# Lectures on PDEs- APMA- E6301
# DRAFT IN PROGRESS

Michael I. Weinstein*

April 28, 2008

## Contents



*Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027; miw2103@columbia.edu

# 1 First order partial differential equations and the method of characteristics

Consider the first order PDE

$$F(x, u(x), Du(x)) \ = \ 0, \quad Du(x) = (\partial_{x_1} u, \dots . \partial_{x_n} u)\,. \tag{1.1}$$

Here, $F(x, z, p)$ denotes a smooth real-valued function of $2n + 1$ variables $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ and $z \in \mathbb{R}^1$, with $x = (x_1, \dots, x_n)$, $\quad p = (p_1, \dots, p_n)$.
**Transport equation:** $\partial_t u(x, t) + v \cdot u(x, t) \ = \ 0, \quad x \in \mathbb{R}^n, \quad t \in \mathbb{R}^1$.
**Linear first order equations:** $b(x) \cdot Du(x) + c(x)u(x) = 0$
**Quasilinear first order equations:** $b(x, u(x)) \cdot Du(x) + c(x, u(x)) = 0$
**Fully nonlinear equations:** equations which are nonlinear in $Du(x)$, *e.g.* The eikonal equation, $|\nabla u(x)|^2 = 1$.

# 2 The Laplacian - Laplace's and Poisson's equations

## 2.1 The Newtonian potential in $\mathbb{R}^n$

## 2.2 Boundary value problems and Green's functions

### 2.2.1 Green's function for the ball in $\mathbb{R}^n$

### 2.2.2 Green's function for the upper half plane in $\mathbb{R}^2$, $\mathbb{R}^n$

## 2.3 Single and Double Layer Potentials and their boundary limits

## 2.4 Boundary integral formulation of the Dirichlet and Neumann Problems – Interior and Exterior

## 2.5 Fredholm Operators in Banach Spaces, Solvability of Fredholm Integral Equations on $C^0(S)$ and the Fredholm Alternative

## 2.6 Application of Fredholm Integral Operator Theory to the solution of the Dirichlet and Neumann Problems

# 3 Distributions and Fundamental Solutions

[1]

**Definition 3.1** *Let $C_0^\infty(\Omega)$ or $C_c^\infty(\Omega)$ denote the set of compactly supported $C^\infty$ functions.*

$$C_0^\infty(\Omega) \;=\; \{ \; f \;:\; \text{there exists } K \text{ compact subset of } \Omega,\; f(x) \equiv 0,\; x \in \Omega - K \; \}$$

*One often denotes $C_0^\infty(\Omega)$ by $\mathcal{D}(\Omega)$ or $\mathcal{D}$* [2]*. Functions in $\mathcal{D}$ are often called test functions.*

**Definition 3.2** *A linear functional, $T$, on $\mathcal{D}$ is linear mapping which associates to any $\phi \in \mathcal{D}$ a number $T[\phi]$, i.e.*

$$T[\;\lambda\phi + \mu\psi\;] \;=\; \lambda\, T[\phi] \;+\; \mu\, T[\psi]$$

*for any $\lambda, \mu$ scalars and any $\phi, \psi \in \mathcal{D}$.*

**Examples:** (a) *Distributions obtained from locally integrable functions:* Let $f \in L^1_{loc}$ be any *locally integrable function*[3]. Define

$$T_f[\phi] \;=\; \int_{\mathbb{R}^n} \phi(x) f(x)\; dx \tag{3.1}$$

Then, $T_f$ is a linear functional on $\mathcal{D}$. For simplicity, we will often write $f[\phi]$ instead of $T_f[\phi]$.
(b) *The Dirac "delta function":*

$$\delta_\xi[\phi] \;=\; \phi(\xi) \tag{3.2}$$

defines a linear functional on $\mathcal{D}$.

**Continuity of linear functionals on $\mathcal{D}$, notion of distribution**

**Definition 3.3** *Consider an arbitrary sequence of functions $\{\phi_k\}$ in $\mathcal{D}$, which tend to zero in the sense of the following two conditions: (a) $\phi_k$ vanish outside a fixed compact subset, $K$, of $\Omega$. and (b) For any multi-indx $\alpha$, $\lim_{k\to\infty} \partial^\alpha \phi_k(x) = 0$ uniformly in $x \in K$.*
*A linear functional $T$ is continuous if $\lim_{k\to\infty}\; T[\phi_k] \;=\; 0$ for any sequence $\{\phi_k\}$, which tends to zero in the above sense. A continuous linear functional on $\mathcal{D}$ is called a distribution. The set of distributions on $\mathcal{D}(\Omega)$ is denoted by $\mathcal{D}'(\Omega)$.*

**Examples:** $T_f$, defined for any locally integrable $f$ and $\delta_\xi$ are examples of distributions.

**Differentiation of distributions:**

[1]References: Chapter 3, F. John PDEs; Section E of Chapter 0 in G.B. Folland Introduction to PDEs
[2]Here, $\Omega$ is taken to be an open set. A compact set is a set which is a set which is closed and bounded.
[3]$f \in L^1_{loc}$ means that for any compact set $C$, $\int_C |f(x)| dx < \infty$.

**Notation:** A *multi-index* is a vector of non-negative integers $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}_0^m$. The *order* of a multi-index is given by $|\alpha| \equiv \alpha_1 + \cdots + \alpha_m$.

$$\partial^\alpha f(x) = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_m}^{\alpha_m} f(x_1, \ldots, x_m).$$

Note that if $f, \phi \in \mathcal{D}$, then

$$\int \partial^\alpha f(x)\phi(x)\ dx \ = \ (-1)^{|\alpha|} \int\ f(x)\ \partial^\alpha \phi(x)\ dx, \tag{3.3}$$

which we can write as

$$\partial^\alpha\ f[\phi] \ = \ (-1)^{|\alpha|}\ f[\ \partial^\alpha \phi\ ]. \tag{3.4}$$

Equation (3.4) motivates our definition of the derivative of a distribution. Let $T \in \mathcal{D}'$. Then,

$$\partial^\alpha T[\phi] \ = \ (-1)^{|\alpha|}\ T[\ \partial^\alpha \phi\ ] \tag{3.5}$$

### Distribution solutions to PDEs

Let $L$ denote the linear partial differential operator $L = \sum_{\alpha \in \mathbf{I}}\ a_\alpha(x)\partial^\alpha$. Suppose we have a PDE

$$Lu = T, \tag{3.6}$$

where $T \in \mathcal{D}'$. Given the above definition of derivative of a distribution, we say that the PDE has a distribution solution $u$, provided

$$u[L^t \phi] \ = \ T[\phi], \quad \phi \in \mathcal{D}. \tag{3.7}$$

Here, $L^t\ \phi \ = \ \sum_{\alpha \in \mathbf{I}} (-1)^{|\alpha|}\partial^\alpha\ (\ a_\alpha(x)\ \phi(x)\ )$ denotes the *adjoint* of the differential operator $L$, which satisfies the generalization of (3.3): for $\phi, \psi \in \mathcal{D}$

$$\int\ L\psi(x)\ \phi(x)\ dx \ = \ \int\ \psi(x)\ L^t\phi(x)\ dx$$

### Fundamental Solutions

**Definition 3.4** *A fundamental solution of a linear differential operator $L$ with pole at $x$ is a distribution solution, $u$, of $Lu = \delta_x$.*

**Examples: Laplace Operator** $L \ = \ \Delta$. $L^t \ = \ \Delta$. We have shown that a fundamental solution for $\Delta$ with pole at $x$ is:

$$\begin{aligned} \Phi(x-y) \ &= \ \frac{1}{2\pi}\ \log\ |x-y|, \quad n = 2 \\ &= \ \frac{1}{(2-n)\omega_n}\ |x-y|^{2-n}, \quad n \geq 3 \end{aligned}$$

**Exercise:** In one space dimension, solve

$$\frac{d^2\Phi}{dx^2} \ = \ \delta_x$$

where $\delta_x$ denotes the Dirac delta function (distribution).

# 4 Introduction to Hilbert Space

**Definition 4.1** [4][5]

*(a)* *S is a vector space or linear space:* $u, v \in S$ *implies, for all* $\lambda, \mu$ *scalars that* $\lambda u + \mu v \in S$*, plus vector spaces axioms.*

*(b)* *S is an inner product space if it is a linear space and there is a function*

$$\begin{aligned} (\cdot,\cdot) : S \times S \to \mathbb{R} \\ u, v \in S \ \mapsto \ (u, v) \in \mathbb{R} \end{aligned} \tag{4.1}$$

*with the following properties*

*(P1)* $(u, v) = (v, u)$

*(P2)* $(\lambda u + \mu v, w) \ = \ \lambda(u, v) \ + \ \mu(v, w), \ \lambda, \mu \in \mathbb{R}.$

*(P3)* $(u, u) \geq 0$ *and* $(u, u) = 0 \implies u = 0.$

**Example 1a:** $S = \mathbb{R}^n$. $u, v \in \mathbb{R}^n$. $(u, v) \ = \ \sum_{i=1}^n \ u_i v_i$.
**Example 2a:** $S = C^\infty(\Omega)$, $\Omega \subset \mathbb{R}^n$ bounded open, $(f, g) \ = \ \int_\Omega u(x) \ v(x) \ dx$.

**Definition 4.2** *A normed linear space is a vector space equipped with a notion of length, called the norm,* $\|\cdot\|$*, satisfying the following properties*

*(n1)* $\|u\| = 0$ *implies* $u = 0$

*(n2)* $\|\lambda u\| = |\lambda| \ \|u\|, \ \ \lambda \in \mathbb{R}$

*(n3)* $\|u + v\| \leq \|u\| + \|v\|$

Given an inner product space, $S$, there is a natural *norm*, or measure of size of a vector in $S$

$$\|u\| \ = \ \sqrt{(u, u)} \tag{4.2}$$

[4]References: We seek to give the required working knowledge of the basic functional analysis, *e.g.* Hilbert space, linear operators, spectral theory, ..., with a view toward its application to PDEs. See, for example, (a) L.C. Evans, *PDEs*, Appendix D on Linear Functional Analysis and Appendix E on Measure and Integration; (b) F. John, *PDEs*, Springer 4th Ed., Chapter 4, section 5; G.B. Folland, *Introduction to PDEs*, Princeton University Press, Chapter 0. A good text on elementary notions in analysis (convergence, uniform convergence, integration,... is the text: W. Rudin, *Principles of Mathematical Analysis*, McGraw Hill. a more advanced text is that of E. Lieb and M. Loss, *Analysis*, Am. Math. Soc.

[5]Throughout we will construct spaces of functions, with notions of size and distance within these spaces defined in terms of integrals. These integrals are assumed throughout to be taken in the *Lebesgue* sense, an important extension notion of Riemann integral. Note that if $f$ is integrable in the sense of Riemann, it is in the sense of Lebesque and the values of these integrals is the same. Appendix E of L.C. Evans PDEs contains a very terse treatment of measure theory and the Lebesgue integral.

Any norm satisfies the *Cauchy Schwarz* inequality

$$|(u,v)| \;\leq\; \|u\|\,\|v\| \tag{4.3}$$

**Example 1b:** $S = \mathbb{R}^n$. $u, v \in \mathbb{R}^n$. $\|u\| \;=\; \sqrt{\sum_{k=1} |u_k|^2}$.

**Example 2b:** $S = C^\infty(\Omega)$, $\Omega \subset \mathbb{R}^n$ bounded open. $\|u\|_2 \;=\; \left( \int_\Omega |u(x)|^2 dx \right)^{\frac{1}{2}}$.

**Convergence:**
The sequence $\{u^j\}_{j\geq 1}$ in $S$ converges to a point $u_* \in S$ provided

$$\lim_{j\to\infty} \|u_j - u_*\| = 0. \tag{4.4}$$

In this case we say that the sequence $\{u^j\}_{j\geq 1}$ converges *strongly to* $u_*$ or *converges to* $u_*$ *in norm.* Later, we shall introduce the notion of *weak convergence.*

**Cauchy sequence:**
The sequence $\{u^j\}_{j\geq 1}$ in $S$ is called a *Cauchy sequence* if $\|u^j - u^k\| \to 0$ as $j, k \to \infty$.

**Completeness:**
The normed linear space $S$ is complete if every Cauchy sequence in $S$ converges to some element of $S$. A complete normed linear space is called a *Banach space.* A complete normed linear space, whose norm derives from an inner product, as in (4.2), is called a *Hilbert space.*

**Completion Theorem:** Every normed linear space can be completed. Specifically, if $S$ is a normed linear space, then it is possible to naturally extend $S$ to a possibly larger set, $\tilde{S}$, and the norm $\|\cdot\|$ to be defined on all $\tilde{S}$, in such a way that the set $\tilde{S}$, equipped with the norm $\|\cdot\|$. is complete. The completion is constructed by identifying all Cauchy sequences with points in the extended space $\tilde{S}$. If $v \in \tilde{S}$ but $v \notin S$, then there is a sequence $\{v_k\}$ in $S$, such that $\|v_k - v\| \to 0$ as $k \to \infty$. We say that the space $S$ is *dense* in $\tilde{S}$.

**Example 1c:** Let $\mathbb{Q}$ denote the set of rational numbers, numbers of the form $p/q$, where $p$ and $q$ are integers and $q \neq 0$. Consider the inner product space $S = \mathbb{Q}^n$. $u, v \in \mathbb{Q}^n$. $\|u\| \;=\; \sqrt{\sum_{k=1} |u_k|^2}$, the Euclidean norm. While $\mathbb{Q}^n$ equipped with this norm is not complete, $\mathbb{Q}^n$ has a completion, namely $\mathbb{R}^n$ equipped with the norm $\|u\|$.

**Example 2c:** $C^\infty(\Omega)$ is not complete with respect to the norm $\|u\|_2 \;=\; \int_\Omega |u(x)|^2 dx$. There are many important completions of $C^\infty(\Omega)$ with respect to different norms:

(1) **The Hilbert Space** $L^2(\Omega)$: The completion of $C^\infty(\Omega)$ with respect to the norm $\|\cdot\|_2$ is called $L^2(\Omega)$. Since the norm on $L^2$ derives from an inner product, $f, g \mapsto \int_\Omega fg$, $L^2$ is a Hilbert space.

(2) **Sobolev Spaces - $H^k(\Omega)$, $k \geq 0$:** Define the inner product of two functions $f, g \in H^1$ by

$$(f,g)_{H^1} \ = \ \int_\Omega \nabla f \cdot \nabla g \ + \ fg \ dx \tag{4.5}$$

The $H^1$ Sobolev norm is defined by:

$$\|u\|^2_{H^1} \ = \ (u,u)_{H^1} \ = \ \int_\Omega |\nabla u(x)|^2 \ + \ |u(x)|^2 \ dx \tag{4.6}$$

More generally, for any integer $k \geq 0$, define $H^k(\Omega)$ to be the completion of $C^\infty(\Omega)$ with respect to the norm

$$\|u\|^2_{H^k} \ = \ \sum_{|\alpha|\leq k} \int_\Omega |\partial^\alpha u(x)|^2 \ dx \tag{4.7}$$

Note that $H^0(\Omega) = L^2(\Omega)$.

(3) **Sobolev Spaces - $H^k(\mathbb{R}^n)$, $k \geq 0$:** Replace the bounded open set, $\Omega$, with $\Omega = \mathbb{R}^n$. First consider $C_0^\infty(\mathbb{R}^n)$ and define $H^k(\mathbb{R}^n)$ to be the completion of $C_0^\infty(\mathbb{R}^n)$ with respect to the $H^k(\mathbb{R}^n)$ norm:

$$\|u\|^2_{H^k} \ = \ \sum_{|\alpha|\leq k} \int_{\mathbb{R}^n} |\partial^\alpha u(x)|^2 \ dx \tag{4.8}$$

**Remark 4.1** *Sobolev spaces are important in analysis of functions largely due to their role in relating global (integral) properties and local (pointwise) properties. For example, here is the most elementary*

*Sobolev inequality:* There is a constant $K > 0$ such that for any $f \in C_0^\infty(\mathbb{R})$ we have

$$|f(x)|^2 \ \leq \ K \left( \int |f(x)|^2 \ + \ |f'(x)|^2 \ dx \right) \ = \ K \ \|f\|^2_{H^1} \tag{4.9}$$

*Proof:* **Exercise**

**Weak convergence:** A sequence $\{u^j\}_{j\geq 1}$ in $\mathcal{H}$ converges to a limit $u_*$ in $\mathcal{H}$ if for all $w \in \mathcal{H}$

$$(u^j, w)_{\mathcal{H}} \ \to \ (u_*, w)_{\mathcal{H}} \tag{4.10}$$

**Compactness** A closed subset, $C$, of $\mathcal{H}$ is *compact* if any sequence $\{u^j\}_{j\geq 1}$ in $C$ has a convergent subsequence.

**Remark 4.2** *Any closed and bounded subset of $\mathbb{C}^n$ is compact. However, in infinite dimensional spaces, e.g. $L^2$, this is not the case. See section 12.4.*

## 4.1 Representation of Linear Functionals on $\mathcal{H}$

(1) A *linear functional on* $S$ is a mapping $\phi : S \to \mathbb{R}, \quad u \mapsto \phi(u)$, such that for all $u, v \in S$ and all scalars $\lambda, \mu$ we have

$$\phi(\lambda u \ + \ \mu v) \ = \ \lambda\phi(u) \ + \ \mu\phi(v).$$

One sometimes represents the action of $\phi$ on a vector $u$ with notation

$$\phi(u) = \langle \phi, u \rangle.$$

(2) A *bounded linear functional* is a linear functional for which there is a constant, $M > 0$, such that for any $u \in S$

$$|\phi(u)| \ \le \ M \ \|u\| \tag{4.11}$$

The set of bounded linear functionals on a Hilbert Space, $\mathcal{H}$, is denoted $\mathcal{H}^*$ and is called the *dual space* of $\mathcal{H}$. The smallest constant, $M$, for which the inequality (4.11) holds is called the norm of the linear functional $\phi$, $\|\phi\|_{\mathcal{H}^*}$.

**Example 1d:** For any vector $v \in \mathbb{R}^n$, $\phi(u) = (u, v) = u \cdot v = \sum_{k=1}^n u_k v_k$ is a bounded linear functional with norm $\|\phi\|_{(\mathbb{R}^n)^*} = \|v\|_{\mathbb{R}^n}$. Moreover, it is straightforward to see that given *any* linear functional on $\mathbb{R}^n$, $\phi$, there is a unique vector $v_\phi \in \mathbb{R}^n$ such that $\phi(u) = (u, v_\phi)$ for all $u \in \mathbb{R}^n$.

**What about bounded linear functionals on Hilbert space?** Let $\mathcal{H}$ denote any Hilbert space and let $v$ be a fixed vector in $\mathcal{H}$. Then, in analogy with the construction in **Example 1d**, $\phi(\cdot) = (\cdot, v)$ defines a bounded linear functional on $\mathcal{H}$ with norm $\|v\|$. The joke is that *every* bounded linear functional on a Hilbert space is obtained in this way:

**Theorem 4.1 Riesz Representation Theorem:** *Let $\mathcal{H}$ denote a Hilbert space and $\mathcal{H}^*$ the set of all bounded linear functionals on $\mathcal{H}$. For any $\phi \in \mathcal{H}^*$, there is a unique vector $v_\phi \in \mathcal{H}$, such that*

$$\phi(u) \ = \ (u, v_\phi), \quad \text{for any} \ \ u \in \mathcal{H}. \tag{4.12}$$

*Furthermore*[6], $\|\phi\|_{\mathcal{H}^*} \ = \ \|v_\phi\|_{\mathcal{H}}$.

The proof of this theorem relies on some elementary geometry of Hilbert space, in particular, the projection theorem.

**Proof of the Riesz Representation Theorem:** If $\phi(u) = 0$ for every $u \in \mathcal{H}$, then we take $v_\phi = 0$. Thus we suppose that $\phi(q) \neq 0$ for some $q \in \mathcal{H}$. By linearity, $\phi(q \ \phi(q)^{-1}) = 1$. The proof can be broken down into a sequence of claims.

*Claim 1: Let $m = \inf\{ \ \|u\| \ : \ \phi(u) = 1 \ \}$. The infimum, $m$, is attained, i.e. there exists $w \in \mathcal{H}$ such that $\phi(w) = 1$ and $\|w\| = m$.*

We begin with the identity (exercise), which holds for all $a, b \in \mathcal{H}$:

$$\text{Parallelogram Law :} \quad \frac{1}{4}\|a+b\|^2 + \frac{1}{4}\|a-b\|^2 = \frac{1}{2}\|a\|^2 + \frac{1}{2}\|b\|^2, \tag{4.13}$$

[6] $\mathcal{H}$ and $\mathcal{H}^*$ are isometrically isomorphic

which follows by direct computation, using that $\|a\|^2 = (a,a)$.
Now let $w_k$ denote a minimizing sequence. Thus, $\|w_k\| \downarrow m$, as $k \to \infty$, $\phi(w_k) = 1$. We use (4.13) to show that $\{w_k\}$ is a Cauchy sequence. Set $a = w_k$ and $b = w_k$. Then, (4.13)

$$\frac{1}{4}\|w_k - w_j\|^2 = \frac{1}{2}\|w_k\|^2 + \frac{1}{2}\|w_j\|^2 - \left|\left|\frac{w_k + w_j}{2}\right|\right|^2 \leq \frac{1}{2}\|w_k\|^2 + \frac{1}{2}\|w_j\|^2 - m^2$$

The latter inequality holds because $\phi(\frac{w_k+w_j}{2}) = 1$, and therefore $\left|\left|\frac{w_k+w_j}{2}\right|\right| \geq m$. Now let $k, j$ tend to infinity and we have, since $\|w_{k,j}\| \downarrow m$, that $\|w_k - w_j\| \to 0$, *i.e.* the sequence $\{w_k\}$ is Cauchy. It therefore has a limit $w \in \mathcal{H}$.
*Claim 2: The minimizer, $w$, is unique.*

Again this follows from the parallelogram law (4.13). If $w_A$ and $w_B$ both belong to $\mathcal{H}$ with $\|w_A\| = \|w_B\| = m$, we have $\left|\left|\frac{w_A+w_B}{2}\right|\right|^2 \geq m^2$. Substitution into (4.13) gives:

$$m^2 \ + \ \frac{1}{4}\|w_A - w_B\|^2 \ \leq \ \left|\left|\frac{w_A + w_B}{2}\right|\right|^2 \ + \ \frac{1}{4}\|w_A - w_B\|^2 \ = \ m^2.$$

This gives a contradiction.

*Claim 3: $w$ is orthogonal to the null space of $\phi$, i.e. for all $u$ such that $\phi(u) = 0$, we have $(w,u) = 0$.*

If $\phi(u) = 0$ then

$$\phi\left( \ w \ - \ \frac{(w,u)}{(u.u)}u \ \right) \ = \ \phi(w) \ = \ 1 \tag{4.14}$$

Since $w$ has minimal norm,

$$\left|\left| \ w - \frac{(w,u)}{(u,u)}u \ \right|\right|^2 \ \geq \ \|w\|^2.$$

Expanding and cancelling terms gives $|(w,u)|^2 \leq 0$ and therefore $(w,u) = 0$.

*Claim 4: $w \neq 0$.*

$1 \ = \ \phi(w_k) \ \to \ \phi(w)$ by continuity of $\phi$. If $w = 0$ then $\phi(w) = 0$, a contradiction.

*Claim 5 and Completion of the Proof:*
*Define $v_\phi = \|w\|^{-2}w$. Then, for all $u \in \mathcal{H}$, $\phi(u) = (u, v_\phi)$.*

Let $u \in \mathcal{H}$ be arbitrary. Recall that by Claims 3 and 4, $w \neq 0$ and $\{x : \phi(x) = 0\} \perp w$. Now express $u$ as the sum of two orthogonal components, one in the direction of $w$ and one lying in the null space of $\phi$.

$$\begin{aligned} u \ &= \ [ \ u \ - \ w\phi(u) \ ] \ + \ w\phi(u) \\ (u,w) \ &= \ (u \ - \ w\phi(u) \ , w) \ + \ \phi(u) \ (w,w) \end{aligned}$$

Note: $(u \ - \ w\phi(u) \ , w) = 0$ since $\phi(u \ - \ w\phi(u)) = 0$. Therefore,

$$(u,w) \ = \ \phi(u) \ (w,w) \ \iff \ (u, v_\phi) \ = \ \phi(u), \ v_\phi \equiv \frac{w}{\|w\|^2}$$

This completes the proof of the Riesz Representation Theorem.

# 5 Applications of Hilbert Space Theory to the Existence of Solutions of Poisson's Equation

We begin with Poisson's equation for a function $u$ defined on an open bounded subset $\Omega$ of $\mathbb{R}^n$:

$$-\Delta v \ = \ f, \quad v\,|_{\partial\Omega} = 0 \tag{5.1}$$

Here, we begin by taking $f \in C(\bar{\Omega})$.

We seek a *weak formulation* of (5.1). That is, we seek a formulation which is satisfied by classical solutions, *i.e.* $u$ of class $C^2(\bar{\Omega})$ satisfying (5.1). To derive a weak formulation, let $u \in C_0^\infty(\Omega)$ and suppose that $v$ is a classical solution. Now multiply (5.1) by $u$ and integrate over $\Omega$. Integration by parts, using the Gauss' divergence theorem and that $u$ vanishes on $\partial\Omega$ gives

$$\begin{aligned} \int_\Omega uf \ dx \ = \int_\Omega u\ (-\Delta v) \ &= \ - \int_\Omega \nabla \cdot (u \nabla v) \ dx \ + \ \int_\Omega \nabla u \cdot \nabla v \ dx \\ &= \ \int_\Omega \nabla u \cdot \nabla v \ dx. \end{aligned} \tag{5.2}$$

Define

$$\begin{aligned} \phi(u) \ &= \ \int_\Omega u(x) f(x) dx \\ (u,v)_D \ &= \ \int_\Omega \nabla u(x) \cdot \nabla v(x) \ dx \end{aligned}$$

The result (5.2) can be expressed as follows: If $v$ is a classical solution to Poisson's equation with zero (Dirichlet) boundary conditions ($v(x) = 0, \ x \in \partial\Omega$), then for all $u \in C_0^\infty(\Omega)$, we have $\phi(u) \ = \ (u,v)_D$. Note, on the other hand, that $\phi(u)$ is well-defined for any $u, v \in L^2(\Omega)$ and $(u,v)_D$ is well-defined if $\nabla u$ and $\nabla v$ are both in $L^2(\Omega)$. With the Riesz Representation Theorem in mind, it is natural to regard $(u,v)_D$ as defining an inner product with respect to which we shall represent the bounded linear functional $\phi(u)$.
Toward the correct formulation, we define the

**Dirichlet norm** on $C_0^\infty(\Omega)$

$$\|u\|_D \ = \ \sqrt{(u,u)_D}. \tag{5.3}$$

The key to showing that $\|\cdot\|_D$ is indeed a norm is to observe that if $\|u\|_D = 0$, then $u$ is identically constant, and because $u$ vanishes on $\partial\Omega$, we must have $u \equiv 0$.

Next, to put our problem in a Hilbert space setting we define the

**Sobolev space** $H_0^1(\Omega)$, which encodes zero boundary conditions, to be the completion of the space $C_0^\infty(\Omega)$ with respect to the Dirichlet norm, $\|\cdot\|_D$.

We can now formulate the **Weak Dirichlet Problem**:

*Let $f \in L^2(\Omega)$. Find $v \in H_0^1(\Omega)$ such that for all $u \in H_0^1(\Omega)$, we have*

$$(u,v)_D \ = \ \int_\Omega \nabla u(x) \cdot \nabla v(x) \ dx \ = \ \int_\Omega u(x) \ f(x) \ dx \ = \ \phi_f(u). \tag{5.4}$$

**Theorem 5.1** *The weak Dirichlet problem has a unique $H^1_0(\Omega)$ solution.*

**Proof:** We have set things up so that the result follows directly from the Riesz Representation Theorem. The main thing to check is that $\phi(u)$ is a bounded linear functional on $H^1_0(\Omega)$, *i.e.* there exists a constant $M > 0$, such that

$$|\phi(u)| \ \equiv \ \left| \int_\Omega uf \ dx \right| \ \leq M_f \ \|u\|_D, \tag{5.5}$$

for some $M_f$, depending on $f$ but independent of $u$. By the Cauchy Schwarz inequality[7]

$$\left| \int_\Omega uf \ dx \right| \ \leq \ \|f\|_{L^2} \ \|u\|_{L^2}$$

Therefore, to prove that $\phi(u)$ is a bounded linear functional and therewith the theorem, we must prove the:
**Poincaré Inequality:** Let $\Omega$ be an open connected subset in $\mathbb{R}^n$ , which lies between two planes, a distance $2a$ apart[8]. There exists a positive constant, $p_\Omega$, depending on $\Omega$, such that for any $u \in H^1_0(\Omega)$

$$\|u\|_{L^2} \ \leq p_\Omega \ \|u\|_D \tag{5.6}$$

Note that by the Poincaré inequality, $\phi(u)$ is a bounded linear functional on $H^1_0(\Omega)$, with norm bound $M_f = p_\Omega \|f\|_2$.

**Proof of the Poincaré inequality:** After possibly rotating the region $\Omega$, we may assume it lies between the two planes $x_1 = a$ and $x_1 = -a$. Note that it suffices to prove the inequality for $u$ belonging to a dense subset of $H^1_0(\Omega)$, $u \in C_0^\infty(\Omega)$. Let $x = (x_1, x')$, with $x' = (x_2, \ldots, x_n)$. For fixed $u(x)$, we have

$$u(x_1, x') \ = \ \int_{-a}^{x_1} \ \partial_{x_1} u(\xi_1, x') \ d\xi_1$$

Squaring and applying the Cauchy-Schwarz inequality, we have

$$\begin{aligned} u^2(x_1, x') \ &= \ \left( \int_{-a}^{x_1} \ 1 \cdot \partial_{x_1} u(\xi_1, x') \ d\xi_1 \right)^2 \\ &\leq \ \left( \int_{-a}^{x_1} 1^2 \ d\xi_1 \right) \ \cdot \ \left( \int_{-a}^{x_1} \ |\partial_{x_1} u(\xi_1, x')|^2 \ d\xi_1 \ \right) \\ &\leq \ (x_1 + a) \ \int_{-a}^{a} \ |\partial_{x_1} u(\xi_1, x')|^2 \ d\xi_1 \end{aligned}$$

[7] $\int \ |fg| \ \leq \|f\|_2 \ \|g\|_2$
[8] Thus we allow domains, more general than bounded domains.

Integration with respect to $x_1$ gives

$$\int_{-a}^{a} |u(x_1, x')|^2 \ dx_1 \ \leq \ 2a^2 \int_{-a}^{a} |\partial_{x_1} u(\xi_1, x')|^2 \ d\xi_1$$

Finally, integation with respect to the remaining $n-1$ variables, $x'$ gives:

$$\int_\Omega |u(x_1, x')|^2 \ dx_1 \ d\xi_1 dx' \ \leq \ 2a^2 \int_\Omega |\partial_{x_1} u(\xi_1, x')|^2 \ d\xi_1 dx'$$

This proves Poincaré's inequality with $p_\Omega = \sqrt{2}a$.

## 5.1 Generalizations of the Weak Dirichlet Problem

We briefly discussed how the forgoing discussion can be generalized from the Laplacian, $-\Delta$, to a treatment of the Dirichlet problem for for more general linear elliptic partial differential operators, $L$:

$$Lv = f, \quad x \in \Omega, \qquad v\,|_{\partial\Omega} = 0$$

(a) $L = -\nabla \cdot a(x)\nabla$, where $a(x)$ is a smooth function satisfying the *strong ellipticity condition*: $\min_{x\in\bar{\Omega}} a(x) \geq \theta$. Define the Dirichlet inner produce

$$(u, v)_a = \int_\Omega a(x)\nabla u(x) \cdot \nabla v(x) \ dx.$$

We then introduce the Dirichlet norm is then $\|u\|_{\mathcal{H}^1_a}$. Let $\mathcal{H}^1_a$ denote the completion of $C_0^\infty(\Omega)$ with respect to $\|u\|_{\mathcal{H}^1_a}$. Then, we can prove (**Exercise using Poincaré's inequality!**) that $\phi(u) = (u, f)_{L^2}$ is a bounded linear functional on $\mathcal{H}^1_a$. We can then apply the Riesz Representation Theorem to show that there exists a unique $v_\phi \in \mathcal{H}^1_a$, such that for all $u \in \mathcal{H}^1_a$, $\phi(u) \ = \ (u, v_\phi)_a$.

(b) *More general strongly elliptic operators:* $L = -\sum_{jk} \ \partial_{x_j} a_{jk}(x)\partial_{x_k}$, where $A = (a_{jk}(x))$ is a $n$ by $n$ symmetric and positive definite matrix function, whose minimum eigenvalue, $\lambda_{min}(x)$ satisfies $\min_{x\in\bar{\Omega}} \lambda_{min}(x) \geq \theta > 0$, for some $\theta$. Note that example (a) is a special case with $A = \delta_{jk}a(x)$. Introduce the Dirichlet norm: $\|u\|_{\mathcal{H}^1_A}$ and $\mathcal{H}^1_A$, the completion of $C_0^\infty(\Omega)$ with respect to $\|u\|_{\mathcal{H}^1_A}$. Then, (**Exercise using Poincaré's inequality**) that $\phi(u) = (u, f)_{L^2}$ is a bounded linear functional on $\mathcal{H}^1_A$. We can then apply the Riesz Representation Theorem to show that there exists a unique $v_\phi \in \mathcal{H}^1_A$, such that for all $u \in \mathcal{H}^1_A$, $\phi(u) \ = \ (u, v_\phi)_A$.

(c) *Yet more general elliptic operators - elliptic operators with non-symmetric lower order terms:* Garding's inequality and the Lax-Milgram Lemma (representation theorem, which generalizes Riesz Rep Th'm, which is applicable to non-symmetric forms).

## 5.2 Relation between weak and strong (classical) solutions of elliptic PDE

For strongly elliptic linear partial differential operators (see 5.1 one has the following important theorem, showing that "weak implies strong".

**Theorem 5.2** *(Elliptic Regularity Theorem) Consider the Dirichlet problem*

$$Lv(x) = f(x),\ x \in \Omega, \quad v(x) = 0,\ x \in \partial\Omega, \tag{5.7}$$

*where $\Omega$ is an open and bounded subset of $\mathbb{R}^n$ with smooth boundary and $f \in L^2(\Omega)$. Let $v \in H_0^1(\Omega)$ be a weak solution of $Lv = f$, where $f \in C^\infty(\Omega)$. Then, $v \in C^\infty$ and $\lim_{x\to\partial\Omega,\ x\in\Omega} v(x) = 0$.*

The proof is very technical; see L.C. Evans, *PDE*, section 6.3.

## 5.3 Introduction to the Finite Element Method

[9]

The notion of *weak solution* of the Dirichlet problem for Poisson's equation $-\Delta v(x) = f(x),\ v(x) = 0,\ x \in \partial\Omega$ (see (5.4) ) motivates *The Finite Element Method - (FEM)*, a method of central importance in the numerical solution of PDEs. We introduce FEM via a simple example.

Consider the two-point boundary value problem:

$$-\partial_x^2 v(x) \ = \ f(x), \quad v(0) = v(1) = 0. \tag{5.8}$$

Above we saw that this can be formulated, in terms of the **Weak Dirichlet Problem** (5.4), which we present in modified form as follows.

Define

$$\mathcal{U} \ = \ \{u :\ u \in C^0[0,1], u \in C^1_{piecewise}[0,1]\ \}. \tag{5.9}$$

Here, $f \in C^1_{piecewise}[0,1]$ if there are finitely many points $0 < \xi_1 < \cdots < \xi_r < 1$ such that $f \in C^1(\xi_j, \xi_{j+1}),\ j = 1, \ldots, r-1$

Our modified form of the weak Dirichlet problem is:
*Find $u \in \mathcal{U}$ such that for all $v \in \mathcal{U}$*

$$(\partial_x u, \partial_x v) \ = \ (f, v) \tag{5.10}$$

To obtain the Galerkin or Finite Element Method for the (5.8) we will replace $\mathcal{U}$, by a finite-dimensional approximation, $\mathcal{U}_h$, which we now construct.

[9]Reference: *Numerical solution of partial differential equations by the finite element method*, Claes Johnson, Cambridge 1992

Partition the interval $[0,1]$ into subintervals:

$$\begin{aligned} [0,1] &= [x_0,x_1)\cup(x_1,x_2)\cup\cdots\cup(x_{M-1},x_M)\cup(x_M,x_{M+1}] = \cup_{j=1}^{M+1} I_j \\ 0 &= x_0 < x_1 < x_2 < \cdots < x_{M-1} < x_M < x_{M+1} = 1 \end{aligned}$$

The interval lengths in the partition are $h_j = x_{j+1} - x_j$. If the points, $x_j$, are equally spaced than $h_j = h = (M+1)^{-1}$.

**Definition of $\mathcal{U}_h$:**

$$\mathcal{U}_h = \{u: \ u \text{ linear on } I_j, \ j = 1,\ldots,M+1, \ v \text{ continuous}, \ u(0) = u(1) = 0\ \} \tag{5.11}$$

**Proposition:** $\mathcal{U}_h$ has dimension $M$ and has the basis (spanning set which is linearly independent) consisting of the "triangular tent functions" defined by:

$$\begin{aligned} \varphi_j &\in \mathcal{U}_h \\ \varphi_j(x_i) &= \delta_{ij}, (1, i=j; 0, i \neq j) \\ j &= 1,2,\ldots,M \end{aligned}$$

**proof:** Let $u \in \mathcal{U}_h$. Then, $u(x) = \sum_{j=1}^M \eta_j \varphi_j(x)$, where $\eta_j = u(x_j)$. Therefore, the functions $\varphi_i(x), \ i = 1,\ldots,M$ span $\mathcal{U}_h$. It is simple to check they are linearly independent.

**The Galerkin / Finite Element Method**

Find $u_h \in \mathcal{U}_h$ such that

$$(\partial_x v_h, \partial_x u) = (f,u), \quad \text{for all } u \in \mathcal{U}_h$$

or equivalently

$$(\partial_x v_h, \partial_x \varphi_j) = (f,\varphi_j), \quad \text{for all } j = 1,\ldots,M \tag{5.12}$$

Exercise: Using the ideas from our discussion of the Dirichlet's principle, show that (5.12) is equivalent to finding $u_h \in \mathcal{U}_h$, such that

$$I[v_h] = \min_{u \in \mathcal{U}_h} I[u] \tag{5.13}$$

**Implementation of the finite element method**

Since $v_h \in \mathcal{U}_h$, we have that $v_h(x) = \sum_{k=1}^M \xi_k \varphi_k(x)$, where $\xi_k = v_h(x_k)$ is to be determined. Substitution into (5.12) gives

$$\left( \sum_{k=1}^M \xi_k \partial_x \varphi_k \ , \partial_x \varphi_j \right) = (f,\varphi_j), \ j = 1,\ldots,M.$$

Equivalently,

$$A\,\xi = b, \tag{5.14}$$

where $A = (a_{ij})$ denotes the *stiffness matrix* and $b = (b_i)$ denotes the *load* vector given by:

$$\begin{aligned} a_{ij} &= (\ \partial_x \varphi_i, \partial_x \varphi_j\ ), \quad i,j = 1, \ldots, M \\ b_i &= (f, \varphi_i). \end{aligned}$$

Clearly, $A$ is symmetric.

**Computation of and properties of the stiffness matrix**

(1) Note that $(\partial_x \varphi_i, \partial_x \varphi_j) = 0, \quad |i-j| > 1$. Therefore, $A$ is a *tridiagonal* matrix. The diagonal and off-diagonal elements are given by:

$$a_{ii} = (\partial_x \varphi_i, \partial_x \varphi_i) \ = \ h_i^{-1} + h_{i+1}^{-1}, \quad a_{i,i+1} = a_{i+1,i} = (\partial_x \varphi_i, \partial_x \varphi_{i+1}) = -h_i^{-1}$$

(2) $A$ is symmetric and positive definite. Symmetry was noted above. Positive definiteness is seen as follows: For any $\eta \in \mathbb{R}^M$

$$\begin{aligned} \eta^T \ A \ \eta &= \sum_{i,j=1}^{M} \eta_i \ (\partial_x \varphi_i, \partial_x \varphi_j) \eta_j \\ &= \left( \sum_{i=1}^{M} \eta_i \partial_x \varphi_i, \sum_{j=1}^{M} \eta_j \partial_x \varphi_j \right) \ = \ \|v\|^2 \geq 0. \end{aligned}$$

where $v = \sum_{i=1}^{M} \eta_i \partial_x \varphi_i$. Moreover, $\eta^T A \eta = 0$ only if $\eta = 0$ because $\{\varphi_i : i = 1, \ldots, M\}$ is a linearly independent set.

**Theorem:** The finite element method reduces to the solution of the system of $M$ nonhomogeneous linear algebraic equations $A\xi = b$, (5.14), with stiffness matrix, $A$, and load vector, $b$. $A$ is symmetric and positive definite. It is therefore invertible and the system can be solved for any load vector $b$.

**Special case:** Uniform partition of $[0,1]$: Here, $h_j = x_{j+1} - x_j = h = (M+1)^{-1}$. Therefore, $a_{ii} = 2/h$ and $a_{i,i+1} = a_{i+1,i} = -1/h$. Thus,

$$A \ = \ \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & \ldots & 0 \\ 0 & -1 & 2 & -1 & \ldots & 0 \\ . & . & . & . & \ldots & . \\ 0 & . & 0 & -1 & 2 & -1 \\ 0 & 0 & \ldots & 0 & -1 & 2 \end{pmatrix} \tag{5.15}$$

# 6 Operators on Hilbert spaces

The setting: complex Hilbert space with inner product, $(\cdot,\cdot)_{\mathcal{H}}$.

## 6.1 Bounded operators

**Definition 6.1 Bounded linear transformation or bounded linear operator** *A linear transformation on $\mathcal{H}$ is bounded if there is a constant, $C \geq 0$, such that for all $x \in \mathcal{H}$,*

$$\|T(x)\| \leq C\|x\| \tag{6.1}$$

*The smallest constant, $C$, for which (6.1) holds, is called the norm of $T$ and is denoted by $\|T\|$,*

$$\|T\| \ = \ \sup_{x\neq 0} \frac{\|T(x)\|}{\|x\|} \ = \ \sup_{\|x\|=1} \|T(x)\| \tag{6.2}$$

*We sometime write $\|T\|_{\mathcal{L}(\mathcal{H},\mathcal{H})}$ to explicitly denote the operator norm of $T$, which maps $\mathcal{H}$ to $\mathcal{H}$, and more generally write,*

$$\|T\|_{\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)} \ = \ \sup_{\|x\|_{\mathcal{H}_1}=1} \|T(x)\|_{\mathcal{H}_2}$$

*when speaking of $T$ as an operator between different spaces $\mathcal{H}_1$ and $\mathcal{H}_2$.*

**Definition 6.2** *(***Self-adjoint operator***) $T : \mathcal{H} \to \mathcal{H}$ is self-adjoint if*

$$(Tx, y) \ = \ (x, Ty), \quad \text{for all } x, y \in \mathcal{H} \tag{6.3}$$

**Example 1:** $\mathcal{H} = \mathbb{C}^n$, $(x,y)_{\mathbb{C}^n} = \sum_j x_j\overline{y_j}$. Linear transformation on $\mathbb{C}^n$: $T(x) = Tx$, $T = (t_{ij})$, $n$ by $n$ matrix. $T$ is self-adjoint if $T = (t_{ij}) = (\overline{t_{ji}}) = T^*$.

$$T = T^* \implies \|T\| \ = \ max\{\ |\lambda| \ : \ \lambda \in \sigma(T)\ \},$$

where $\sigma(T)$ denotes the set of eigenvalues, the *spectrum*, of the matrix $T$.

**Example 2:** $\mathcal{H} = L^2(\mathbb{R}^n)$. Fix $\tau \in \mathbb{R}^n$. Define $T_\tau f = f(x+\tau)$. By change of variables $\|T_\tau f\| = \|f\|$. Thus, $T_\tau$ is unitary on $L^2$; $\|T_\tau\| = 1$.

**Example 3:** $\mathcal{H} = L^2(\mathbb{R}^n)$. Define $\mathcal{F}[f](\xi) = \hat{f}(\xi) = \int e^{-2\pi i x\cdot\xi} f(x)\,dx$, the Fourier transform of $f$; see section 12.5. The Plancherel Theorem states: $\|\mathcal{F}[f]\| \ = \ \|f\|$. Thus, $\mathcal{F}$ is unitary on $L^2$; $\|\mathcal{F}\| = 1$.

**Example 4:** $\mathcal{H} =$ the set of $1-$ periodic $L^2$ functions:

$$L^2(S^1) = \{f : f(x+1) = f(x),\ x \in \mathbb{R},\ \int_0^1 |f|^2 < \infty\ \}$$

Define, for $n \in \mathbb{Z}$, the Fourier coefficients

$$\mathcal{F}[f](n) = \hat{f}(n) \ = \ (f, \phi_n) \ = \ \int_0^1 f(z)e^{-2\pi i n z}\,dz$$

Parseval - Plancherel Theorem states

$$\|f\|^2_{L^2(S^1)} \;=\; \int_0^1 |f(z)|^2 dz \;=\; \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$$

Thus, the mapping $\mathcal{F} : L^2(S^1) \to l^2(\mathbb{Z})$, where $l^2(\mathbb{Z})$ is the set of *sequences*, which are square summable, is unitary in $L^2$; $\|\mathcal{F}\|_{\mathcal{L}(L^2(S^1);l^2(\mathbb{Z}))} = 1$.

**Definition 6.3** *An operator $T : \mathcal{H} \to \mathcal{H}$ has finite rank if its range is finite dimensional.*

**Example 5:** Let $\phi \in \mathcal{H}$, and $\|\phi\|_{\mathcal{H}} = 1$. Define the *projection operator*

$$P_\phi f \;=\; (f, \phi)_{\mathcal{H}}\ \phi$$

An important example is the family of projection operators, $P_j$, defined by

$$P_j f(x) \;=\; \hat{f}(j)\ e^{2\pi i j x}$$

for any $f \in L^2(S^1)$. Note that the family of operators $\{P_j\}$ are *orthogonal projections*. That is, $P_j\ P_j \;=\; P_j^2 \;=\; P_j$ and $P_j\ P_k \;=\; 0$ if $j \neq k$.

**Example 6:** Let $\mathcal{P}^{(N)} = \text{span}\{\ e^{2\pi i j x} :\ |j| \le N\ \}$. Let $P^{(N)} = \sum_{|j|\le N} P_j$; see the previous example. The operator $\frac{d}{dx} P^{(N)}$ maps $\mathcal{P}^{(N)}$ to itself. What is its matrix representation with the respect to the basis $\{\ e^{2\pi i j x} :\ |j| \le N\ \}$?

## 6.2 Compact operators

This would be a good time to review the background material on compactness and convergence in section 12.4.

**Definition 6.4** *Let $T : \mathcal{H}_1 \to \mathcal{H}_2$ denote a bounded linear transformation between Hilbert spaces. $T$ is compact if for any bounded sequence, $\{u_k\}$ in $\mathcal{H}_1$, the sequence $\{T(u_k)\}$ has a convergent subsequence in $\mathcal{H}_2$. That is, if for all $k$, $\|u_k\| \le M$, then there exists a subsequence $\{T(u_{k_j})\}$ and a $v_* \in \mathcal{H}_2$ such that $\|T(u_{k_j}) - v_*\|_{\mathcal{H}_2} \to 0$ as $j \to \infty$.*

**Example 6.1** *Finite rank operators are compact. The proof is an exercise, using the compactness of closed and bounded subsets of $\mathbb{C}^n$.*

The importance of finite rank operators in the general theory is that *any* compact operator in a separable Hilbert space can be approximated arbitrarily well by finite rank operators.

**Theorem 6.1** *If $\mathcal{H}$ is a separable Hilbert space and $T$ is compact, then $T$ is the norm limit of operators of finite rank. That is, there exists $T_j$ compact and finite rank such that $\|T_j - T\|_{\mathcal{L}(\mathcal{H},\mathcal{H})} \to 0$.*

A very important class of operators is the class of **Hilbert Schmidt operators**:

**Definition 6.5** *$K(x,y) : S \times S \to \mathbb{R}$ is a called a Hilbert-Schmidt kernel if*

$$\int\int_{S\times S} |K(x,y)|^2 \; dx \; dy \; < \; \infty$$

*We use $K(x,y)$ to generate an operator on the Hilbert space $L^2(S)$.*

$$T_K[f](x) \;=\; \int_S K(x,y) \; f(y) \; dy \tag{6.4}$$

**Theorem 6.2** *If $K(x,y)$ is a Hilbert-Schmidt kernel, then $T_K$ is a compact operator.*

**Proof:** Let $\{\phi_j(x)\}_{j\geq 1}$ denote an orthonormal basis for $L^2(S)$. Then, $\psi_{jk}(x,y) = \phi_j(x)\phi_k(y)$ is an orthonormal basis for $L^2(S\times S)$. Define $K_N(x,y)$ to be the "partial sum" approximation to $K(x,y)$ given by:

$$K_N(x,y) \;=\; \sum_{j+k\leq N} a_{jk}\psi_{jk}(x,y) \tag{6.5}$$

Here, $a_{jk} = \int_S \int_S K(x,y)\phi_j(x) \; dx \; \phi_k(y)dy$ is the $(j,k)$ Fourier coefficient of $K$. Then,

$$\|T_K - T_{K_N}\| \;\leq\; \|K - K_N\|_{L^2(S\times S)} \;\to\; 0, \quad N \to \infty$$

**Definition 6.6** *Let $T$ denote a bounded operator on $\mathcal{H}$. The resolvent set of $T$, denoted $\rho(T)$, given by*

$$\rho(T) \;=\; \{ z \in \mathbb{C} \;:\; (T - zI)^{-1} \text{ exists and is a bounded operator on } \mathcal{H} \;\}. \tag{6.6}$$

*The spectrum of $T$, denoted by $\sigma(T)$, is the complement of $\rho(T)$:*

$$\sigma(T) \;=\; \mathbb{C} \;-\; \rho(T)$$

**Theorem 6.3** *(***Variant of the Riesz-Schauder Theorem***)*
*Let $T$ denote a compact operator on $\mathcal{H}$.*

*(1) $\sigma(T)$ is discrete*

*(2) If $\mathcal{H}$ is infinite dimensional, then $\sigma(T)$ is infinite and consists of eigenvalues accumulating at zero. Thus, $0 \in \sigma(T)$.*

*(3) The eigenvalues of $T$ have finite multiplicity. If $\lambda \neq 0$ is an eigenvalue of $T$, then $\dim\{x : Tx = \lambda x\ \}$ is finite.*

The following result generalizes the well-known fact that Hermitian symmetric or self-adjoint $n \times n$ matrices have a complete set of eigenvectors that span $\mathbb{C}^n$.

**Theorem 6.4** *(***Hilbert Schmidt Theorem***) Let $T : \mathcal{H} \to \mathcal{H}$ be a compact and* **self-adjoint** *operator. Then, there exists an orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of $T$. That is, there exists $\{x_j\}_{j\geq 1}$, $Tx_j = \lambda_j x_j$, $(x_j, x_k) = \delta_{jk}$, such that for any $x \in \mathcal{H}$*

$$\left\| \; x \;-\; \sum_{j=1}^{N} \; (x, x_j)x_j \; \right\|_{\mathcal{H}} \;\to 0$$

# 7 Applications of compact operator theory to eigenfunction expansions for elliptic operators

By Theorem 5.1, the Dirichlet problem for Poisson's equation

$$-\Delta v \ = \ f, \ x \in \Omega; \quad u\,|_{\partial\Omega} = 0$$

has a weak $H_0^1(\Omega)$ solution, *i.e.* There exists a unique $v_f \in H_0^1(\Omega)$ such that

$$\int \nabla u \cdot \nabla v_f \ dx \ = \ \int uf \ dx, \quad u \in H_0^1(\Omega) \tag{7.1}$$

The *solution operator* $T : f \mapsto v_f, \quad L^2(\Omega) \to H_0^1(\Omega)$ is a bounded operator. Think of this mapping as

$$(-\Delta_D)^{-1} : \ f \mapsto (-\Delta_D)^{-1} f,$$

where $-\Delta_D$ denotes the *Dirichlet Laplacian*, the Laplace operator acting in the space of functions with zero boundary conditions.

We next view $T = (-\Delta_D)^{-1}$ as an operator from $L^2(\Omega)$ to itself and can estimate its norm.

Setting $u = v_f$ in the weak formulation (7.1) we obtain

$$\int_\Omega |\nabla v_f|^2 dx \ = \ \int_\Omega v_f f dx.$$

By the Cauchy Schwarz inequality,

$$\int_\Omega |\nabla v_f|^2 dx \ \le \ \left( \int_\Omega |v_f|^2 dx \right)^{\frac{1}{2}} \ \left( \int_\Omega |f|^2 dx \right)^{\frac{1}{2}}.$$

By the Poincaré inequality (5.6), ($\|g\|_{L^2} \le p_\Omega \|\nabla g\|_{L^2}$, for $g \in H_0^1(\Omega)$), we have

$$p_\Omega^{-2} \|v_f\|_{L^2}^2 \ \le \ \int_\Omega |\nabla v_f|^2 dx \ \le \ \|v_f\|_{L^2} \ \|f\|_{L^2}$$

In other words,

$$\|Tf\|_{L^2} \ \le \ p_\Omega^2 \ \|f\|_{L^2(\Omega)}; \tag{7.2}$$

$T$ is bounded on $L^2$ and $\|T\| \le p_\Omega^2$.

We now claim that $T$ is compact and self-adjoint.

**Theorem 7.1** *The solution operator $T : f \mapsto Tf = v_f$, the solution of the Dirichlet problem for Poisson's equation is compact and self-adjoint operator for $L^2(\Omega)$ to $L^2(\Omega)$. Therefore, by the Hilbert-Schmidt Theorem 6.4 there exists an orthonormal set of eigenfunctions of $T$ which is complete in $L^2(\Omega)$.*

**Proof:** To show that $T$ is compact, we must show that if $f_j$ is a bounded sequence in $L^2(\Omega)$, then $Tf_j$ has a convergent subsequence. The key tool here is the *Rellich compactness Lemma*, Theorem 12.11[10]

Suppose that $f_j$ is a sequence for which there is a constant $M$ such that $\|f_j\|_{L^2} \leq M$. By (7.2)

$$\|Tf_j\|_D \ \leq \ p_\Omega^2 \ \|f\|_{L^2(\Omega)} \ \leq \ p_\Omega^2 \ M \tag{7.3}$$

Thus, $Tf_j$ is a bounded sequence in $H_0^1(\Omega)$. By Theorem 12.11, this sequence has an $L^2(\Omega)$ convergent subsequence, $\{Tf_{j_k}\}$.

Next, we verify self-adjointness of $T$. By the weak formulation of the Dirichlet problem and the definition of $T$:

$$\int \nabla Tf \cdot \overline{u} \ dx \ = \ \int f \ \overline{u} \ dx, \quad u \in H_0^1(\Omega) \tag{7.4}$$

Setting $u = Tg$, where $g \in L^2(\Omega)$, we have

$$\int \nabla Tf \cdot \overline{\nabla Tg} \ dx \ = \ \int f \cdot \overline{Tg} \ dx, \tag{7.5}$$

Similarly,

$$\int \nabla Tg \cdot \overline{\nabla Tf} \ dx \ = \ \int g \cdot \overline{Tf} \ dx, \quad u \in H_0^1(\Omega) \tag{7.6}$$

Therefore,

$$\int Tf \ \overline{g} \ dx \ = \ \int f \ \overline{Tg} \ dx. \tag{7.7}$$

That is, $(Tf, g)_{L^2} = (f, Tg)_{L^2}$; $T$ is self-adjoint. The existence of a complete set of eigenfunctions of $T$ is now a consequence of the Hilbert Schmidt Theorem, Theorem 6.4.

**Corollary 7.1** *$L^2(\Omega)$ has an orthonormal basis of eigenfunctions of the $-\Delta$, satisfying Dirichlet boundary conditions.*

**Proof:** Let $\phi_j \in H_0^1(\Omega)$ denote the orthonormal sequence of eigenfunctions of $T$. By the definition of $T$

$$\begin{aligned}
&\Longleftrightarrow \ (T\phi_j, u)_D \ = \ (\phi_j, u)_{L^2} \\
&\Longleftrightarrow \ (\alpha_j \phi_j, u)_D \ = \ (\phi_j, u)_{L^2} \\
&\Longleftrightarrow \ \alpha_j \ (\phi_j, u)_D \ = \ (\phi_j, u)_{L^2}, \quad \alpha_j \text{ real} \\
&\Longleftrightarrow \ \int \nabla \phi_j \cdot \overline{\nabla u} \ dx \ = \ \int \frac{1}{\alpha_j} \phi_j \ \overline{u} \ dx
\end{aligned}$$

Therefore, the eigenfunctions of $T$ are weak solutions of $-\Delta \phi_j \ = \ \lambda_j \phi_j, \ \lambda_j = \frac{1}{\alpha_j}, \ \ \phi_j \in H_0^1(\Omega)$. "Interior" elliptic regularity results ensures that these are, in fact, $C^\infty(\Omega)$. It can

[10]**Theorem:** ( Rellich compactness Lemma) Suppose $\{u_j\}$ is a sequence in $H_0^1(\Omega)$ such that $\|u_j\|_{H_0^1(\Omega)} \leq C$. Then, there exists a subsequence $u_{j_k}$ and an element $u_* \in H_0^1$ such that $\|u_{j_k} - u_*\|_{L^2} \to 0$ as $j_k \to \infty$.

also be shown, using "boundary" elliptic regularity results that if $\partial\Omega$ is sufficiently smooth, then $\lim_{x\to\partial\Omega} \phi_j(x) = 0$[11].

Moreover, setting $u = \phi_j$ above yields positivity of the eigenvalues

$$\lambda_j = \frac{1}{\alpha_j} = \int_\Omega |\nabla\phi_j|^2 \, dx.$$

FInally, since $\alpha_j \to 0$ as $j \to \infty$, $\lambda_j \to \infty$.

[11] See L.C. Evans, *PDE*, section 6.3 for a discussion of interior and boundary elliptic regularity

# 8 Introduction to Variational Methods

In these notes we introduce *Variational Methods* or *The Calculus of Variations* via two classical problems:

(1) **Dirichlet's principle**, which characterizes the solution of the Poisson equation as the solution of an optimization (minimization) problem[12].

(2) The **Rayleigh-Ritz**, characterization of the smallest eigenvalue of the Dirichlet problem for the Laplacian as the solution of a minimization problem[13]

## 8.1 Variational Problem 1: Dirichlet's Principle

We introduce a class of *admissible functions*, $\mathcal{A}$

$$\mathcal{A} \equiv \{ w \in C^2(\bar{U}) : w = g \text{ on } \partial U \} \tag{8.1}$$

For $w \in \mathcal{A}$, define the *Dirichlet energy*:

$$I[w] \ = \ \int_U \frac{1}{2} |\nabla w(x)|^2 - w(x) f(x) dx \tag{8.2}$$

**Theorem 8.1** *(Dirichlet's principle[14])*

*(1) If $u \in \mathcal{A}$ and $-\Delta u = f$ in $U$, then*

$$I[u] = \min_{w \in \mathcal{A}} I[w]. \tag{8.3}$$

*(2) Conversely, if $u \in \mathcal{A}$ and $I[u] = \min_{w \in \mathcal{A}} I[w]$, then $-\Delta u = f$ and $u = g$ for $x \in \partial U$.*

We have studied, by Hilbert Space - functional analytic methods, the solution of the Dirichlet problem for Poisson's equation, for the case $g = 0$. The above Theorem gives an alternative characterization of the solution.

## 8.2 Variational Problem 2: Smallest eigenvalue of $-\Delta$

*Poincaré inequality*[15]: Let $\Omega$ denote a domain in $\mathbb{R}^n$ which is open, connected and can be bounded between two planes. There exists a positive constant, $p_\Omega$, depending on the domain $\Omega$, such that for any $u \in H_0^1(\Omega)$

$$\int_\Omega |u|^2 \ dx \ \le \ p_\Omega \int_\Omega |\nabla u|^2 \ dx. \tag{8.4}$$

[12] Section 2.2.5 of L.C. Evans PDEs, pages 42-43

[13] Section 6.5.1,Theorem 2, We follow an approach to Theorem 2, using the methods of Chapter 8 of L.C. Evans, PDEs.

[14] Theorem 17, page 42 of L.C. Evans *PDEs*

[15] We established Poincaré's inequality in our study of weak solutions to Poisson's equation.

We proved, in particular, that if $\Omega$ is bounded between two planes a distance $2a$ apart, then $p_\Omega$ can be taken to be $2a^2$.

**Question:** *What is the smallest constant, $p^*_\Omega$ for which the Poincaré's inequality holds?*

To answer this question, we need to compute

$$\begin{aligned}\frac{1}{p^*_\Omega} &\equiv \inf_{0\neq u\in H^1_0(\Omega)} R[u]\\ &= \inf_{0\neq u\in H^1_0(\Omega)} \frac{\int_\Omega |\nabla u|^2\ dx}{\int_\Omega |u|^2\ dx}\end{aligned} \tag{8.5}$$

**Remark 8.1** *As in Dirichlet's principle, we have an admissible class of functions,, $\mathcal{A} = H^1_0(\Omega)$ in this case, and a functional $R[u]$, defined on $\mathcal{A}$ which we seek to minimize.*

Suppose the minimium in (8.5) is attained a function $u_* \in H^1_0(\Omega)$. Let $\eta \in C^\infty_0(\Omega)$ be fixed. Then, we have for any $\epsilon$, $u_* + \epsilon\eta \in H^1_0(\Omega)$ and

$$R[u_* + \epsilon\eta]\ \geq\ R[u_*]$$

Thus the function $r(\epsilon) = R[u_* + \epsilon\eta]$ is minimized at $\epsilon = 0$. In particular, we must have

$$\frac{d}{d\epsilon} R[u_* + \epsilon\eta]\,|_{\epsilon=0}\ =\ 0. \tag{8.6}$$

Equation (8.6) can be rewritten equivalently as

$$\int_\Omega \nabla u\cdot\nabla\eta\ dx\ -\ R(u_*)\int_\Omega u\eta = 0$$

Thus, for any $\eta \in C^\infty_0(\Omega)$

$$\int\ (\ -\nabla u\ \nabla\eta\ -\ R(u_*)\ u\eta\ )\ dx = 0 \tag{8.7}$$

In other words, $(u_*, R(u_*))$ is a weak solution of the eigenvalue problem

$$-\Delta\varphi\ =\ \lambda\varphi,\ \varphi = 0,\quad x\in\partial\Omega. \tag{8.8}$$

$u_*$ is an eigenfunction of $-\Delta$ with corresponding eigenvalue $R(u_*)$.

We have previously seen that the eigenvalue problem (8.8) has an infinite sequence of "Dirichlet eigenvalues" $\lambda_j$, with $\lambda_0 < \lambda_1 \leq \lambda_2 \ldots$, and corresponding eigenfunctions $\varphi_j \in H^1_0(\Omega)$ [16]. Note that $R(\varphi_{\lambda_j}) = \lambda_j$ and therefore we have $\lambda_1 = R(u_*)$. Borrowing from terminology in quantum physics, $\lambda_1(\Omega)$ is sometimes called the ground state energy and $u_*$ the ground state eigenfunction. We therefore have the following equivalence

[16] This we shown by proving that the mapping $(-\Delta)^{-1} : f \mapsto u_f = (-\Delta)^{-1}f$ which maps $f \in L^2(\Omega)$ to the *weak* $H^1_0(\Omega)$ solution of the Dirichlet problem is a compact linear operator from $L^2(\Omega)$ to itself. We then applied the general spectral theorem for compact self-adjoint operators in Hilbert space.

**Theorem 8.2** *(a) $\lambda_0(\Omega) > 0$ is the smallest Dirichlet eigenvalue of the Laplacian, $-\Delta$.*

*(b) $\lambda_0(\Omega) \; = \; \min R(u)$, the minimum of the Rayleigh quotient over $u \in H^1_0(\Omega)$.*

*(c) $(\lambda_0(\Omega))^{-1}$ is the best (smallest) constant for which the Poincaré inequality holds, i.e. for any $u \in H^1_0(\Omega)$*

$$\int_\Omega |u(x)|^2 \; dx \;\le\; \frac{1}{\lambda_0(\Omega)} \int_\Omega |\nabla u(x)|^2 \; dx. \tag{8.9}$$

## 8.3 Direct methods in the calculus of variations

In both variational problems, we **assumed** that the minimum of of the Dirichlet energy, (8.3), and the Rayleigh quotient, $R(u)$ in (8.5) are *attained* at an admissible function. The question of whether this is in fact the case is quite subtle. In both problems we have the following set up:

(a) a functional $J[u]$ defined on an admissible class of functions, $\mathcal{A}$ ($J[u] = I[u]$ in variational problem 1 and $J[u] = R[u]$ in variational problem 2).

(b) $J[u] \; > -\infty$ for $u \in \mathcal{A}$.

Thus, $\inf_{u\in\mathcal{A}} J[u] \; = \; J_*$ exists. In particular, there is a *minimizing sequence* of functions $u_j \in \mathcal{A}$ such that $J[u_j] \downarrow J_*$.

**Exercise 1:** Prove that for both Dirichlet's and Rayleigh-Ritz principles that $J[u]$ has a lower bound, when $u$ varies over the appropriate admissible set of functions, $\mathcal{A}$.

**THE QUESTION:** Does the minimizing sequence $\{u_j\}$ converge to some $u_*$ which is (a) admissible, *i.e.* $u_* \in \mathcal{A}$ and (b) a minimizer, *i.e.* $J[u_*] = \lim_{j\to\infty} J[u_j] = J_*$?

The property that $\{u_j\}$ is a *minimizing* sequence can often be used to conclude some kind of

(c) uniform upper bound on norms of the functions $u_j$.

The goal is to deduce from (a), (b) and (c) that

(d) the sequence of $\{u_j\}$, or possibly some subsequence of it, has "good convergence properties to a limiting admissible function $u_* \in \mathcal{A}$", for which $J[u_*] = j_*$.

Assuming this and if we set aside the issue of whether minimizers are $C^2$ functions[17], the arguments of the previous section apply to give variational characterizations to solutions of PDEs.

*For the remainder of these notes, we shall focus primarily on the solution of Variational Problem 2: Lowest eigenvalue of $-\Delta$ and Poincaré's inequality.*

[17]This is an important and very technical issue. See the discussion in L.C. Evans PDEs, Chapter 8.3

## 8.4 Ideas and general discussion

We now turn to the following questions:
**(Q1)** How do we deduce a uniform bound on norms of the function $u_j$ from (a) and (b)?, and
**(Q2)** How do we show that, perhaps by passing to some subsequence of the $u_j's$ that we can find a sequence which converges to a function $u_*$ in the admissible set, at which the functional attains its minimum.

The approach we take to the existence of solutions of a PDE via the proof that a functional is minimized is called *the direct method in the calculus of variations*[18].
**Exercise 2:** Concerning **(Q1)**, prove that if $u_j$ is a minimizing sequence for either of the variational problems in section 8, then there is a constant $C$, such that $\|u_j\|_{H_0^1(\Omega)} \leq C$ for all $j$.

Concerning **(Q2)**, we are interested in when bounds on a sequence of functions imply the existence of a convergent subsequence. This is the subject of *compactness*; see section 12.4.

## 8.5 Compactness and minimizing sequences

**Exercise 5:** Prove that minimizing sequences for the variational problems in section 8 have subsequences which converge weakly in $H^1$.

To show that subsequence of any minimizing sequence for the problems of section 8 converge to an admissible minimizer, we use the following key properties:

**Property A** *Weak compactness of minimizing sequences*: There is a subsequence $\{u_j\}$ such $u_j \to u_* \in H_0^1(\Omega)$ weakly.

**Property B** *(Strong) compactness in $L^2(\Omega)$ of minimizing sequences*: There is a subsequence for which $\|u_j - u_*\|_2 \to 0$.

**Property C** *Weak lower semicontinuity of a functional with respect to weak convergence*[19]:

$$\int_\Omega |\nabla u_*|^2 \, dx \ \leq \ \liminf_{j\to\infty} \int_\Omega |\nabla u_j|^2 \, dx \tag{8.10}$$

Here's how Properties A, B and C are used to established the minimum is attained. Consider variational problem 2 to minimize $R(u)$. Recall that $J_* = 1/p_\Omega^* > 0$ denotes the positive infimum. By scaling, we can assume that $\{u_j\}$ is such that $\|u_j\|_2 = 1$. Moreover, it is clear since $R(u_j)$ approaches its positive infimum, that $\|\nabla u_j\|_2$ is bounded. Thus $u_j$ is a bounded sequence in $H_0^1(\Omega)$. By weak compactness, there's a subsequence which we'll, for simplicity call $u_j$, which converges weakly to some $u_* \in H_0^1(\Omega)$. Furthermore, Rellich's Lemma implies there is a subsequence, which we'll call once again $u_j$ which converges strongly to some $u_*$. (Note: the weak limit is unique.). By strong convergence in $L^2$ we

$$| \ \|u_j\|_2 \ - \ \|u_*\|_2 \ | \ \leq \|u_j - u_*\|_2 \to 0$$

[18] An excellent introductory reference is the book of Gelfand and Fomin: Calculus of Variations

[19] This is a special case of the lower semi-continuity of any convex functional with respect to weak convergence. Compare also, with Fatou's Lemma from basic real analysis.

Thus, $u_*$ is admissible, *i.e.* $u_* \in H_0^1(\Omega)$ and $\|u_*\|_2 = 1$.

By weak lower semicontinuity and the fact that $\{u_j\}$ is a minimizing sequence

$$J_* \; = \; \lim_j R[u_j] = \liminf_j R[u_j] \geq R[u_*].$$

But $u_*$ is admissible, so $R[u_*] \; \geq \; J_*$ and the minimum is attained $R[u_*] = j_*$.

What's left is the **Justification of Properties A, B and C:**
**Property A** follows from the above theorem on weak compactness.

**Property B** is a consequence of the following theorem, which was at the heart of the Hilbert space approach to the Dirichlet problem and eigenfunction expansions.

**Proof of Property C - weak lower semicontinuity:** To prove (8.10) we begin with the following observation:

$$\int_\Omega |\nabla u_k|^2 \; - \; \int_\Omega |\nabla u_*|^2 \; = \; \int_\Omega |\nabla(u_k - u_*)|^2 \; + 2\int_\Omega \nabla u_* \cdot \nabla(u_k - u_*), \tag{8.11}$$

which is proved by simply expanding $\nabla(u_k - u_*) \cdot \nabla(u_k - u_*)$. Dropping the positive term on the right hand side gives

$$\int_\Omega |\nabla u_k|^2 \; - \; \int_\Omega |\nabla u_*|^2 \; \geq \; 2\int_\Omega \nabla u_* \cdot \nabla(u_k - u_*), \tag{8.12}$$

Letting $k$ tend to infinity and using that $u_k$ tends weakly to $u_*$ in $H^1$ we see that the term on the right hand side approaches zero as $k \to \infty$. This implies

$$\int_\Omega |\nabla u_*|^2 \; \leq \; \liminf_{k\to\infty} \int_\Omega |\nabla u_k|^2.$$

# 9 The Heat / Diffusion Equation

**Initial value problem**

$$\begin{aligned} \partial_t u &= \Delta u \\ u(x,0) &= f(x) \end{aligned}$$

**Solution by Fourier transform**

$$\begin{aligned} \partial_t \hat{u}(\xi,t) &= -4\pi^2|\xi|^2\hat{u}(\xi,t), \quad \hat{u}(\xi,0) = \hat{f}(\xi) \\ \hat{u}(\xi,t) &= e^{-4\pi^2|\xi|^2 t}\ \hat{f}(\xi) \\ u(x,t) &= \int e^{2\pi i\xi\cdot x}\ e^{-4\pi^2|\xi|^2 t}\ \hat{f}(\xi)\ d\xi \\ &= \int f(y)\ dy \int e^{2\pi i\xi\cdot(x-y)}\ e^{-4\pi^2|\xi|^2 t}\ d\xi \\ &= \int K_t(x-y)\ f(y)\ dy, \quad K_t(z) = \frac{1}{(4\pi t)^{\frac{n}{2}}}\ e^{-z^2/4t} \end{aligned}$$

**Theorem 9.1** *Let $f$ be bounded and continuous on $\mathbb{R}^n$. For $t > 0$, define*

$$u(x,t) = \int_{\mathbb{R}^n} K_t(x-y)\ f(y)\ dy \equiv e^{\Delta t} f. \tag{9.1}$$

*Then,*

*(a)* $u \in C^\infty$ *for any* $t > 0$,

*(b)* $u(x,t)$ *satisfies the heat equation, and*

*(c)* $\lim_{t\to 0} u(x,t) = f(x)$.

$K_t(x)$ *is called the fundamental solution for the operator* $L = \partial_t - \Delta$, *i.e.*

$$(\ \partial_t - \Delta\ )\ K_t(x) = 0,\ t > 0 \quad \lim_{t\downarrow 0} K_t(x) = \delta_0,$$

*where* $\delta_0$ *denotes the Dirac delta distribution.*

**Proof:** The explicit construction guarantees that (a) and (b) hold. To prove part (c) we note from the following properties of $K_t(x)$.

(K1) $K_t(x) > 0, \quad t > 0$

(K2) $\int K_t(x)\ dx = 1, t > 0.$

(K3) Let $\delta > 0$ be fixed. Then,

$$\lim_{0<t\to 0} \int_{|x-y|\geq\delta} K_t(y)\ dy = 0,$$

uniformly in $x$.

We need to show $u(x,t) - f(\xi) \to 0$ as $(x,t) \to (\xi, 0)$.

$$
\begin{aligned}
|\ u(x,t)\ -\ f(\xi)\ | &= \left| \int_{\mathbb{R}^n} K_t(x-y)\ (\ f(y) - f(\xi)\ )\ dy \right| \\
&= \int_{\mathbb{R}^n} K_t(x-y)\ |f(y)-f(\xi)|\ dy, \quad (K_t > 0) \\
&= \int_{|x-y|\le\delta} K_t(x-y)\ |f(y)-f(\xi)|\ dy\ +\ \int_{|x-y|>\delta} K_t(x-y)\ |f(y)-f(\xi)|\ dy \\
&= I_1\ +\ I_2
\end{aligned}
$$

Here, we have used property (K1). *Estimation of $I_1$:* Here, $|x-y| \le \delta$: Let $\varepsilon > 0$ be arbitrary. Using the continuity of $f$, we can choose $\delta = \delta(\varepsilon)$ so that if $|y-\xi| < 2\delta$ then $f(y)-f(\xi)| < \varepsilon$. By taking $|x-\xi| < \delta$, we have $|y-\xi| \le |y-x| + |x-\xi| \le 2\delta$. Thus,

$$I_1\ \le\ \varepsilon \int_{|x-y|\le\delta} K_t(x-y)\ dy\ \le\ \int_{\mathbb{R}^n} K_t(x-y)\ dy\ =\ \varepsilon,$$

where we have used property (K2).
*Estimation of $I_2$:* Here, $|x-y| \ge \delta$.

$$I_2\ \le\ 2\|f\|_{L^\infty} \int_{|x-y|\ge\delta(\varepsilon)} K_t(x-y)\ dy, \tag{9.2}$$

which tends to zero, as $t \downarrow 0$ uniformly in $x$, by property (K3).

Thus we have shown

$$|u(x,t)\ -\ f(\xi)|\ \le\ \varepsilon\ +\ o(1)$$

as $t \downarrow 0$, where $\varepsilon$ is arbitrary. This proves the theorem.

## 9.1 Remarks on solutions to the heat / diffusion equation

Let $u(x,t) = e^{\Delta t} f$, the solution of the heat equation with initial data $f(x)$.

(1) *Instantaneous smoothing and infinite propagation speed* $f \ge 0$, bounded and of compact support $\Longrightarrow$ for all $t > 0$, $e^{\Delta t} f \in C^\infty$, $e^{\Delta t} f > 0$ for all $x$.
Contrast this with the transport (wave) equation $(\partial_t + c\partial_x) u(x,t) = 0$ with initial data $u(x,0) = f(x)$. The solution,$u(x,t) = f(x - ct)$ is clearly no smoother than the initial data.

(2) $f \ge 0$, $u(x,0) = f \in L^1 \Longrightarrow \int u(x,t)dx\ =\ \int u(x,0)dx,\ t > 0$.

(3) $\|e^{\Delta t} f\|_{L^2}\ \le\ \|f\|_{L^2},\quad t \ge 0$. More precisely,

$$\frac{d}{dt} \int |u(x,t)|^2\ dx\ =\ -2 \int |\nabla u(x,t)|^2\ dx \tag{9.3}$$

(4) $|u(x,t)|\ \le\ C\ t^{-\frac{n}{2}}\ \|f\|_{L^1},\quad t > 0$.

(5) $\inf f \ \leq\ e^{\Delta t} f \ \leq\ \sup f, \quad t > 0.$

(6) The basic properies of $e^{\Delta t}$ can be proved for $e^{-Lt}$, where $L$ is a strictly elliptic divergence form operator:

$$L \ = \ -\sum_{j,k} \frac{\partial}{\partial x_j} a_{jk}(x) \frac{\partial}{\partial x_k}, \tag{9.4}$$

where $\sum_{j,k} a_{jk}(x)\xi_j\xi_k \geq \theta\|\xi|^2$, where $\theta > 0$ is independent of $x$ and $\xi \in \mathbb{R}^n$ is arbitrary.

**Theorem 9.2** *( Nash (1958) and Aronson (1967) )*

$$e^{-Lt} f \ = \ \int K_t^L(x,y)\ f(y)\ dy$$

*and there exist constants $a, a', b, b' > 0$ depending on $\theta$ and such that for all $x, y, \in \mathbb{R}^n$, $K_t^L$ satisfies Gaussian upper and lower bounds:*

$$a'\ t^{-\frac{n}{2}}\ \exp(-b'\frac{|x-y|^2}{t}) \ \leq\ K_t^L(x,y) \ \leq\ a\ t^{-\frac{n}{2}}\ \exp(-b\frac{|x-y|^2}{t})$$

## 9.2 Inhomogeneous Heat Equation - DuHamel's Principle

Consider the scalar ODE

$$\frac{du}{dt} \ = \ au \ + \ f(t), \quad u(0) = u_0$$

where $a$ denotes a constant. Multiplying by the *integrating factor* $e^{-at}$, the resulting equation can be rewritten as

$$\frac{d}{dt}(e^{-at}u(t)) \ = \ e^{-at} f(t)$$

which we can now integrate and conclude:

$$u(t) \ = \ e^{at}u_0 \ + \ \int_0^t e^{a(t-s)}\ f(s)\ ds$$

The formula generalizes immediately to the case $u(t)$ is the solution of the system of ordinary differential equations:

$$\frac{du}{dt} \ = \ Au \ + \ f(t), \quad u(0) = u_0$$

where $A$ is a constant $n \times n$ matrix and $f(t)$ is a $n \times 1$ vector function of $t$:

$$u(t) \ = \ e^{At}u_0 \ + \ \int_0^t e^{A(t-s)}\ f(s)\ ds \tag{9.5}$$

Under suitable hypotheses this can be generalized to the case where $A$ is an operator acting on a Hilbert space and $f(t)$ takes values in a Hilbert space[20].

[20] A. Pazy, *Semigroups of linear operators and application to partial differential equations*, Springer 1983

Equation (9.5) is often called DuHamel's formula or DuHamel's principle. Applying it with $A = \Delta$, we obtain that the solution to the inhomogeneous heat equation:

$$\partial_t u(t,x) \ = \ \Delta u(t,x) \ + f(t,x), \quad u(0,x) = u_0(x)$$

is given by

$$u(t,x) = K_t \star u_0 \ + \ \int_0^t K_{t-s} \star f(s)\ ds, \tag{9.6}$$

where $e^{\Delta t} g = K_t \star g$ and $K_t$ denotes the heat-kernel

$$K_t(z) \ = \ \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|z|^2}{4t}} \tag{9.7}$$

## 9.3 Heat equation on a bounded domain, $\Omega$

$$\begin{aligned} \partial_t u \ &= \ \Delta u, \quad (t,x) \in (0,\infty) \times \Omega \\ u(t=0,x) \ &= \ f(x) \\ u(t,x) \ &= \ 0, x \in \partial\Omega, \ t > 0 \end{aligned} \tag{9.8}$$

**Solution:** From section 7, we have that $L^2(\Omega)$ is spanned by the complete orthonormal set of eigenfunctions of $-\Delta$, $F_j(x), \quad j = 0,1,2,\dots$ with eigenvalues $\lambda_j > 0$:

$$\begin{aligned} -\Delta F_j(x) \ &= \ \lambda_j \ F_j(x), \ x \in \Omega \\ F_j(x) \ &= \ 0, \quad x \in \partial\Omega \end{aligned}$$

$$\int F_j(x)\ F_k(x)\ dx \ = \ \delta_{jk}$$

We write the solution $u(x,t)$ as a weighted sum of solutions of the form $e^{-\lambda_j t}\ F_j(x)$:

$$u(x,t) \ = \ \sum_{j=0}^{\infty} a_j \ F_j(x) \ e^{-\lambda_j t} \tag{9.9}$$

The formal sum (9.9), by construction, solves the heat equation and satisfies the Dirichlet boundary condition. To complete the solution, it suffices to choose the constants $a_j$ so that the initial condition $u(0,x) = f(x)$ is satisfied:

$$f(x) \ = \ u(0,x) \ = \ \sum_{j=0}^{\infty} a_j F_j(x)$$

Since the eigenfunctions $F_j$ form a complete orthonormal set we have

$$a_j \ = \ (f, F_j)_{L^2} \tag{9.10}$$

Therefore,

$$u(x,t) \ = \ \sum_{j=0}^{\infty} (f, F_j)_{L^2}\ F_j(x)\ e^{-\lambda_j t} \tag{9.11}$$

It is clear from (9.11) and the fact that $\lambda_0 < \lambda_1 < \ldots$, that each term in the expansion of $u(x,t)$ decays exponentially in $t$. The dominant behavior of $u(t,x)$ for large $t$, if $a_0 \neq 0$, is given by

$$u(t,x) \ \sim \ (f, F_0)_{L^2}\ e^{-\lambda_0 t}, \tag{9.12}$$

where $\lambda_0$ denotes the smallest eigenvalue of the $-\Delta_D$, the Laplacian on $\Omega$ with Dirichlet (zero) boundary conditions.

## 9.4 Energy inequalities and heat equation

Another, more general, approach to time-decay estimates is based on energy inequalities or energy estimates.

Multiplication of the heat equation by $u(t,x)$ and integration over $\Omega$ gives

$$\frac{1}{2}\ \frac{d}{dt}\ \int_\Omega u^2(t,x)\ dx \ = \ -\int_\Omega |\nabla u(t,x)|^2\ dx \tag{9.13}$$

By Theorem 8.2 we have the sharp form of Poincaré's inequality, valid on $H_0^1(\Omega)$:

$$\int_\Omega\ |u(x)|^2\ dx \ \le \ \frac{1}{\lambda_0(\Omega)}\ \int_\Omega\ |\nabla u(x)|^2\ dx. \tag{9.14}$$

Therefore, by (9.13)

$$\frac{d}{dt}\ \int_\Omega u^2(t,x)\ dx \ \le \ -2\lambda_0(\Omega) \int_\Omega |u(t,x)|^2\ dx \tag{9.15}$$

Thus,

$$\begin{aligned} \int_\Omega u^2(t,x)\ dx \ &\le \ e^{-2\lambda_0(\Omega)t}\ \int_\Omega u^2(0,x)\ dx \\ \|u(t)\|_{L^2} \ &\le \ e^{-\lambda_0 t}\ \|u_0\|_{L^2}, \end{aligned} \tag{9.16}$$

consistent with the asymptotic result (9.12).

It is simple to extend this approach to parabolic partial differential equation, in which, $\Delta$ is replaced by a second order strictly ellptic operator.

## 9.5 The maximum principle for the heat / diffusion equation

## 9.6 Application: Burger's equation and Shock Waves

# 10 The Wave Equation

## 10.1 One dimensional wave equation

$$\partial_t^2 u = c^2 \partial_x^2 \tag{10.1}$$

**Characteristic variables:**

$$\xi \ = \ x + ct, \qquad \eta \ = \ x - ct \tag{10.2}$$

Wave equation is equivalent to $\partial_\xi \partial_\eta\ u(\xi,\eta)\ =\ 0$.
**General Solution**

$$u(x,t)\ =\ F(x+ct)\ +\ G(x-ct) \tag{10.3}$$

consisting of right and left going waves.
**Solution to the initial value problem with data:**

$$u(x,0)\ =\ f(x),\ \ \partial_t u(x,0)\ =\ g(x) \tag{10.4}$$

**D'Alembert's solution**

$$u(x,t)\ =\ \frac{1}{2}\left(\ f(x+ct)+f(x-ct)\ \right)\ +\ \frac{1}{2c}\int_{x-ct}^{x+ct}\ g(\sigma)\ d\sigma \tag{10.5}$$

**Remarks:** Domain of Dependence, Range of Influence, Forward and Backward light cones" in $(x,t)$

## 10.2 Three dimensional wave equation

$$\partial_t^2 u = c^2 \Delta u \tag{10.6}$$

with initial data $u(x,0) = f(x)$ and $\partial_t u(x,0) = g(x)$. Introduce the spherical mean of a continuous function, $h$ on a ball of radius $r$ about a point $x$:

$$\begin{aligned} M_h(x,r)\ &=\ \frac{1}{\omega_n r^{n-1}}\ \int_{|y-x|=r}\ h(y)\ dS_y \\ &=\ \frac{1}{\omega_n}\ \int_{|\xi|=1}\ h(x+r\xi)\ dS_\xi \end{aligned} \tag{10.7}$$

The latter expression for $M_h$ permits extension of $M_h$ to all $r\in\mathbb{R}$ as an even function. Note that if $h$ is continuous then

$$\lim_{r\to 0}\ M_h(x,r)\ =\ h(x). \tag{10.8}$$

**Euler-Poisson-Darboux**

$$\frac{1}{r^{n-1}}\ \partial_r\ r^{n-1}\ \partial_r\ M_h(x,r)\ =\ \Delta_x\ M_h(x,r) \tag{10.9}$$

The operator on the LHS, acting of $M_h$, is the *radial* part of the $n$-dimensional Laplace operator.

Now consider the spherical mean of a solution of the wave equation, $M_{u(\cdot,t)}(x,r)$. We have, using (10.9), that

$$\partial_t^2\ M_{u(\cdot,t)}(x,r)\ =\ c^2\Delta_r\ M_{u(\cdot,t)}(x,r) \tag{10.10}$$

Remarkably, in the **three** space dimensions, $n=3$, this implies that $r\ M_{u(\cdot,t)}(x,r)$ satisfies the **one** dimensional wave equation!!

$$\partial_t^2\ \left(\ r\ M_{u(\cdot,t)}(x,r)\ \right)\ =\ c^2\ \partial_r^2\ \left(r\ M_{u(\cdot,t)}(x,r)\ \right)$$

with initial conditions

$$r\ M_{u(\cdot,0)}(x,r)\ =\ r\ M_f(x,r),\quad \partial_t\ M_{u(\cdot,0)}(x,r)\ =\ r\ M_g(x,r) \tag{10.11}$$

D'Alembert's formula (10.5) applies and yields an explicit formula for $rM_{u(\cdot,t)}(x,r)$. By (10.8) $\lim_{r\to 0} M_{u(\cdot,t)}(x,r) = u(x,t)$, and we obtain an expression for $u(x,t)$ in terms of the data $f$ and $g$.

The result is

**Theorem 10.1** *Consider the three-dimensional wave equation with initial data $u(x,0) = f(x)$ and $\partial_t u(x,0) = g(x)$. The solution is given by:*

$$\begin{aligned} u(x,t) &= tM_g(x,ct)\ +\partial_t\ (\ t\ M_f(x,ct)\ ) \\ &= \frac{t}{4\pi}\int_{|\xi|=1} g(x+ct\xi)dS_\xi\ +\ \frac{\partial}{\partial t}\left(\ \frac{t}{4\pi}\int_{|\xi|=1} f(x+ct\xi)dS_\xi\ \right) \\ &= \frac{1}{4\pi c^2 t}\int_{|x-y|=ct} g(y)\ dS_y\ +\ \frac{\partial}{\partial t}\left(\ \frac{1}{4\pi c^2 t}\int_{|x-y|=ct} f(y)\ dS_y\ \right) \end{aligned} \tag{10.12}$$

This is Kirchoff's formula; see Evans, page 73.

**Remarks on the solution of the three-dimensional wave equation:**

- **Loss of pointwise smoothness:** In contrast to one-dimensional wave equation, solution is not as smooth as data in a pointwise sense. This is evident upon differentiating out the expression for $u(x,t)$ on the second line of (10.12):

$$\begin{aligned} u(x,t) &= \frac{t}{4\pi}\int_{|\xi|=1} g(x+ct\xi)dS_\xi\ +\ \frac{1}{4\pi}\int_{|\xi|=1} f(x+ct\xi)dS_\xi \\ &+\ \frac{ct}{4\pi}\int_{|\xi|=1} \xi\cdot\nabla_x f(x+ct\xi)dS_\xi \end{aligned} \tag{10.13}$$

- **Conservation of energy and preservation of $L^2$ smoothness:** However, there is a preservation of $L^2$ type regularity

$$\frac{d}{dt}\int (\partial_t u)^2\ +\ c^2\ |\nabla u|^2\ dx\ =\ 0 \tag{10.14}$$

- **Diffractive spreading and time decay estimate:** $f, g$ of compact support $\Longrightarrow$ $|u(x,t)| = \mathcal{O}(t^{-1})$. This can be seen as follows. Rewrite (10.13) using the change of variables $y = x + ct\xi$. Thus, $dS_y = c^2t^2 dS_\xi$. We obtain:

$$u(x,t) = \frac{1}{4\pi c^2 t^2}\int_{|x-y|=ct} [\ tg(y) + f(y) + (y-x)\cdot\nabla_y f(y)\ ]\, dS_y \tag{10.15}$$

  Assume $f$ and $g$ are supported within $B_\rho = \{x : |x| < \rho\}$. Each integral is the form $\int_A h$, where $h$ is supported in a ball of radius $\rho$ and $A = \{y : |x-y| = ct\}$. Clearly,

$$\begin{aligned} \left|\int_A h\right| &= \left|\int_A h\ 1_{\{h\neq 0\}}\right| \le \|h\|_\infty \int_A 1_{\{h\neq 0\}} \\ &= \|h\|_\infty\ meas(A\cap\{h\neq 0\}) \\ &\le \|h\|_\infty\ meas(A\cap B_\rho) \le \|h\|_\infty\ 4\pi\rho^2 \end{aligned}$$

Applying this bound to each term in (10.15) yields the time-decaying upper bound:

$$|u(x,t)| \ \leq \ K \ \rho^2 \ t^{-1} \ \left( \ \|g\|_\infty + \|f\|_\infty + \|Df\|_\infty \ \right) \tag{10.16}$$

- **Huygen's principle:** Sharp arrival and departure of signals in odd space dimensions, $n = 1, 3$.

## 10.3 2D wave equation - Hadamard's method of descent

**Wave equation in $\mathbb{R}^2$:** $\partial_t^2 u = (\partial_{x_1}^2 + \partial_{x_2}^2) u$

View solution to the two-dimensional initial value problem as a solution to the three-dimensional initial value problem with special data which is constant in $x_3$:

$$u(x_1, x_2, x_3, 0) = f(x_1, x_2), \quad \partial_t u(x_1, x_2, x_3, 0) = g(x_1, x_2) \tag{10.17}$$

Note then that $\partial_{x_3} u(x_1, x_2, x_3, t)$ solves the three-dimensional wave equation with zero initial data. Hence, $\partial_{x_3} u(\cdot, t) = 0$ for all $t \neq 0$ and $u(x_1, x_2, x_3, t) = u(x_1, x_2, 0, t)$.

Evaluating Kirchoff's formula (10.12) we have

$$\begin{aligned} u(x_1, x_2, 0, t) \ &= \ \frac{1}{4\pi c^2 t} \int_{(y_1 - x_1)^2 + (y_2 - x_2)^2 + y_3^2 = c^2 t^2} g(y_1, y_2) dS_y \\ &+ \ \frac{\partial}{\partial t} \left( \frac{1}{4\pi c^2 t} \int_{(y_1 - x_1)^2 + (y_2 - x_2)^2 + y_3^2 = c^2 t^2} f(y_1, y_2) dS_y \right) \end{aligned}$$

We now parametrize the upper / lower hemispheres of the sphere via $y_3(y_1, y_2) = \pm\sqrt{c^2 t^2 - r^2}$, where $r^2 = (y_1 - x_1)^2 + (y_2 - x_2)^2$. Then,

$$dS_y \ = \ \sqrt{1 + \left( \frac{\partial y_3}{\partial y_1} \right)^2 + \left( \frac{\partial y_3}{\partial y_2} \right)^2} \ dy_1 dy_2 \ = \ \frac{ct}{\sqrt{c^2 t^2 - r^2}} \ dy_1 dy_2$$

Taking into account contributions from both upper and lower hemispheres, we obtain:

**Theorem 10.2** *Solution of IVP for wave equation in 2 dimensions*

$$u(x_1, x_2, t) = \frac{1}{2\pi c} \int_{r < ct} \frac{g(y_1, y_2)}{\sqrt{c^2 t^2 - r^2}} \ dy_1 dy_2 \ + \ \frac{\partial}{\partial t} \frac{1}{2\pi c} \int_{r < ct} \frac{f(y_1, y_2)}{\sqrt{c^2 t^2 - r^2}} \ dy_1 dy_2$$

See, for example, Evans, page 80.

## 10.4 Principle of causality

Computation of energy flow across light cone - See, *e.g.* Evans Theorem 6, page 84.

## 10.5 Inhomogeneous Wave Equation - DuHamel's Principle

## 10.6 Initial boundary value problem for the wave equation

**Example 1 :**
1-d wave equation defined on $x > h(t)$ where Cauchy data, $u(x,0), \partial_t u(x,0)$ is given for $t = 0,\ x \geq h(0) = 0$ and along the curve (moving boundary), *i.e.* $u(h(t), t)$ prescribed.
**Example 2:**
Example 1 with $h(t) = c_0 t,\ c > c_0.$ and $u(c_0 t, t) = \cos(\omega t)$. Obtain *Doppler effect*.

# 11 The Schrödinger equation

In this section we introduce the Schrödinger equation in two ways. First, we mention how it arises in the fundamental description of quantum atomic phenomena. We then show its role in the description of diffraction of classical waves.

## 11.1 Quantum mechanics

The hydrogen atom: one proton and one electron of mass $m$ and charge $e$. The state of the atom is given by a function $\psi(x,t)$, complex-valued, defined for all $x \in \mathbb{R}^3$ and $t \in \mathbb{R}$. $\psi$ is often called the *wave function*.
Let $\Omega \subset \mathbb{R}^3$. $|\psi(x,t)|^2\ dx$ is a probability measure with the interpretation

$$\int_D |\psi(x,t)|^2\ dx \ = \ \text{Probability}\,(\text{electron} \ \in \ \Omega \text{ at time } t)$$

Thus, we require

$$\int_{\mathbb{R}^3} |\psi(x,t)|^2\ dx \ = \ \text{Probability}\,\big(\text{electron} \ \in \mathbb{R}^3 \text{ at time } t\big) \ = \ 1$$

Given an initial wave function, $\psi_0$

$$\begin{aligned} i\hbar\ \partial_t \psi \ &= \ H\ \psi \\ H \ &= \ -\frac{\hbar^2}{2m}\ \Delta \ + \ V(x) \end{aligned} \tag{11.1}$$

Here, $\hbar$ denotes Planck's constant divided by $2\pi$. The operator $H$ is called a Schrödinger operator with potential $V$, a real-valued function determined by the nucleus. For the special case of the hydrogen atom

$$H = \ -\frac{\hbar^2}{2m}\ \Delta \ - \ \frac{e^2}{r}, \quad r = |x| \tag{11.2}$$

The *free electron* (unbound to any nucleus) is governed by the *free Schrödinger equation* ($V \equiv 0$):

$$i\hbar\ \partial_t \psi \ = \ -\frac{\hbar^2}{2m}\ \Delta\psi \tag{11.3}$$

## 11.2 Diffraction of classical waves

Propagation of waves in a homogeneous medium, described by the classical wave equation in dimensions $n \geq 2$ gives rise to the phenomenon of diffractive spreading of the wave-field. By considering a class of initial value problems with so-called *wave packet* initial conditions, we demonstrate how on "short time scales" waves propagate along classical straight-lined rays. However, on "long (but finite) time scales" waves propagate according to the (free) Schrödinger equation.

Consider the initial value problem for the wave equation:

$$\partial_t^2 \, u(x,t) \; = \; c^2 \Delta \, u(x,t)$$
$$u(x,0), \; \partial_t u(x,0) \;\; \text{given}$$

Fourier representation of the solution

$$u(x,t) \; = \; \int_{\mathbb{R}^n} e^{2\pi i \eta \cdot x} \; \hat{u}(\eta,t) \; d\eta$$
$$\hat{u}(\eta,t) \; = \; \cos(2\pi c|\eta| t) \; \hat{u}(\eta,0) \; + \; \frac{\sin(2\pi c|\eta| t)}{2\pi c|\eta|} \; \partial_t \hat{u}(\eta,0)$$

**Remark 11.1** *It is an interesting (optional) exercise to deduce the finite propagation speed of signals from the Fourier representation. Note: The Fourier representation of the solution does not make explicit central features of the solution such as finite propagation speed, domain of dependence... A proof of finite propagation speed, starting with the Fourier representation, requires arguments closely related to a Paley-Wiener Theorem from complex variables. Reference: F. John - PDEs, $4^{th}$ edition, Springer, 1982, Chapter 5, Section 2.*

**Multiscale (wave packet) initial data - an example**

$$u(x,0) \; = \; e^{2\pi i \frac{x_1}{\varepsilon}} \; f(x_1, x_\perp), \; \partial_t u(x,0) \; = \; 0, \tag{11.4}$$

where $x_\perp = (x_2, \ldots, x_n)$. We call $x_1$ the longitudinal variable and $x_\perp$ the transverse variables. We take $\varepsilon << 1$ so that this data is oscillatory (with spatial scale $\varepsilon$) with slowly varying envelope (with spatial scale of order one).

**Theorem 11.1** *The solution $u(x,t)$ is a superposition of terms of the following* **multiple scale** *type*

$$u(x,t) \; \sim \; e^{2\pi i \frac{x_1 - ct}{\varepsilon}} \; A(x_1 - ct, \varepsilon t, x_\perp) \; + \; \mathcal{O}(\varepsilon^2 t),$$

*giving a good approximation of the solution for times $t \leq o(\varepsilon^{-2})$. The pre-factor $e^{2\pi i \frac{x_1 - ct}{\varepsilon}}$ is a plane wave with rapid spatial variation propagating in the $x_1$ direction at speed $c$. This plane wave is modulated by an amplitude function, which is more slowly varying. $A(x_1 - ct; T, x_\perp)$ satisfies the free Schrödinger equation*

$$i\partial_T \; A(\cdot; T, x_\perp) \; = \; -\Delta_\perp \; A(\cdot; T, x_\perp) \tag{11.5}$$

*governing the diffraction of waves and spreading over energy on time scales $t = o(\varepsilon^{-2})$. Here, $\Delta_\perp$ denotes the Laplacian with respect to the transverse variables, $x_\perp$.*

## 11.3 Free Schrödinger - initial value problem

**Initial Value Problem**

$$i\partial_t u \; = \; -\Delta \, u, \quad u(x,0) \; = \; f(x) \tag{11.6}$$

Unlike the heat equation, $\partial_t u = -\Delta u$, which has an exponentially decaying Gaussian fundamental solution, the fundamental solution of the Schrödinger equation is a oscillatory Gaussian with no spatial decay. For this reason, the derivation of the solution to the initial value problem is more subtle. One approach is to *regularize* the Schrödinger equation by adding a small ($\varepsilon > 0$) diffusive term, which we then take to zero ($\varepsilon \to 0$).

**Regularized initial value problem** Take $\varepsilon > 0$.

$$i\partial_t u^\varepsilon \ = \ -(1-i\varepsilon)\Delta\ u^\varepsilon, \quad u^\varepsilon(x,0) \ = \ f(x) \tag{11.7}$$

$$\begin{aligned} i\partial_t \hat{u}^\varepsilon \ &= \ 4\pi^2(1-i\varepsilon)|\xi|^2\hat{u} \\ \hat{u}^\varepsilon(\xi,t) \ &= \ e^{-4\pi^2(i+\varepsilon)|\xi|^2 t}\ \hat{f}(\xi) \\ u^\varepsilon(x,t) \ &= \ \int K_t^\varepsilon(x-y)\ f(y)\ dy \\ K_t^\varepsilon(x) \ &= \ \int e^{-2\pi i x\cdot\xi}\ e^{-4\pi^2(i+\varepsilon)|\xi|^2 t}\ d\xi \\ &= \ (4\pi(i+\varepsilon)t)^{-n/2}\ e^{-\frac{|x|^2}{4t(i+\varepsilon)}} \end{aligned} \tag{11.8}$$

For $f \in L^1$, we can pass to the limit $\varepsilon \to 0^+$ and define the free Schrödinger evolution by:

$$\begin{aligned} u(x,t) \ &= \ e^{i\Delta t}f \ = \ \int\ K_t(x-y)\ f(y)\ dy \\ K_t(x) \ &= \ (4\pi i t)^{-n/2}\ e^{i\frac{|x|^2}{4t}} \end{aligned} \tag{11.9}$$

For $f \in L^2$, we can use that the Fourier transform is defined (and unitary) on all $L^2$ to define, by (11.8),

$$u(x,t) \ = \ \Big(e^{-4\pi^2 i|\xi|^2 t}\ \hat{f}(\xi)\ \Big)\check{\ }(x,t) \ = \ e^{i\Delta t}f \tag{11.10}$$

## 11.4 Free Schrödinger in $L^p$

<u>$e^{i\Delta t}f$ **for** $f \in L^1(\mathbb{R}^n)$**:**</u> In this case,

$$|u(x,t)| \ = \ \left|\int\ K_t(x-y)\ f(y)\ dy\right| \ \le \ \int |f|\ dy$$

Therefore, if $f \in L^1(\mathbb{R}^n)$, then $e^{i\Delta t}f \in L^\infty(\mathbb{R}^n)$ for $t \neq 0$ and

$$\|\ e^{i\Delta t}f\ \|_{L^\infty} \ \le \ |4\pi t|^{-\frac{n}{2}}|\ \|f\|_{L^1} \tag{11.11}$$

<u>$e^{i\Delta t}f$ **for** $f \in L^2(\mathbb{R}^n)$**:**</u> In this case

$$\int |u(x,t)|^2 dx \ = \ \int |\hat{u}(\xi,t)|^2 d\xi \ = \ \int |e^{-4\pi^2 i|\xi|^2 t}\ \hat{f}(\xi)|^2 d\xi \ = \ \int |f(\xi)|^2 d\xi$$

Thus, if $f \in L^2(\mathbb{R}^n)$, then $e^{i\Delta t} f \in L^2(\mathbb{R}^n)$ and

$$\|e^{i\Delta t} f\|_{L^2} \;=\; \|f\|_{L^2} \tag{11.12}$$

Extension to $L^p$: Suppose $f \in L^p(\mathbb{R}^n)$ with $1 \le p \le 2$. Using a theorem of M. Riesz on interpolation of linear operators[21], one can show:

**Theorem 11.2** *Let $1 \le p \le 2$ and $2 \le q \le \infty$, where $p^{-1} + q^{-1} = 1$. If $f \in L^p(\mathbb{R}^n)$, then for $t \ne 0$ $e^{i\Delta t} f \in L^q(\mathbb{R}^n)$ and*

$$\|e^{i\Delta t} f\|_{L^q} \;\le\; |4\pi t|^{-\left(\frac{n}{2}-\frac{n}{q}\right)} \; \|f\|_{L^p} \tag{11.13}$$

## 11.5 Free Schrödinger evolution of a Gaussian wave packet

Consider the evolution of a Gaussian wave-packet in one-space dimension. Let $\eta_0 = 2\pi\xi_0$.

$$\begin{aligned} i\partial_t u \;&=\; -\partial_x^2 u \\ u(x,0) \;&=\; e^{i\eta_0 x}\, e^{-\frac{x^2}{2L^2}} \;=\; g_{L,\eta_0}(x) \end{aligned} \tag{11.14}$$

Thus $u(x,0)$ is an oscillatory and localized initial condition with *carrier oscillation* period $\xi_0^{-1}$ or frequency $\xi_0$. Its evolution has an elegant and illustrative form:

**Theorem 11.3**

$$u(x,t) \;=\; e^{i\Delta t} g_{L,\eta_0} \;=\; \frac{e^{i\eta_0(x-\eta_0 t)}}{\left(1+\frac{2it}{L^2}\right)^{\frac{1}{2}}}\; e^{-\frac{(x-2\eta_0 t)^2}{2L^2\left(1+\frac{2it}{L^2}\right)}} \tag{11.15}$$

**Proof:** The Fourier representation of the solution is:

$$u(x,t) \;=\; \int e^{2\pi i\xi x}\, \hat{g}_{L,2\pi\xi_0}(\xi)\; d\xi \tag{11.16}$$

Note: $\hat{g}_{L,2\pi\xi_0}(\xi) \;=\; \hat{g}_{L,0}(\xi-\xi_0) \;=\; (2\pi)^{\frac{1}{2}}\; L\; e^{-2\pi^2 L^2(\xi-\xi_0)^2}$. Substitution into (11.16) and grinding away with such tools as completing the square yields the result.

**Remarks:**

- Phase propagates with velocity $\eta_0$, the *phase velocity*
- Energy $\sim |u(x,t)|^2$ propagates with velocity $2\eta_0$, the *group velocity*
- Solution disperses (spreads and decays) to zero as $t \uparrow$. This is seen from the general estimate (11.11) as well as the explicit solution (11.15).
- However, solution does not decay in $L^2$. The Schrödinger evolution is unitary in $L^2$; see (11.12).
- Concentrated (sharp) initial conditions ($L$ small) disperse more quickly than spread out initial conditions ($L$ large). The time scale of spreading is $t \sim L^2$.

[21]See Chapter 5 of *Introduction to Fourier Analysis on Euclidean Spaces*, E.M. Stein and G. Weiss, Princeton 1971

## 11.6 The Uncertainty Principle

Recall that $|\psi(x,t)|^2$ has the interpretation of a probability density for a quantum particle to be at position $x$ at time $t$. $|\hat{\psi}(\xi,t)|^2$ has the interpretation of a probability density for a quantum particle to be at momentum $\xi$ at time $t$.The expected value of an *observable or operator* , $A$, is formally given by[22]

$$\langle A\rangle \;=\; (\psi, A\psi) \;=\; \int \bar{\psi} A\psi \tag{11.17}$$

**Examples**

(i) $\langle X\rangle$, the average position $=\ \int x|\psi(x,t)|^2dx$.

(ii) $\langle \Xi\rangle$, the average momentum $=\ \int \xi|\hat{\psi}(\xi,t)|^2d\xi$.

(iii) $\langle |X|^2\rangle$, the variance or uncertainty in position $\;=\; \int |x|^2|\psi(\xi,t)|^2dx$.

(iv) $\langle |\Xi|^2\rangle$, the variance or uncertainty in momentum $\;=\; \int |\xi|^2|\hat{\psi}(\xi,t)|^2d\xi$.

**Theorem 11.4** *(Uncertainty Inequality) Suppose $xf$ and $\nabla f$ are in $L^2(\mathbb{R}^n)$. Then,*

$$\int |f|^2 \;\le\; \frac{2}{n}\left(\int |xf|^2\right)^{\frac{1}{2}}\left(\int |\nabla f|^2\right)^{\frac{1}{2}}$$

*or equivalently*

$$\int |f|^2 \;\le\; \frac{4\pi}{n}\left(\int |xf|^2\right)^{\frac{1}{2}}\left(\int |\xi\hat{f}|^2\right)^{\frac{1}{2}} \tag{11.18}$$

**Exercise:** (a) Prove the uncertainty inequality, using the pointwise identity

$$x\cdot\nabla|f|^2 \;=\; \nabla\cdot(x|f|^2) \;-\; n\,|f|^2,$$

(b) Prove that the inequality (11.18) is sharp in the sense equality is *attained* for the Gaussian $f(x) = \exp(-|x|^2/2)$.

Applying the uncertainty inequality (11.18) to a solution of the Schrödinger equation, with initial condition $\|\psi(\cdot,0)\|_{L^2} = 1$ and we have,

$$1 \;=\; \int |\psi(x,0)|^2dx = \int |\psi(x,t)|^2dx \le \frac{4\pi}{n}\left(\int |x\psi(x,t)|^2\right)^{\frac{1}{2}}\left(\int |\xi\hat{\psi}(\xi,t)|^2\right)^{\frac{1}{2}} \tag{11.19}$$

The latter, can be written as

$$\frac{n}{4\pi} \;\le\; \sqrt{\langle |X|^2\rangle(t)}\;\sqrt{\langle |\Xi|^2\rangle(t)} \tag{11.20}$$

and is called *Heisenberg's uncertainty principle.*

[22] We proceed formally, without any serious attention to operator domains *etc.* For a fully rigorous treatment, see M. Reed and B. Simon, Modern Methods of Mathematical Physics - Volumes 1-4

# 12 Background

## 12.1 Linear Algebra

**Solvability of linear systems:** $A\vec{x} = \vec{b}$, where $A$ denote an $m \times n$ matrix.

**Theorem 12.1** *(i) The system of $m$ equations in $n$ unknowns: $A\vec{x} = \vec{b}$ is solvable if and only if $\langle \vec{z}, \vec{b} \rangle \equiv \vec{z} \cdot \vec{y} = 0$ for all $\vec{z}$ such that $A^T \vec{z} = 0$. That is, a solution exists if and only if $\vec{b}$ is orthogonal to the null space of the transpose of $A$.*
*(ii) Thus, either $A\vec{x} = \vec{b}$ has a unique solution or it has it has an $n - r(A)$ parameter family of solutions, where $r(A)$ denotes the rank of the matrix $A$.*

**The eigenvalue problem:** $A\vec{x} = \lambda\vec{x}$

*Diagonalization:* Let $\lambda_j, \;\; j = 1, \ldots, n$ denote the $n$ eigenvalues of $A$, a $n \times n$ (square) matrix. This list of eigenvalues may have repetitions. If $A$ has $n$ linearly independent eigenvectors $v_1, \ldots, v_n$, then we say that $A$ has a complete set of eigenvectors. Denote by $V$ the $n \times n$ matrix whose $j^{th}$ column in $v_j$. Then, $AV = V\Lambda$, where $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, is the diagonal matrix with eigenvalues along the diagonal. Note that $A = V\Lambda V^{-1}$ or $\Lambda = V^{-1}AV$. We say $A$ is diagonalized by $V$.

*Symmetric matrices:* The eigenvalues of a symmetric matrix are real. Moreover, for any symmetric matrix $A$, there exists a complete set of eigenvectors $v_1, \ldots, v_n$, which form an orthonormal set, *i.e.* $\langle v_i, v_j \rangle = \delta_{ij}$. In other words, the matrix $V$ whose columns are the eigenvectors of $A$ satisfies $V^T V = V V^T$, *i.e.* $V$ is an orthogonal matrix.

## 12.2 Calculus

Let $\Omega$ denote an open subset of $\mathbb{R}^n$ and $\partial\Omega$ denote its boundary, assumed smooth.

**Definition 12.1** *$C^k(\Omega)$ is the set of all functions $f(x)$, defined on $\Omega$, whose derivatives up to order $k$ are continuous functions on $\Omega$.*

**Definition 12.2** *A $C^k$ vector field on $\mathbb{R}^n$ is a vector function $u : \mathbb{R}^n \to \mathbb{R}^n$, mapping $x \in \mathbb{R}^n$ to $u(x) \in \mathbb{R}^n$. Furthermore, the components of $u_j, \; j = 1, \ldots, n$ of $u$ are each $C^k(\mathbb{R}^n)$. If $k \geq 1$, then the Jacobian matrix of the vector field $D_x u(x)$ is well-defined. It is the matrix $(\; \partial_{x_i} u_j(x) \;), \;\; 1 \leq i, j \leq n$.*

**Definition 12.3** *Given a $C^1$ vector field on $\mathbb{R}^n$, we can define its divergence:*

$$div\ u \; = \; \nabla \cdot u \; = \; \sum_{i=1}^{n} \partial_{x_i} u_i(x) \tag{12.1}$$

The following profound generalizations of the fundamental theorem of calculus and the integration by parts formulae play a very important role in the study of PDEs.

**Theorem 12.2** *Gauss-Green Theorem*
*Let $u$ be a $C^1$ vector field defined on $\Omega$, a subset of $\mathbb{R}^n$ with smooth boundary. Then,*

$$\int_\Omega \nabla \cdot u \ dx \ = \ \int_{\partial\Omega} u \cdot n \ dS, \tag{12.2}$$

*where $n$ denotes the unit normal, pointing outward from the region $\Omega$.*
*Applying (12.2) to the vector field $V(x)\delta_{ij}$, where $V$ is a scalar $C^1$ function, we obtain:*

$$\int_\Omega \partial_{x_i} V(x) \ dx \ = \ \int_{\partial\Omega} V \ n^i \ dS \tag{12.3}$$

Substitution of $U(x)\ V(x)$ for $V(x)$ in (12.3) we obtain:

**Theorem 12.3** *Integration by parts*
*Let $U(x)$ and $V(x)$ denote scalar $C^1$ functions*

$$\int_\Omega \partial_{x_i} U \ V \ dx \ = \ -\int_\Omega U \ \partial_{x_i} V \ dx \ + \ \int_{\partial\Omega} UV \ n^i \ dS \tag{12.4}$$

**Theorem 12.4** *Change of variables*
*Suppose $f : y \mapsto f(y)$ is a continuous function on the region $\Omega \subset \mathbb{R}^n$. Let $\Omega_0$ denote another region and $F : x \mapsto y = F(x)$ a mapping from $\Omega_0$ to $\Omega$, such that (i) $F \in C^1$ and (ii) $F$ is one to one and onto. Then,*

$$\int_\Omega f(y) \ dy \ = \ \int_{\Omega_0 = F^{-1}(\Omega)} f(F(x)) \ | \det D_x F(x) \ | \ dx \tag{12.5}$$

**Theorem 12.5** *Implicit Function Theorem*
*Consider a function $F : \mathbb{R}^n_x \times \mathbb{R}^m_\sigma \to \mathbb{R}^m, \quad (x,\sigma) \mapsto F(x,\sigma)$, where $F$ is sufficiently smooth. We wish to understand the solution set of $F(x,y) \ = \ 0$. Suppose*
*(i) Solvability at a point $(x_0,\sigma_0)$: $F(x_0,\sigma_0) \ = \ 0$.*
*(ii) Non-vanishing of the Jacobian determinant: Let $D_\sigma F(x,\sigma)$ denote the $m$ by $m$ Jacobian matrix of first partials: $\partial_{\sigma_k} F_l(x,\sigma), \quad k,l \ = \ 1,\ldots,n$. We assume that $D_\sigma(x_0,\sigma_0)$ is nonsingular, i.e.* $\det \ D_\sigma F(x_0,\sigma_0) \neq 0$
*Then, defined in an open neighborhood of $x_0 \in \mathbb{R}^n$ , $U$, is a smooth function, $\sigma = g(x)$, such that $g(x_0) = \sigma_0$ and*

$$F(x, g(x)) \ = \ 0, \quad \text{for all} \quad x \in U. \tag{12.6}$$

*Therefore, the equation $F(x,\sigma) = 0$, locally defines a surface, which can be represented as a graph of a function $\sigma = g(x)$.*

**Theorem 12.6** *The Inverse Function Theorem*
*Assume $f : \mathbb{R}^m \to \mathbb{R}^m$ is a smooth function and such that*
*(i) $y_0 \ = \ f(x_0)$*
*(ii) The $n$ by $n$ Jacobian matrix, $D_x f(x_0)$ is non-singular, i.e.* $\det D_x f(x_0) \neq 0$.
*Then, defined in a neighborhood $U$ of $y_0$, there is function $g : \mathbb{R}^m \to \mathbb{R}^m$, such that $x_0 = g(y_0)$ and for all $y \in U$, $f(g(y)) = y$.*

**Proof of the Inverse Function Theorem:** Apply the implicit function theorem to the function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ defined by $F(x, y) \equiv f(x) - y$ and the equation $F(x, y) = 0$.

**References:** See for example,

(1) R.C. Buck, Advanced Calculus", McGraw-Hill

(2) Appendix C of L.C. Evans, Partial Differential Equations, AMS - Graduate Studies in Mathematics, Volume 19

## 12.3 Local existence theorem for ordinary differential equations (ODEs)

Consider the system of $n-$ ordinary differential equations:

$$\frac{dx(t)}{dt} = f(x(t), t), \quad x(t_0) = \alpha, \tag{12.7}$$

where $f$ is a smooth (sufficiently differentiable) vector function $\mathbb{R}^n \times \mathbb{R}^1 \to \mathbb{R}^n$ (vector field on $\mathbb{R}^n$). The condition on $x(t)$ at $t = t_0$ is called an initial condition. The differential equation, together with the initial condition is called the *initial value problem* or IVP.

When it exists, the solution of the IVP is a function of $t$ as well the initial condition, $\alpha$. The following theorem ensures the existence of a *local* solution of the differential equation, $\phi(t; \alpha)$, which depends continuously on the initial condition, $\alpha$.

**Theorem 12.7** *There exists $T$, which depends on $\xi$ and the details of $f$, and a function $\phi(t; \alpha)$, defined and smooth for $|t - t_0| < T_*$, such that for $|t - t_0| < T_*$ the function $\phi(t; \alpha)$ satisfies the ODE and initial condition (12.7). Moreover, for any $0 < T < T_*$ there is a positive constant, depending on $T$, $r(T)$, such that*

$$\max_{|t| \le T} \| \phi(t; \alpha) - \phi(t; \eta) \| \le r(T) \| \alpha - \eta \| \tag{12.8}$$

*Here $\|y\|$ denote the norm of a vector in $\mathbb{R}^n$, e.g. the Euclidean norm: $\|y\|^2 = \sum_{j=1}^n |y_j|^2$.*

*The upper bound, (12.8), on the deviation between trajectories quantifies the continuous dependence of the solution with respect to initial conditions. In particular, on any fixed closed interval of existence, as two initial conditions are brought close to each other, the corresponding solution trajectories become uniformly close.*

**References:** See, for example,

(1) V.I. Arnol'd, Ordinary Differential Equations, MIT Press, 1980

(2) M.W. Hirsch and S. Smale, Differential Equations, Dynamical Systems, and Linear Algebra, Academic Press, 1974; Chapter 8

(3) E.A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, 1955

## 12.4 Compactness and convergence results

The simplest compactnes result, and the one to which many more sophisticated results reduce, is the classical theorem of analysis of Bolzano and Weierstrass. First, recall that a subset, $X$, of $\mathbb{R}^n$ is *compact* if any sequence in $\{x_n\} \subset X$ has a convergent subsequence, *i.e.* there exists a subsequence $\{x_{n_k}\} \subset X$ and $x_* \in X$ such that $\|x_{n_k} - x_*\| \to 0$ as $n_k \to \infty$.

**Theorem 12.8** *(Bolzano and Weierstrass) If $X \subset \mathbb{R}^n$ is closed*[23] *and bounded*[24]*, then $X$ is compact.*

**Exercise 3: Note that this theorem does not hold if $\mathbb{R}^n$ is replaced by an infinite dimensional space.** Let $X = \{ f \in L^2(\mathbb{R}^n) \ : \ \|f\|_2 \le 1 \ \}$.

(a) Show that $X$ is closed and bounded.

(b) Fix any function $f_1 \in X$ with $\|f_1\|_2 = 1$. Define $f_m(x) = m^{n/2} f_1(mx)$. Show that $\|f_m\|_{L^2} = 1$ and that $m^n f_1^2(mx) \to \delta_0$, the Dirac delta distribution, as $m \to \infty$. Thus, **no** subsequence of $f_m$ converges to a limiting $f_* \in L^2$ as $m \to \infty$; the Dirac delta distribution is not even a function!

**Exercise 4:** Let $X$ be as in Exercise 3 above and for any $m = 0, 1, 2 \ldots$ set $g_m(x) = f_1(x - m)$. The sequence $\{g_m\}$ lies in the closed and bounded subset of $L^2$, $X$. Show that $\{g_m\}$ does not have a subsequence which converges to an $L^2$ function.

**It is therefore clear from the preceding two exercises that boundedness of a sequence of *functions* itself is not sufficient to ensure compactness in the strong sense**. That is, if $\{x_n\}$ is a bounded sequence (for all $n \ge 0$, $\|x_n\| \le C$, this does not imply the existence of a subsequence $x_{n_k}$ which converges in norm.

A second classical result, whose proof uses the Bolzano-Weierstrass theorem shows how (i) boundedness **and** (ii) *equicontinuity* of a sequence implies compactness.

**Definition 12.4** *Let $C(\mathbb{R}^n)$ denote the set of continuous real-valued functions on a $R^n$ and $F$ denote a subset of $C(\mathbb{R}^n)$. The set $F$ is equicontinuous if for any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon; F)$ such that $\|x - y\|_{\mathbb{R}^n} < \delta$ implies $|f(x) - f(y)| < \varepsilon$ for all $f \in F$.*

**Theorem 12.9** *(Ascoli-Arzela Theorem) If $F \subset C(\mathbb{R}^n)$ is a family of functions which is uniformly bounded*[25] *and equicontinuous, then there is a sequence in $F$ which converges uniformly on any compact subset of $X$ to a function in $C(X)$.*

Note that although closed and bounded subsets in a Hilbert space $\mathcal{H}$ are not generally compact, THEY ARE HOWEVER *WEAKLY COMPACT*:

[23] $X$ contains all its limit points

[24] There exists $r_X > 0$ such that $X$ lies within the ball of radius $r_X$ about the origin

[25] There exists $M_F > 0$, such that $|f(x)| \le M_F$ for all $x \in \mathbb{R}^n$ and $f \in F$

**Theorem 12.10** [26] *(Weak compactness - special case of Banach -Alagolou ) If $\{x_n\}$ is a bounded sequence in a Hilbert space $\mathcal{H}$, then there is a subsequence $\{x_{n_k}\}$ and $x_* \in \mathcal{H}$ such that $(x_{n_k}, y)_{\mathcal{H}} \to (x_*, y)_{\mathcal{H}}$ for all $y \in \mathcal{H}$.*

A compactness result, rooted in the Ascoli-Arzela Theorem 12.10 and at the heart of showing that $(-\Delta_D)^{-1}$ is a self-adjoint compact operator is

**Theorem 12.11** *( Rellich compactness Lemma)*[27] *Suppose $\{u_j\}$ is a sequence in $H_0^1(\Omega)$ such that $\|u_j\|_{H_0^1(\Omega)} \le C$. Then, there exists a subsequence $u_{j_k}$ and an element $u_* \in H_0^1$ such that $\|u_{j_k} - u_*\|_{L^2} \to 0$ as $j_k \to \infty$.*

## 12.5 Fourier Transform

**Definition:** Schwartz class, $\mathcal{S}(\mathbb{R}^n)$, is defined to be the space of all functions which are $C^\infty$ and which, together with all their derivatives, decay faster than any polynomial rate. Specifically, if $f \in \mathcal{S}$, then for any $\alpha, \beta \in \mathbb{N}_0^n$, there exists a constant $C_{\alpha,\beta}$ such that

$$\sup_{x\in\mathbb{R}^n} |x^\alpha \partial_x^\beta f(x)| \ \le\ C_{\alpha,\beta}$$

For $f \in \mathcal{S}$ define the *Fourier transform*, $\hat{f}$ or $\mathcal{F}f$ by

$$\hat{f}(\xi) \ =\ \mathcal{F}f(\xi) \ =\ \int e^{-2\pi i x\cdot\xi} f(x)\ dx \tag{12.9}$$

**Proposition 12.1** *Assume $f \in \mathcal{S}$.*

*(a)* $\hat{f} \in C^\infty$ *and* $\partial^\beta \hat{f}(\xi) \ = \left[\ (-2\pi i x)^\beta f(x)\ \right]\hat{}\,(\xi)$

*(b)* $\widehat{\partial^\beta f}(\xi) \ =\ (2\pi i\xi)^\beta\ \hat{f}(\xi)$

*(c)* $\hat{f} \in \mathcal{S}$. *Thus, the Fourier transform maps $\mathcal{S}$ to $\mathcal{S}$.*

**Theorem 12.12** *(Riemann-Lebesgue Lemma)*
$f \in L^1(\mathbb{R}^n) \ \implies\ \lim_{\xi\to\infty}\ \hat{f}(\xi) \ =\ 0.$

**Proof:** Approximate by step functions, for which the result can be checked. Note: no rate of decay of the Fourier transform is implied by $f \in L^1$.

**Theorem 12.13** *Let $f(x) \ \equiv\ e^{-\pi a|x|^2}$, $\Re a > 0$. Then,*

$$\hat{f}(\xi) \ =\ a^{-\frac{n}{2}}\ e^{-\pi\frac{|\xi|^2}{a}} \tag{12.10}$$

*The formula (12.10) also extends to $\Re a = 0,\ a \ne 0$.*

**Proof:** Write out definition of $\hat{f}$. Note that the computation factors into computing the Fourier transform of $n$ independent one dimensional Gaussians. For the one-dimensional Gaussian, complete the square in the exponent, deform the contour using analyticity of the integrand, and finally use that $\int_{\mathbb{R}} e^{-\pi y^2} dy = 1$.

[26] See Theorem 3, Appendix D, page 639 of L.C. Evans *PDEs*

[27] This is a special case of the more general compactness theorem of Rellich and Kondrachov. For a proof see L.C. Evans, *PDEs*; see section 5.2, Theorem 1 on page 272. Its proof relies on the Arzela-Ascoli theorem, given above.

## 12.6 Fourier inversion on $\mathcal{S}$ and $L^2$

**Definition 12.5** *For $g \in \mathcal{S}$ define $\check{g}$*

$$\check{g}(x) \equiv \int e^{2\pi i\xi\cdot x} g(\xi)\, d\xi = \hat{g}(-x)$$

**Proposition 12.2**

$$\int \hat{f}\ g\ dx = \int f\ \hat{g}\ dx \tag{12.11}$$

**Proof:** Interchange orders of integration (Fubini's Theorem)

**Remark 12.1** *Equation (12.11) is used to define the Fourier transform of a distribution. Namely, if $T$ is a distribution, then the distribution $\hat{T}$ is defined to be that distribution whose action on $C_0^\infty(\mathbb{R}^n)$ functions is:*

$$\hat{T}(\phi) = T[\hat{\phi}]$$

*Thus, for example we can compute the Fourier transform of the delta function as follows:*

$$\begin{aligned}\hat{\delta_x}[\phi] &= \delta_x[\hat{\phi}] = \hat{\phi}(x)\\ &= \int_{\mathbb{R}^n} e^{-2\pi ixy} f(y)dy.\end{aligned}$$

*Thus, we identify the $\hat{\delta_x}$ with the $L^1_{loc}$ function $e^{-2\pi ix\cdot y}$, i.e. $\hat{\delta_x} = e^{-2\pi ix\cdot y}$.*

**Theorem 12.14** *Fourier inversion formula Assume $f \in \mathcal{S}$. Then,*

$$f \in \mathcal{S} \implies \check{g}(x) = f(x), \quad \text{where } g(\xi) = \hat{f}(\xi) \tag{12.12}$$

**Proof:** We shall prove Fourier inversion in the following sense.

$$\lim_{\varepsilon\to 0} \int e^{-\pi\varepsilon^2|\xi|^2}\ e^{2\pi ix\cdot\xi}\ \hat{f}(\xi)\ d\xi = f(x)$$

For any $\varepsilon > 0$ define

$$\phi_\varepsilon(\xi) = e^{2\pi x\cdot\xi\ -\ \pi\varepsilon^2|\xi|^2}$$

whose Fourier transform is

$$\hat{\phi}_\varepsilon(y) = \frac{1}{\varepsilon^n}\ e^{-\pi\frac{|x-y|^2}{\varepsilon^2}} = \frac{1}{\varepsilon^n} g\left(\frac{x-y}{\varepsilon}\right) \equiv g_\varepsilon(x-y)$$

Now,

$$\begin{aligned}\int e^{-\pi\varepsilon^2|\xi|^2}\ e^{2\pi ix\cdot\xi}\ \hat{f}(\xi)\ d\xi &= \int \phi_\varepsilon(\xi)\ \hat{f}(\xi)\ d\xi\\ &= \int \hat{\phi}_\varepsilon(y)\ f(y)\ dy\\ &= \int g_\varepsilon(x-y)\ f(y)\ dy \ \to\ f(x)\end{aligned}$$

as $\varepsilon \downarrow 0$ because $g_\varepsilon$ is an *approximation of the identity*[28].

[28] *Approximation of the identity*: Let $K(x) > 0$ and $\int K(x)dx = 1$. Define $K_N(x) = N^n K(Nx)$, $N \geq 1$.

**Theorem 12.15** *Plancherel Theorem*

$$f \in \mathcal{S} \quad \Longrightarrow \quad \int |f(x)|^2 \, dx \; = \; \int |\hat{f}(\xi)|^2 \, d\xi$$

*That is the Fourier transform preserves the $L^2$ norm on $\mathcal{S}$ ($\|\hat{f}\|_2 \; = \; \|f\|_2$).*

**Corollary 12.1** *The Fourier transform can be extended to a unitary operator defined on all $L^2$ such $\|\hat{f}\|_2 \; = \; \|f\|_2$.*

**Proof:** $\mathcal{S}$ is dense in $L^2$. If $f \in L^2$, there exists a sequence $f_j \in \mathcal{S}$ such that $\|f_j - f\|_2 \to 0$. Define $\hat{f} = \lim_{j\to\infty} \; \hat{f}_j$.

---

Let $f$ be a bounded and continuous function on $\mathbb{R}^n$ and consider the *convolution*

$$K_N \star f(x) \; = \; \int K_N(x-y) \; f(y) \; dy \tag{12.13}$$

Prove: $K_N \star f(x) \to f(x)$ uniformly on any compact subset, $C$, of $\mathbb{R}^n$ as $N \uparrow \infty$, *i.e.* $\max_{x\in C} | \; [K_N \star f](x) - f(x)| \to 0$.