)$ for $(x,t) \in U$. We prove this subject to a regularity assumption on u.

Theorem 26. *Suppose u is a solution of* (7.2),(7.3) *and* $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial x}$, Δu *are uniformly bounded and uniformly continuous then for all $(x,t) \in U$ we have that*

$$u_\nu(x,t) \to u(x,t).$$

Proof. The uniform continuity and boundedness guarantee that

$$\frac{u(x,t+h) - u(x,t)}{h} - \frac{\partial u}{\partial t}$$

and

$$\frac{u(x+k,t) - 2u(x,t) + v(x-k,t)}{k^2} - \Delta u$$

tend to zero uniformly as $h, k \to 0$ that is as $\nu \to \infty$. (use L'Hopital's rule) Given $\epsilon > 0$, we have for ν sufficiently big we that

$$|\Lambda_\nu u - d| < \epsilon.$$

We therefore have that

$$|\Lambda_\nu (u - u_\nu)| < \epsilon$$

with initial data zero. So from our estimate (7.1), we have

$$|u(x,t) - u_\nu(x,t)| < \epsilon t,$$

which proves the result. □

8. Appendix

8.1. **Integration.** For convenience, we have used some theorems from Lebesgue integration. Here we just run over the notions and theorems we need.

The basic space is $L^1(\mathbb{R}^n)$ - a function f will be in $L^1(\mathbb{R}^n)$ if both it and its modulus are Riemann integrable and the integrals converge at infinity. Much more general functions are in the space but will not be needed in this course.

We say a function is locally integrable if it is integrable when multiplied by the characteristic function of any ball. This space is denoted by L^1_{loc}.

The main theorem we will use is (a weakened version of) the dominated convergence theorem.

Theorem 27. *Let h_n be a sequence of functions in L^1 such that there exists $g \in L^1$ with $|h_n| \leq g$ and such that h_n converges to a function h pointwise then h is in L^1 and*

$$\lim_{n\to\infty} \int h_n = \int h.$$

8.2. **Taylor's Theorem.** While a smooth function does not in general have a convergent power series, it can be approximated accurately by a finite series to whatever order we like.

Theorem 28. *(i) Let $f(t,x)$ be a smooth function then for all $N \in \mathbb{N}$*

$$f(t,x) = \sum_{j=0}^{N-1} \frac{f^{(j)}(0,x)}{j!} t^j + \int_0^t \frac{(t-s)^{N-1}}{(N-1)!} f^{(N)}(s,x) ds.$$

[Note that the error is a smooth function vanishing to order N at $t = 0$.]

(ii) Under the conditions of (i), if $f^{(j)}(0,x) = 0$ for $0 \leq j \leq N-1$ then

$$f(t,x) = t^n \phi(t,x)$$

with $\phi(t,x)$ a smooth function.

Proof. (i) (This is just the proof you know and love dressed up a bit to make a change.) By the fundamental theorem of calculus,

$$f(t,x) = f(0,x) + \int_0^t f^{(1)}(s,x) ds.$$

If we iterate this we get,

$$f(t,x) = \sum_{j=0}^{N-1} \frac{f^{(j)}(0,x)}{j!} s^j + \int_0^t \int_0^{s_1} \cdots \int_0^{s_{N-1}} f^{(N)}(s_N) ds_N ds_{N-1} \ldots ds_1.$$

The result then follows by reversing the order of integration.

(ii) By part (i)

$$f(t,x) = \int_0^t \frac{(t-s)^{N-1}}{(N-1)!} f^{(N)}(s,x) ds,$$

so making the linear change of variable $s = tu$ we have

$$f(t,x) = t^n \phi(t,x)$$

with

$$\phi(t,x) = \int_0^1 \frac{(1-u)^{N-1}}{(N-1)!} f^{(N)}(ut,x)du.$$

Standard theorems about differentiating under the integral sign show that ϕ is smooth. □

It's important to remember that there are function whose Taylor series vanishes to infinite order at $t = 0$ but are positive for $t \neq 0$. For example $e^{-t^{-2}}$.

References

[1] F.G. Friedlander, Introduction to the Theory of Distributions, *CUP 1982, QA324*

[2] F. John, Partial Differential Equations, *Springer-Verlag 1982, QA1.A647*

[3] I.G. Petrovksy, Lectures on Partial Differential Equations, *Dover 91, QA377.P433*

[4] M.E. Taylor, Partial Differential Equations,*Springer 96, QA1.A647*