REVIEW OF ANALYSIS FOR STUDENTS OF MICROLOCAL ANALYSIS

CHARLES L. EPSTEIN

Version: 0.5; Revised: 9–24–97; Run: September 24, 1997

1. Introduction

This semester we will study the basic parts of microlocal analysis. These are techniques which grew out of Fourier analysis and mathematical physics that are useful for studying the properties of solutions to linear partial differential equations:

$$\sum_{|\alpha|\leq m} A_\alpha(x)\partial_x^\alpha u(x) = f(x). \qquad x\in\mathbf{R}^n. \tag{1}$$

If the coefficients, $\{A_\alpha(x)\}$ do not depend on x, then one can use the Fourier transform, more or less directly to solve this equation. The techniques of microlocal analysis include a calculus which allows one to efficiently quantify and handle the errors which arise from variability of the coefficients. More generally these techniques lead to a precise and quantitative analysis of the singularities of distributions. The connection between these two problems is the following:

There is often a linear operator K, such that a solution of (1) is given by $u = Kf$. Formally we can express this linear operator as an integral:

$$u(x) = Kf(x) = \int k_x(y)f(y)dy.$$

In general $k_x(y)$ is a family of distributions. The singularities of this family of distributions determine the relationships between the data, f and the solution, $u(x)$.

The elementary prototype for (1) is a finite system of linear equations:

$$Ax = y, \tag{2}$$

where $A : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, we quickly review the theory of such equations. To generalize the finite dimensional theory of (2) to an infinite dimensional context one first needs to put topologies on the domain and range of A and consider the continuity properties of linear transformations in this context. To that end we review the elements of functional analysis and Fredholm theory. Finally we consider the elementary theory of the Fourier integral. At that point we will be prepared to begin the study of microlocal analysis in earnest and begin to follow the lecture notes of R. Melrose. These can be obtained at http://www-math.mit.edu/~rbm/lecture_notes. html

2. Solving Linear Equations

The prototypic finite dimensional problem we would like to solve is a system of linear equations:

$$Ax = y \tag{3}$$

where $A : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation. In case $m = n$ there is a simple criterion for (3) to be solvable for any choice of y:

Theorem 2.1. *: If $A : \mathbf{R}^n \to \mathbf{R}^n$ is linear then (0.1) has a solution for every $y \in \mathbf{R}^n$ if and only if the only solution to*

$$Ax = 0$$

is $x = 0$. More geometrically: A is surjective if and only if it is injective.

Note that this condition for solvability implies that the solution to (3) is unique. The theorem is purely algebraic in character but we can also consider the dependence of the solution, x on the data, y. To that end we need to introduce a topology on $\mathbf{R}^n$. For the purposes of studying linear equations the natural topology on $\mathbf{R}^n$ is that defined by a norm. A norm is a function, $||\cdot|| : \mathbf{R}^n \to \mathbf{R}^n$ which satisfies the conditions:

$$||x|| \geq 0 \quad \forall x \in \mathbf{R}^n,\ ||x|| = 0 \,\text{iff}\, x = 0 \tag{4}$$

$$||x + y|| \leq ||x|| + ||y|| \qquad \text{(triangle inequality)}, \tag{5}$$

$$\forall x \in \mathbf{R}^n,\ \lambda \in \mathbf{R} \quad ||\lambda x|| = |\lambda|\, ||x||. \tag{6}$$

For example if $p \geq 1$ then

$$||x||_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$$

defines a norm. Note that

$$\lim_{p\to\infty} ||x||_p = \max\{|x_i| : \quad i = 1 \dots n\} = ||x||_\infty.$$

This also a norm

Exercise 2.1. : Show that if $1 \leq p,\ q \leq \infty$ then there are constants c, C such that

$$c||x||_q \leq ||x||_p \leq C||x||_q, \tag{7}$$

Briefly: $||\cdot||_q$ and $||\cdot||_p$ are equivalent norms.

A consequence of (7) is that all the norms, $\{||\cdot||_p\}$ define the same topology on $\mathbf{R}^n$; in fact all norms on $\mathbf{R}^n$ are equivalent and therefore define the same topology. Note in particular that the set $\{x : ||x|| \leq 1\}$ is compact. An especially useful norm is

$$||x||_2^2 = \sum_{i=1}^n x_1^2.$$

This is usually called the Euclidean norm. What distinguishes this norm is that it is defined by an inner product. An inner product is a mapping

$$\langle \cdot, \cdot \rangle : \quad \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$$

such that

$$\langle x, y \rangle = \langle y, x \rangle, \tag{8}$$
$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \tag{9}$$
$$\langle ax, y \rangle = a\langle x, y \rangle, \tag{10}$$
$$\langle x, x \rangle \geq 0 \text{ with equality only if } x = 0. \tag{11}$$

In the case at hand

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

Evidently

$$||x||_2^2 = \langle x, x \rangle.$$

A collection of vectors $\{e_1 \dots e_n\}$ is a basis for $\mathbf{R}^n$ iff every vector x has a unique representation as

$$x = \sum_{i=1}^{n} x_i e_i.$$

We say that a basis is orthonormal if

$$\langle e_i, e_j \rangle = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

Exercise 2.2. Suppose we are given a basis, $\{f_1 \dots f_n\}$ for $\mathbf{R}^n$ show that there is an orthonormal basis, $\{e_1 \dots e_n\}$ such that for each $1 \leq j \leq n$

$$\{\sum_{i=1}^{j} x_i f_i : \quad x_1 \in \mathbf{R} \quad i = 1 \dots j\} =$$

$$\{\sum_{i=1}^{j} x_i e_i : \quad x_i \in \mathbf{R} \quad 2 = 1 \dots j\}.$$

Give an algorithm to construct $\{e_i\}$ from $\{f_i\}$.

Let $\{e_i \dots e_n\}$ be an orthonormal basis for $\mathbf{R}^n$. For each $x \in \mathbf{R}$ we can express $x = \sum_{i=1}^{n} x_i e_i$ and thus

$$Ax = \sum_{i=1}^{n} x_i A e_i.$$

By the triangle inequality

$$||Ax|| \leq \sum_{i=1}^{n} |x_1| \, ||Ae_i||$$

So if $M = \max\{||Ae_1||, \dots ||Ae_n||\}$ then

$$||Ax|| \leq M||x||_1.$$

Since all norms on $\mathbf{R}^n$ are equivalent there is a constant C s.t. $||x||_1 \leq C||x||$ and therefore we've shown:

Lemma 2.1. *If $A : \mathbf{R}^n \to \mathbf{R}^n$ is a linear transformation and $||\cdot||$ is a norm on $\mathbf{R}^n$ then there is a constant C such that*

$$||Ax|| \leq C||x|| \qquad \forall x \in \mathbf{R}^n.$$

Corollary 2.1. *: If $A : \mathbf{R}^n \to \mathbf{R}^n$ is a linear transformation then $A : \mathbf{R}^n \to \mathbf{R}^n$ is continuous.*

Proof. Since A is linear $Ax - Ay = A(x-y)$, thus $||Ax - Ay|| = ||A(x-y)|| \leq C||x-y||$, for some constant C. □

Proposition 2.1. *: If $A : \mathbf{R}^n \to \mathbf{R}^n$ is a linear transformation then A is surjective iff there is a constant, $C > 0$ such that*

$$c||x|| \leq ||Ax||. \tag{12}$$

Proof. By theorem 1.1 A is surjective iff A is injective. Thus $||Ax|| \neq 0, \quad \forall x \neq 0$. Since $S = \{x : ||x|| = 1\}$ is compact and A is continuous the function $x \to ||Ax||$ assumes its minimum value at some point of $x_0 \in S_1$. As $||x_0|| = 1$ it is clear that $||Ax_0|| = c > 0$. For $X \in S_1$ we have

$$c||x|| \leq ||Ax||,$$

since $||\lambda x|| = |\lambda|\,||x||$ and $A\lambda x = \lambda Ax$ this shows that if A is surjective then (12) holds. If (12) holds then evidently $Ax = 0$ iff $x = 0$ and so by Theorem 1.1 A is surjective. □

Corollary 2.2. *: If $A : \mathbf{R}^n \to \mathbf{R}^n$ is a surjective linear transformation then*

$$||A^{-1}x|| \leq \frac{1}{c}||x|| \tag{13}$$

where c is the constant appearing in (13).

In summary we see that

$$\max_{\{||x||=1\}} ||Ax||$$

gives a quantitative measure of the continuity of A whereas

$$\min_{\{||x||=1\}} ||Ax||$$

gives a quantitative measure of the continuity of A^{-1}.

Now we consider the equation (3) when $n \neq m$. If for example $m > n$ then it seems quite unlikely that (3) could be solvable for arbitrary $y \in \mathbf{R}^m$. In order to obtain conditions on y for (3) to be solvable we need to consider the space of linear functions on $\mathbf{R}^n$. A map $\ell : \mathbf{R}^n \to \mathbf{R}$ is linear if

$$\ell(x+y) = \ell(x) + \ell(y), \tag{14}$$

$$\ell(ax) = a\ell(x).$$

The set of such linear maps clearly is itself a linear space which we denote by $(\mathbf{R}^n)'$. For example, if $y \in \mathbf{R}^n$ then

$$\ell_y(x) = \langle x, y \rangle$$

defines an element of $(\mathbf{R}^n)'$. In fact it is quite easy to prove

Proposition 2.2. *: The map $y \to \ell_y$ is an isomorphism of the linear spaces $\mathbf{R}^n$ and $(\mathbf{R}^n)'$.*

If $A : \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation then for each $y \in \mathbf{R}^m$ we can think of the map

$$x \to \langle Ax, y \rangle$$

as defining an element of $(\mathbf{R}^n)'$. Hence by Proposition 1.2 there is a unique vector $z_y \in \mathbf{R}^n$ such that

$$\langle Ax, y \rangle = \langle x, z_y \rangle.$$

Exercise 2.3. : Show that the map $y \to z_y$ is a linear transformation.

We call this linear transformation the transpose, dual or adjoint of A, it is denoted A^t:

$$\langle Ax, y \rangle = \langle x, A^t y \rangle.$$

For any linear transformation, $A : \mathbf{R}^n \to \mathbf{R}^m$ we define

image of $A = \text{Im}\, A = \{Ax : x \in \mathbf{R}^n\}$,

kernel of $A = \ker A = \{x : Ax = 0\}$,

cokemel of $A = \text{coker}\, A = \mathbf{R}^m / \text{Im}\ A$.

Evidently if $y \in \text{Im}\ A$ and $z \in \text{Ker}\ A^t$ then

$$\langle y, z \rangle = \langle Ax, z \rangle = \langle x, A^t z \rangle = 0.$$

Thus we have a necessary condition for $Ax = y$ to be solvable. This turns out also to be a sufficient condition:

Theorem 2.2. *: The equation $Ax = y$ is solvable iff $\langle y, z \rangle = 0$ for all $z \in \ker\ A^t$.*

Exercise 2.4. : A map $B : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ is called a bilinear form if

$$B(x + y, z) = B(x, z) + B(y, z)$$

$$B(x, y + z) = B(x, y) + B(x, z)$$

$$B(ax, y) = B(x, ay) = aB(x, y).$$

A bilinear form is non degenerate provided $B(x, y) = 0$ for all $x \in \mathbf{R}^n$ iff $y = 0$.

(a) Show that $y \to B(\cdot, y)$ defines an isomorphism, $\mathbf{R}^n \to (\mathbf{R}^n)'$
(b) If B_1 and B_2 are nondegenerate bilinear on $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively forms show that there is a uniquely defined linear transformation A^t such that $B_2(Ax, y) = B_1(x_y A^t y)$.
(c) Show that $Ax = y$ is solvable iff $B_z(y, z) = 0$ for every $z \in \text{Ker}\ A^t$.

Exercise 2.5. : If $A : \mathbf{R}^n \to \mathbf{R}^m$ then show that:

$$\dim \operatorname{Im} A = n - \dim \operatorname{Ker} A,$$

$$\dim \operatorname{Im} A = n - \dim \operatorname{Ker} A'$$

therefore:

$$\dim \operatorname{Ker} A - \dim \operatorname{Ker} A' = n - m.$$

Exercise 2.6. : If $S \leq \mathbf{R}^n$ is a subspace and $||\cdot||$ is a norm on $\mathbf{R}^n$ then we can define a function $N([x])$ on the quotient vector space $\mathbf{R}^n/S$ by setting:

$$N([x_0]) = \inf_{x \in [x_0]} ||x||.$$

Prove that $N(\cdot)$ is a norm on $\mathbf{R}^n/S$

Exercise 2.7. : Let $A : \mathbf{R}^n \to \mathbf{R}^m$ be a linear map and suppose that $||\cdot||$ denotes a norm on $\mathbf{R}^n$ or $\mathbf{R}^m$. Show that there exists a constant C such for $y \in \operatorname{Im} A$ there is an $x \in \mathbf{R}^n$ with $Ax = y$ and $||x|| \leq C||y||$.

Good references for this material are

1. *Linear Algebra* by Peter D. Lax
2. *Intro. to Matrix Analysis* by Richard Bellman
3. *Calculus*, vol II by Tom M. Apostol.

3. Basic Functional Analysis

In finite dimensions the problem of solving linear equations is purely algebraic. That is: there is no necessity to introduce a topology to give necessary and sufficient conditions for the solvability of $Ax = y$. In infinite dimensions there is a simi lar analysis but it requires a topology on the domain and range of the linear map. In finite dimensions there is a unique norm topology, this is closely related to the fact that the unit sphere, with respect to any norm, is compact. To compare two norms , $||\cdot||_1, \quad ||\cdot||_2$ we simply compute

$$c_1 = \inf_{\{||x||_1=1\}} ||x||_2 \quad \text{and} \quad c_2 = \sup_{\{||x||_1=1\}} ||x||_2$$

Then

$$c_1||x||_1 \leq ||x||_2 \leq c_2||x||_1.$$

In infinite dimensions the unit sphere is never compact and there are many different normed linear spaces. Recall that in the analysis of $Ax = y$ the dual space $(\mathbf{R}^m)^*$ played an important role. This feature becomes even more pronounced in infinite dimensions.

Let's briefly consider normed linear spaces in general. Let X be a vector space. We need to specify an underlying field, the field of scalars. It will usually be $\mathbf{C}$ but occasionally we use $\mathbf{R}$. A norm is a map:

$$||\cdot|| : x \to \mathbf{R} \qquad \text{such that}$$

$$||x|| \geq 0 \text{ with equality iff } x = 0, \tag{15}$$
$$||x+y|| \leq ||x|| + ||y|| \text{ (triangle inequality) }, \tag{16}$$
$$||\lambda x|| = |\lambda| \, ||x|| \qquad x \in X, \quad \lambda \text{ a scalar.} \tag{17}$$

A norm defines a distance

$$d(x,y) = ||x-y||$$

because (16) implies the triangle inequality for d. The distance in defines a metric topology on X, a basis for the open sets is given by the metric balls $B_r(x) = \{y \in X | d(x,y) < r\}$.

Definition 3.1. : If $(X, || \cdot ||)$ is complete as a topological space then we say that X is a Banach space.

Examples: 1) $C^0[0,1]$ = continuous functions on $[0,1]$ with $||f|| = \sup\{|f(x)| : x \in [0,1]\}$

2) $\ell^p = \{(a_1, a_2, \ldots) | \sum |a_i|^p < \infty$ with $||a||_p = (\sum_{i=1} |a_1|^p)^{1/p}\}$ is a Banach space for $a \leq p \leq \infty$.

3) $L^p(\mathbf{R}^n)$ = Equivalence classes of bounded measurable functions on $\mathbf{R}^n$ such that $\int_{\mathbf{R}^n} |f|^p dx < \infty$ with

$$||f||_p = \left(\int_{\mathbf{R}^n} |f|^p dx \right)^{1/p},$$

$f \sim g$ if $\{x | f(x) \neq g(x)\}$ has measure zero.

4) $H^p(D_1)$ – holomorphic functions on the unit disk such that

$$||f||_p = \sup_{0<r<1} \left[\int |f(re^{i\theta})|^p d\theta \right]^p < \infty \qquad 1 \leq p \leq \infty.$$

Note that for each $1 \leq p < \infty$, $\| \cdot \|_p$ defines a norm on $C_c^0(\mathbf{R})$, the compactly supported, continuous functions on $\mathbf{R^n}$. For different values of p the completions of C_c^0 with respect to these norms give inequivalent Banach spaces In the sequel we mostly consider the spaces $L^p(\mathbf{R}^n)$. For this purpose Hölder's inequality is a fundamental tool:

Theorem 3.1 (Hölder's Inequality)**.** *Let $1 \leq p, q \leq \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, if $f \in L^p(\mathbf{R}^n), g \in L^q(\mathbf{R}^n)$ then*

$$|\int_{\mathbf{R}^n} fg dx| \leq ||f||_p \, ||g||_q.$$

Another important space of functions is $C_c^\infty(\mathbf{R}^n)$ = {infinitely differentiable functions with compact support }. This space cannot be endowed with a norm, or even a metric topology. One thing which makes this space so useful is:

Proposition 3.1. *For $1 \leq p < \infty$, $\quad C_c^\infty(\mathbf{R}^n)$ is a dense subspace of $L^p(\mathbf{R}^n)$.*

To prove this statement we use convolution. For $f, g \in C_c^\infty(\mathbf{R}^n)$ we define

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

Note in fact that if $f \in C_c^\infty(\mathbf{R}^n)$ and g is locally integrable that $f*g(x)$ is defined and is evidently an infinitely differentiable function. We can prove the following estimate:

Lemma 3.1. *If* $f \in C_c^\infty(\mathbf{R}^n), \quad g \in L^p(\mathbf{R}^n)$ *then*

$$||f * g||_{L^p} \leq ||f||_{L^1}||g||_{L^p}.$$

Proof. We can assume without loss of generality that g has compact support. This allows us to easily justify the following manipulations:

$$\int |f * g|^p dx = \int |\int f(x-y)g(y)dy|^p dx.$$

We apply Hölder's inequality to obtain:

$$\int |f * g|^p dx \leq \int [\int |f(x-y)|dy]^{p/q} \int |f(x-y)||g(y)|^p dy\, dx$$

$$= ||f||_1^{p/q} \int \int |f(x-y)||g(y)|^p dy\, dx$$

By Fubini's theorem.

$$= ||f||_1^{p/q} \int \int |f(x-y)|dx|g(y)|^p dy$$

$$= ||f||_1^{\frac{p}{q}+1} ||g||_p^p$$

Thus $||f * g||_p \leq ||f||_1 ||g||_p$. For arbitrary $g \in L^p(\mathbf{R}^n)$ we choose a sequence $\{g_n\} \leq L^p(\mathbf{R}^n)$ of functions with compact support that converge to g in L^p. We observe that

$$||f * g_n - f * g_m||_p = ||f * (g_n - g_m)||_p \leq ||f||_q ||g_n - g_m||_p,$$

and therefore $\{f * g_n\}$ is an L^p – Cauchy sequence. On the other hand it converges pointwise to $f * g$ and therefore $f * g \in L^p(\mathbf{R}^n)$ with

$$||f * g||_{L^p} \leq ||f||_{L^1}||g||_{L^p}.$$

□

We can now prove Proposition 3.1:

Proof of Proposition 3.1: Let $f \in L^p(\mathbf{R}^n)$ and choose a sequence $\{f_n\}$ of continuous, compactly supported functions which converge to f in $L^p(\mathbf{R}^n)$. THe existence of such a sequence follows easily from Lusin's Theorem. Given $\varepsilon > 0$ there exists an N such that $||f - f_n||_p < \epsilon$ if $n < N$.

Choose a function $\varphi \in C_c^\infty(\mathbf{R}^n)$ such that

$$\text{supp}\varphi \leq B_1(0), \tag{18}$$

$$\varphi \geq 0, \tag{19}$$

$$\int \varphi(x)dx = 1. \tag{20}$$

And define $\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi(\frac{x}{\varepsilon})$. Observe that $\text{supp}\varphi_\varepsilon \leq B_\varepsilon(0)$, $\varphi_\varepsilon \geq 0$ and $\int \varphi_\varepsilon = 1$ for all $\varepsilon > 0$.

Claim: $\varphi_\varepsilon * f_n$ is a sequence contained in $C_c^\infty(\mathbf{R}^n)$ converging to f in $L^p(\mathbf{R}^n)$ as $\varepsilon \to 0$, $n \to \infty$. Since $\text{supp}\varphi_\varepsilon \leq B_\varepsilon(0)$ it is easy to see that $\text{supp}\varphi_\varepsilon * f_n$ is contained in the ε–neighborhood of the $\text{supp} f_n$:

$$\text{supp}_\varepsilon f_n = \{y|d(y, \text{supp} f_n) \leq \varepsilon\}.$$

Note moreover that

$$\varphi_\varepsilon * f_n(x) = \int \varphi_\varepsilon(x-y)f_n(y)dy.$$

Using this formula it is easy to show, using Lebesque dominated convergence theorem that $\varphi_\varepsilon * f_n$ is continuous, then by considering the difference quotient,

$$\frac{\varphi_\varepsilon * f_n(x + he_i) - \varphi_\varepsilon * f_n(x)}{h}$$

we easily show that $\varphi_\varepsilon * f_n$ has continuous partial derivatives with

$$\partial_{x_i}(\varphi_\varepsilon * g)(x) = \int \partial_{x_i}\varphi_\varepsilon(x-y)f_n(y)dy.$$

Repeating this argument we establish that $\varphi_\varepsilon * g$ has continuous partial derivatives of all orders and therefore belongs to $C_c^\infty(\mathbf{R}^n)$.

By the triangle inequality:

$$\begin{aligned}||\varphi_\varepsilon * f - f||_p &\leq ||\varphi_\varepsilon * f - \varphi_\varepsilon * f_n||_p + ||\varphi_\varepsilon * f_n - f_n||_p + ||f_n - t||_p \\ &\leq (1 + ||\varphi_\varepsilon||_1)||f - f_n||_p + ||\varphi_\varepsilon * f_n - f_n||_p.\end{aligned}$$

Here we use Lemma 2.1. It suffices to show that $\varphi_\varepsilon * g$ converges to g in $L^p(\mathbf{R}^n)$ for g a continuous function of compact support. For such a function we can easily complete the proof of the proposition:

$$|\varphi_\varepsilon * g(x) - g(x)| = |\int \varphi_\varepsilon(x-y)(g(y) - g(x))dy|$$

Since g is a continuous function of compact support, it is uniformly continuous thus given $\delta > 0$ there is an η such that

$$|x - y| < \eta \Rightarrow (g(x) - g(y)| < \delta \tag{21}$$

If $\epsilon < \eta$ then (21) and (18) imply that:

$$\begin{aligned}|\varphi_\varepsilon * g(x) - g(x)| &\leq \int \varphi_\varepsilon(x-y)|g(y) - g(x)|dy \\ &\leq \delta \int \varphi_\varepsilon(x-y)dy \\ &= \delta.\end{aligned}$$

On the other hand supp$\varphi_\varepsilon * g \subset \epsilon$–neighborhood of supp$g$ Therefore:

$$||\varphi_\varepsilon * g - g||_{L^p(\mathbf{R}^n)} \leq \delta |\mathrm{Vol}\,\mathrm{supp}_\varepsilon g|^{1/p}. \tag{22}$$

From (22) we obtain the desired conclusion. □

We can use this density statement to extend Lemma 3.1.

Theorem 3.2 (Hausdorff–Young inequality)**.** *If $f \in L^1(\mathbf{R}^n)$ and $g \in L^p(\mathbf{R}^n)$ $\quad 1 \leq p < \infty$ then $f * g \in L^p(\mathbf{R}^n)$ with*

$$||f * g||_p \leq ||f||_1 ||g||_p. \tag{23}$$

Proof. : Choose a sequence $\{f_n\} subset C_c^\infty(\mathbf{R}^n)$ converging to f in $L^1(\mathbf{R}^n)$. From Lemma 2.1 we have that

$$||f_n * g - f_m * g||_p \leq ||t_n - t_m||_1 ||g||_p.$$

Since $\{f_n\}$ is an L^1–Cauchy sequence it follows that $\{f_n * g\}$ is an L^p–Cauchy sequence. As L^p is complete it follows that $\{f_n * g\}$ has a limit which we denote by $f * g$. It is a simple exercise to show that this limit is independent of the choice of sequence, $\{f_n\}$ converging to f in $L^1(\mathbf{R}^n)$. As

$$||f_n * g||_p \leq ||f_n||_1 ||g||_p$$

It is evident that (23) also holds in the limit. □

There is another way to express this result: the bilinear map $C_c^\infty \times C_c^\infty \to C_c^\infty$ defined by $(f, g) \to f * g$ extends as a continuous map $L^1 \times L^p \to L^p$, satisfying (23). We will often see such statements in the theory of distributions and pseudo-differential operators.

Let us take stock, in this simple example of how this extension was accomplished:

1. We have a formula which defines $f * g$ and makes sense for f and g in certain classes of functions, e.g. $f \in C_c^\infty, \quad g \in L^1_{loc.}$

2. We can prove an estimate

$$||f * g||_p \leq ||f||_1 ||g||_p$$

3. Using the density of C_c^∞ in L_p and this estimate we extend the definition of $f * g$ to pairs for which the formula in 1) does not make a priori sense.

This pattern will be repeated over and over again.

Now we turn our attention to the space of linear functionals defined on $L^p(\mathbf{R}^n)$. We start with a lemma:

Lemma 3.2. *A linear functional $\ell : L_p(\mathbf{R}^n) \to \mathbf{C}$ is continuous iff there is a constant $c < \infty$ such that*

$$|\ell(f)| \leq C ||f||_p$$

Proof. If this estimate holds then it is clear that ℓ is continuous. On the other hand if such an estimate does not hold then we can find a sequence $\{f_n\}$ such that $||f_n||_p = 1$ but $|\ell(f_n)| = m_n$ tends to ∞. Then $f_n/m_n \to 0$ in L^p but $|\ell(f_n)| = 1$. This is a contradiction to the continuity of ℓ at 0. □

In infinite dimensions we generally only consider continuous linear functionals, unlike finite dimensions, it is possible to have non-continuous linear functionals. Let $(L^p)'$ denote the vector space of continuous linear functionals on L^p. We define a norm on this space by setting:

$$||\ell||' = \sup\{|\ell(f)|\,|f \in L^p : ||f||_p \leq 1\}.$$

Exercise 3.1. : Prove that this defines a norm.

The basic theorem is the following:

Theorem 3.3. *: If* $1 \leq p < \infty$ *then*

$$(L^p)' = L^q$$

where $\frac{1}{p} + \frac{1}{q} = 1$

From Hölder's inequality it's easy to see that

$$L^q \leq (L^p)'.$$

The reverse inclusion is more involved to prove.

We now consider the special case $p = 2$. $L^2(\mathbf{R}^n)$ has some additional structure which makes it much easier to analyze than $p \neq 2$, the L^2–norm is defined by an inner product:

$$\langle f, g\rangle = \int_{\mathbf{R}^n} f\bar{g}dx.$$

This is an Hermitian pairing:

(1) $\langle f, g\rangle = \overline{\langle g, f\rangle}$

(2) $\langle \lambda f, g\rangle = \lambda\langle f, g\rangle$, $\langle f, \lambda g\rangle = \bar{\lambda}\langle f, g\rangle$, $\lambda \in \mathbf{C}$.

Clearly $||f||_2^2 = \langle f, f\rangle$. A Banach space whose norm is defined by an inner product is called a Hilbert space. On an inner product space there is a notion of orthogonality: Let $S \subset L^2(\mathbf{R}^n)$ be a subspace, define

$$S^\perp = \{f \in L^2(\mathbf{R}^n) | \langle f, g\rangle = 0 \quad \text{for all} \quad g \in S\}.$$

Clearly $S^\perp$ is also a subspace. It has an important topological property:

Proposition 3.2. $S^\perp$ *is a closed subspace*

Proof: If $\{f_n\} \subset S^\perp$ and $f_n \to f$ in L^2 then for $g \in S$ we have:

$$\langle f, g\rangle = \lim_{n\to\infty} \langle f_n, g\rangle = 0.$$

The fundamental result about $L^2(\mathbf{R}^n)$ (or any Hilbert space) is the following:

Theorem 3.4 (The Projection Theorem)**.** *Let S be a closed subspace of the Hilbert space, H. If* $y \in H \setminus S$ *then there is a unique point* $s \in S$ *such that*

$$||y - s||_H = \inf_{x \in S} ||y - x||_H.$$

The proof of this statement relies on an algebraic identity:

For any Hilbert space:

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2) \qquad \text{(Parallelogram Law)}$$

Proof of the Projection Theorem: Choose a sequence $\{s_n\} \subset S$ such that

$$\lim_{n\to\infty} ||y - s_n|| = \inf_{x\in S} ||y - x|| = d.$$

We apply the parallelogram law to $y - s_n$, $y - s_m$ to obtain:

$$||s_n - s_m||^2 + 4||y - \frac{s_n + s_m}{2}||^2 = 2(||y - s_n||^2 + ||y - s_m||^2)$$

So that

$$\leq ||s_n - s_m||^2 = 2(||y - s_n||^2 + ||y - s_m||^2) - 4||y - \frac{s_m + s_n}{2}||^2$$

As m, $n \to \infty$ the lim inf of the R.H.S. is

$$4(d^2 - \lim_{n,m\to\infty} \sup ||y - \frac{s_n + s_m}{2}||^2) \leq 0$$

because $\frac{s_n+s_m}{2} \in S$. This proves that $\{s_n\}$ is a Cauchy sequence as H is a Hilbert space $s = \lim_{n\to\infty} s_n$ must exist. The uniqueness is established somewhat differently: Let $s_1, s_2 \in S$ be such that $||s_i - y|| = d \quad i = 1, 2$. We consider the function $q(\lambda) = ||\lambda s_1 + (1-\lambda)s_2 - y||^2$. This is a quadratic function of λ which assumes its minimum at $\lambda = 0$ and $\lambda = 1$. This is only possible if $q =$ constant i.e.

$$s_1 - s_2 = 0.$$

Observe that if $t \in S$ then the quadratic function

$$q(\lambda) = \langle \lambda t + s - y,\ \lambda t + s - y\rangle$$

assumes its minimum at $\lambda = 0$. If we let $\lambda = re^{i\theta}$ then this implies that

$$\Re(e^{i\theta}\langle t,\ s - y\rangle) = 0$$

By choosing θ so that $e^{i\theta}\langle t, s-y\rangle = |\langle t, s-y\rangle|$ we see that $\langle t, s-y\rangle = 0 \quad \forall t \in S$. In other words $s - y \in S^\perp$. As a corollary we have

Corollary 3.1. *: If $S \subset H$ is a closed subspace then every vector $y \in H$ has a unique representation as $y = y_0 + y_1$ where $y_0 \in S, y, \in S^\perp$.*

Proof. Let $s \in S$ be the point closest to y then

$$y = y - s + s.$$

This gives the desired decomposition. □

The first consequence of the projection theorem is

Theorem 3.5 (Riesz Representation Theorem)**.** *If $\ell \in H'$ then there is a unique $y \in H$ such that*

$$\ell(x) = \langle x, y\rangle$$

i.e. the inner product defines an isomorphism

$$H \simeq H'$$

Proof. Let $K = \{x \in H | \ell(x) = 0\}$. $K = \ell^{-1}(\{0\})$ is a closed subspace. Let $v_1, v_2 \in K^{\perp}\{0\}$ Evidently neither $\ell(v_1)$ nor $\ell(v_2)$ is zero but

$$\ell(\frac{v_1}{\ell(v_1)} - \frac{v_2}{\ell(v_2)}) = 0.$$

Thus $\frac{v_1}{\ell(v_1)} - \frac{v_2}{\ell(v_2)} \in K^{\perp} \cap K = \{0\}$. That is $K^{\perp}$ is one dimensional! Choose a vector $v \in K^{\perp}$ such that $\ell(v) = 1$. Every vector $x \in H$ has a unique representation as $x = x^{\perp} + av$ where $X^{\perp} \in K^{\perp}$

$$\ell(x) = al(v) = a.$$

On the other hand $\langle x, v\rangle = a\langle v, v\rangle$. Thus

$$\ell(x) = \langle x, \frac{v}{\langle v, v\rangle}\rangle.$$

□

As a corollary we have:

Corollary 3.2. *: If $x \in H$ then*

$$||x|| = \sup_{\{y|\ ||y||=1\}} \langle x, y\rangle.$$

There are several proofs that for $a \leq p < \infty$

$$(L^p)' = L^q.$$

Some of these proofs somehow reduce to the L^2–case where the theorem is relatively easy to prove.

4. The Fourier Integral

Now we consider the theory of the Fourier integral on $L^2(\mathbf{R}^n)$. We begin by defining everything for $f \in C_c^{\infty}(\mathbf{R}^n)$:

$$\hat{f}(\xi) = \int_{\mathbf{R}^n} f(x)e^{-ix\cdot\xi}dx, \quad \check{f}(x) = \int f(\xi)\ell^{ix\cdot\xi}\frac{d\xi}{(2\pi)^n}.$$

The main thing we wish to do is prove an estimate for $||\hat{f}(\xi)||_2$ which allows us to extend the Fourier transform to functions in $L^2(\mathbf{R}^n)$. There is a close relationship between the convolution product and the Fourier transform:

Proposition 4.1. *: For $f, g \in C_c^{\infty}(\mathbf{R}^n)$ we have that*

$$(f * g)^(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

Proof. $\widehat{(f * g)} = \int\int f(x-y)g(y)dye^{-ix\cdot\xi}dx$. This is an absolutely convergent integral so we can rearrange the integrations:

$$\begin{aligned}
\widehat{(f * g)} &= \int\int f(x-y)e^{-ix\cdot\xi}dxg(y)dy \\
&= \int \hat{f}(\xi)e^{-iy\cdot\xi}g(y)dy \qquad \text{Let } z = x - y \qquad (24) \\
&= \hat{f}(\xi)\hat{g}(\xi).
\end{aligned}$$

□

For $f \in C_c^\infty(\mathbf{R}^n)$ we can prove some elementary estimates on $\hat{f}(\xi)$. We have the identity:

$$\xi^\alpha \hat{f}(\xi) = \int_{\mathbf{R}^n} f(x)\, D_x^\alpha e^{-ix\xi} dx$$

where $D_x = (i\partial x_1, \ldots, i\partial x_n)$. α is a multiindex. Thus we can integrate by parts to obtain:

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

Thus

$$|\xi^\alpha \hat{f}(\xi)| \leq \int |Dx^\alpha f(x)| dx < \infty.$$

This easily implies that for any $N > 0$ there is a constant, C_N such that

$$|\hat{f}(\xi)| \leq \frac{C_N}{(1+|\xi|)^N}.$$

Note also that $\hat{f}(\xi)$ is differentiable, differentiating we obtain:

$$D_\xi^\alpha \hat{f}(\xi) = \int_{\mathbf{R}^n} x^\alpha f(x) e^{ix\cdot\xi} dx.$$

Evidently

$$\xi^\beta D_\xi^\alpha \hat{f} = \int_{\mathbf{R}^n} (-D_x)^\beta (x^\alpha f(x)) e^{-ix\cdot\xi} dx.$$

So we get the same sort of estimate for $D_\xi^\alpha \hat{f}(\xi)$. We now prove an identity:

$$\int f(x) g(-x) dx = \int \hat{f}(\xi) / \hat{g}(\xi) d\xi.$$

For $f, g \in C_c^\infty(\mathbf{R}^n)$ we use the proposition to obtain:

$$\int\int\int f(y) g(x-y) dy e^{-ix\xi} dx d\xi = \int \hat{f}(\xi)\hat{g}(\xi) d\xi.$$

The integral on the left cannot a priori be reordered, however it makes sense as $f * g \in C_c^\infty(\mathbf{R}^n)$ and $\widehat{(f*g)}(\xi)$ is easily seen to be rapidly decreasing and so $\int \widehat{(f*g)}(\xi) d\xi$ is an absolutely convergent integral. To prove the identity we will want to reorder the integrations so we put a convergence factor in which allows us to use Fubini's theorem:

$$\int\int\int f(y) g(x-y) dy e^{-ix\cdot\xi} dx d\xi = \lim_{\varepsilon \downarrow 0} \int\int\int f(y) g(x-y) e^{ix\cdot\xi} e^{-\varepsilon^2|\xi|^2} dy dx d\xi.$$

For each $\varepsilon > 0$ the right hand side is an absolutely convergent triple integral and so we can reorder the integrations as we please. we replace the integral on the R.H.S. with

$$\int\int\int e^{-ix\cdot\xi} e^{-\varepsilon^2|\xi|^2} d\xi f(y) g(x-y) dy dx$$

Using complex analysis one evaluates the inner integral:

$$\int e^{-\varepsilon^2|\xi|^2} e^{ix\cdot\xi} d\xi = \frac{\pi^{n/2} e^{-\frac{|x|}{4\varepsilon^2}}}{\varepsilon^n}.$$

Thus the integral becomes:

$$c\int\int f(y)g(x-y)\frac{e^{-\frac{|x|^2}{4\varepsilon^2}}}{\varepsilon^n}dxdy$$
$$= c\int f*g(x)\frac{e^{-\frac{|x|^2}{4\varepsilon^2}}}{\varepsilon^n}dx$$

As $\varepsilon \to 0$ this converges to $\tilde{c}f*g(0)$. This proves, for $f, g\epsilon C_c^\infty(\mathbf{R}^n)$:

Proposition 4.2. *: For $f,\, g \in \mathcal{C}_c^\infty(\mathbf{R^n})$ we have the identity:*

$$\int f(y)g(-y)dy = \int \hat{f}(\xi)\hat{g}(\xi)\frac{d\xi}{(2\pi)^n}.$$

Our argument only gave some constant on the R.H.S., working a little more carefully we could have seen directly that the constant is $(2\pi)^{-n}$.

Let $g(y) = \bar{f}(-y)$ then

$$\begin{aligned}\hat{g}(\xi) &= \int \bar{f}(-y)e^{-iy\cdot\xi}d\xi \\ &= \overline{\int f(-y)e^{iy\cdot\xi}d\xi} \\ &= \overline{\hat{f}(\xi)}.\end{aligned} \tag{25}$$

As a corollary we obtain:

Corollary 4.1. *: If $f \in C_c^\infty(\mathbf{R}^n)$ then*

$$\int |\hat{f}(\xi)|^2\frac{d\xi}{(2\pi)^n} = \int |f(y)|^2dy.$$

This is called the Plancherel formula. Using this we can extend the Fourier transform as a map from L^2 to L^2: For $f \in L^2$ we define

$$\hat{f}(\xi) = \text{l.i.m.}_{n\to\infty}\int f_n(x)e^{-ix\cdot\xi}dx$$

l.i.m. = limit in the mean. Here $\{f_n(x)\}$ is a sequence in $C_c^\infty(\mathbf{R}^n)$ converging to $f(x)$ in L^2. The argument above can also be used to prove the Fourier inversion formula:

$$f(x) = \int \hat{f}(\xi)e^{ix\cdot\xi}\frac{d\xi}{\pi^n}. \qquad \text{for} \quad f \in C_c^\infty(\mathbf{R}^n).$$

Exercise 4.1. : Prove that if $f \in L^2(\mathbf{R}^n)$ then

$$\hat{f}(\xi) = \text{l.i.m.}_{R\to\infty}\int_{|x|<R} f(x)e^{\ell x\cdot\xi}dx.$$

Exercise 4.2. Prove that there is a orthonormal basis for $L^2(\mathbf{R}^n)$, i.e. A sequence of functions $\{f_n\}$ such that

$$\begin{aligned}&i) \quad \langle f_n, f_m\rangle = \delta_{nm}, \\ &ii) \quad \bigcup_N\{\sum_{j=1}^N a_jf_j|a_j \in \mathbf{C}\} \qquad \text{is dense in} \quad L^2(\mathbf{R}^n).\end{aligned} \tag{26}$$

Remark 4.1. The Plancherel Theorem is quite a subtle statement for

$$\int_{\mathbf{R}^n} f(x)e^{-ix\cdot\xi}dx$$

is not defined as a convergent integral if $f \in L^1(\mathbf{R^n})$. Of course $L^2(\mathbf{R}^n) \not\subseteq L^1(\mathbf{R}^n)$!

5. Some Applications of the Fourier Transform

The object of study for this semester is variable coefficient partial differential equations:

$$Pu = \sum_{|\alpha|\leq M} a_\alpha(x)D^\alpha u(x) = f(x).$$

Here $a_\alpha(x)$ will usually be smooth functions of x.

If the coefficients are constant then the Fourier transform provides a powerful tool to study such an equation.

Recall that

$$\widehat{(D_x^\alpha f)}(\xi) = (-\xi)^\alpha \hat{f}(\xi).$$

We proceed formally and apply the Fourier transform to both sides of

$$\sum_{|\alpha|\leq M} a_\alpha D^\alpha u = f \qquad (a_\alpha \in \mathbf{C})$$

to obtain:

$$\sum a_\alpha(-\xi)^\alpha \hat{u}(\xi) = \hat{f}(\xi).$$

Continuing to proceed formally we divide to obtain

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{\sum_{|\alpha|\leq M} a_\alpha(-\xi)^\alpha}.$$

This is where the difficulties really begin! The polynomial $\sum a_\alpha(-\xi)^\alpha$ may have nontrivial real roots and we are left with the difficulty of interpreting $1/\sum a_\alpha(-\xi)^\alpha$. Notice for example that if we set

$$P(\xi) = \sum_{|\alpha|\leq M} (-\xi)^\alpha a_\alpha; \quad p_m(\xi) = \sum_{|\alpha|=m} a_\alpha(-\xi)^\alpha,$$

and ξ_0 is a direction such that $p_m(\xi_0) \neq 0$ then $p_m(\lambda\xi_0) = \lambda^m p_m(\xi_0)$ and so, for large enough λ $P(\lambda\xi_0) \sim \lambda^m$ as well. Thus we see that

$$\hat{u}(\lambda\xi_0) = \frac{\hat{f}(\lambda\xi_0)}{P(\lambda\xi_0)} \sim \lambda^{-m}\hat{f}(\lambda\xi_0).$$

If $p_m(\xi) \neq 0 \quad \forall \xi \neq 0$ then it is apparent that the Fourier transform of u decays at infinity m orders faster than the Fourier transform of f. We shall see that this means, in a precise sense, that u has m more derivatives than f.

Lets consider another sort of problem we can employ the Fourier transform to study. We begin with:

Theorem 5.1 (Cauchy's Formula:)**.** *If D is a domain in $\mathbf{C}$ with a smooth boundary and f is a holomorphic function on D smooth up to bD then*

$$f(z) = \frac{1}{2\pi i}\int_{bD} f(\zeta)\frac{d\zeta}{\zeta - z} \quad \textit{for} \quad z \in D.$$

Take for example $D = \{z|\Im z > 0\}$. We'll assume that $f(z) \to 0$ as $(z) \to \infty$ quickly enough that

$$f(z) = \frac{1}{2\pi i} \int \frac{f(x)}{x-z} dx.$$

Note the following: Even if $f(x)$ is not the boundary value of a holomorphic function then the r.h.s. above defines a holomorphic function, $F(z)$ in $\Im z > 0$. Let's set

$$F(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{f(x)dx}{x-z}.$$

What is $\lim_{y \downarrow 0} F(x+iy)$? Unfortunately

$$\int \frac{f(s)ds}{x-s}$$

is not, in general an absolutely convergent integral. Let's try using the Fourier transform:

$$F(x+iy) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{f(s)}{(x-s+iy)} ds.$$

Observe that this is a family of convolutions. We will try to take advantage of the relation:

$$\widehat{(f * g)}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

We need to compute the Fourier transform of $\frac{1}{x+iy}$ for $y > 0$ i.e.

$$\int_{-\infty}^{\infty} \frac{e^{-ix\cdot\xi}dx}{x+iy}.$$

Unfortunately this integral does not converge absolutely. However $\frac{1}{x+iy} \in L^2(\mathbf{R})$ for $y > 0$ and so

$$(\widehat{\frac{1}{x+iy}})(\xi) = \text{l.i.m.} \int_{-R}^{R} \frac{e^{-ix\cdot\xi}dx}{x+iy}.$$

On the other hand if $\xi \neq 0$ the improper integral also converges. We can use contour integration to compute this integral. If $\xi < 0$ then

$$\int_{-\infty}^{\infty} \frac{e^{ix\cdot\xi}dx}{x+iy} = \lim_{R\to\infty} \int_{\Gamma_R^+} \frac{e^{iz\cdot\xi}}{z+iy} dz.$$

Here Γ_R^+ is the contour:

Exercise 5.1. : Prove this formula, including the convergence of the left hand side.

Using the residue formula we obtain:

$$(\widehat{\frac{1}{x+iy}})(\xi) = \begin{cases} e^{-y\xi} & \text{for} \quad \xi > 0 \\ 0 & \quad \xi < 0. \end{cases}$$

So we see that

$$\hat{F}(\xi, y) = \hat{f}(\xi) e^{-y\xi} \chi_{[0,\infty)}(\xi).$$

We can now give an answer to our original question:

$$\lim_{y\downarrow 0} F(x+iy) = \frac{1}{\pi} \int_0^{\infty} \hat{f}(\xi) e^{ix\cdot\xi} d\xi.$$

Corollary 5.1. *: A function $f \in L^2(\mathbf{R})$ is the boundary value of a holomorphic function in $\Im z > 0$ found only if $\hat{f}(\xi) = 0 \quad \forall \xi < 0$.*

A similar analysis shows that $f(x)$ is the boundary value of a holomorphic function in $\Im z < 0$ if and only if $\hat{f}(\xi) = 0$ for $\xi > 0$. Combining these results we obtain:

Corollary 5.2. *: $L^2(\mathbf{R}) = H^2_+ \oplus H^2_-$ where H^2_+ (H_-) are L^2-boundary values of functions holomorphic in the upper (lower) half plane.*

6. Bounded Linear Operators

A linear map $A : L^2 \to L^2$ is continuous if and only if it satisfies the following estimate:

$$||Ax|| \le C||x||.$$

We define the norm of the linear operator A to be the best such constant, i.e.

$$||A|| = \sup_{x \ne 0} \frac{||Ax||}{||x||}.$$

From the Riesz Representation Theorem it follows that as $x \to \langle Ax, y\rangle$ is a continuous linear functional and so there is a z_y such that

$$\langle Ax, y\rangle = \langle x, z_y\rangle.$$

We define the adjoint of A by

$$A^* y = z_y.$$

An operator which $A = A^*$ is called self adjoint. For such an operator the quadratic form $\langle Ax, y\rangle$ is hermitian symmetric and so

$$\langle Ax, x\rangle$$

is real valued. In many cases $A : L^2 \to L^2$ is given by a function on $\mathbf{R}^n \times \mathbf{R}^n$:

$$Af(x) = \int k(x, y) f(y) dy.$$

k is called the kernel of A.

It is important to have criteria in terms of the kernel such that A is a bounded operator. A very simple case is the following: If $g \in L^1$ then the Hausdorff-Young inequality implies that

$$||g * f||_2 \le ||g||_1 ||f||_2.$$

Thus $Af = g * f$ is a bounded linear operator with $||A|| \le ||g||_1$.

There is a very important extension of this simple result.

Theorem 6.1 (Schurs Lemma)**.** *Suppose that $k(x, y)$ is a locally integrable function on $\mathbf{R}^n \times \mathbf{R}^n$ such that*

$$\sup_x \int |k(x, y)| dy \le M$$

$$\sup_y \int |k(x, y)| dx \le M.$$

Then

$$f \to \int k(x, y) f(y) dy \quad f \in c_c^\infty(\mathbf{R}^n)$$

extends to a bounded linear map $K : L^2 \to L^2$ with $||K|| \le M$.

Proof. We use the fact that $||f|| = \sup_{\{g|\ ||g||=1\}} |\langle f, g\rangle|$.

$$|\langle Kf, g\rangle| = |\int\int k(x,y)f(y)g(x)dydx|$$

$$\leq \int\int |k(x,y)|\,|f(y)||g(x)|dydx$$

by Hölder's inequaltiy we obtain that

$$|\langle Kf, g\rangle| \leq (\int\int |k(x,y)||f(y)|^2 dydx)^{1/2}(\int\int |k(x,y)||g(x)|^2 dydx)^{1/2}$$

$$\leq M^{1/2}||f|| \quad M^{1/2}||g||.$$

Thus $|\langle Kf, g\rangle| \leq M||f||$ for g with $||g|| = 1$. This proves the proposition. □

A consequence of the infinite dimensionality of $L^2(\mathbf{R}^n)$ is the fact that the unit ball is not compact. This motivates the introduction of a second topology for which the unit ball is compact. This is called the weak topology.

Suppose that $\{x_n\}$ is a sequence for which there exists an x such that for every $y \in H$

$$\lim_{n\to\infty} \langle x_n, y\rangle = \langle x, y\rangle.$$

We say that $\{x_n\}$ converges weakly to x

$$x_n \rightharpoonup x \qquad \text{w}-\lim_{n\to\infty} x_n = x.$$

There are certain relationships between the norm topology and the weak topology.

Proposition 6.1. : *If $\{x_n\}$ is a weakly convergent sequence then $\{||x_n||\}$ is a bounded set. Let $x = w - \lim x_n$. Then $\{x_n\}$ converges to x in the norm topology if and only if $\lim_{n\to\infty} ||x_n|| = ||x||$. In any case $||x|| \leq \lim_{n\to\infty} \inf ||x_n||$*

Proof. : The first statement is a consequence of the uniform boundedness principle. For each $y \in$ unit ball in H there is an M_y s.t.

$$\sup_n |\langle x_n, y\rangle| \leq M_y.$$

The u.b.p. $\Rightarrow$ there is a constant M such that

$$|\langle x_n, y\rangle| \leq M\|y\| \qquad \forall n,\, y.$$

Next observe that $0 \leq \langle x - x_n, x - x_n\rangle$ so that

$$0 \leq \langle x, x\rangle - [\langle x_n, x\rangle + \langle x, x_n\rangle] + \langle x_n, x_n\rangle$$

Taking $\liminf$ we conclude, as $\lim_{n\to\infty} \langle x_n, x\rangle = \langle x, x\rangle$ that

$$\lim_{n\to\infty} \inf\langle x_n, x_n\rangle - \langle x, x\rangle \geq 0$$

Evidently if $\lim_{n\to\infty} \langle x_n, x_n\rangle = ||x||^2$ then $\lim_{n\to\infty} ||x_n - x||^2 = 0$. □

Example 6.1. Let $f_n(x) = \begin{cases} 1 & n < x \leq n+1 \\ 0 & \text{otherwise} \end{cases}$

If $g \in L^2(\mathbf{R})$ then

$$|\langle g, f_n \rangle| = |\int_n^{n+1} g(x)dx| \leq (\int_n^{n+1} |g(x)|^2 dx)^{1/2}$$

clearly this tends to zero as $n \to \infty$. So $\text{w}-\lim_{n\to\infty} f_n = 0$. On the other hand $||f_n|| = 1 \quad \forall n$ so clearly $f_n \nrightarrow 0$ in norm.

Exercise 6.1. : Show that the unit ball in $L^2(\mathbf{R}^n)$ is weakly compact.

It turns out to be very interesting to consider operators $K : L^2(\mathbf{R}^n) \to L^2(\mathbf{R}^n)$ which are continuous from the weak topology to the strong topology.

Definition 6.1. : If K is a linear operator such that for every weakly convergent sequence $\{x_n\}$ the sequence $\{Kx_n\}$ is strongly convergent then we say that K is a compact operator.

Compact operators are closely related to finite rank operators. An operator is of finite rank if $\operatorname{Im} A$ is finite dimensional or equivalently $(\operatorname{Ker} A)^\perp$ is finite dimensional. A finite rank operator has a very simple useful representation. Let $\{x_1, \ldots x_n\}$ be an orthonormal basis for $(\operatorname{Ker} A)^\perp$ and set $y_1 = Ax_i$. Then

$$Az = \sum_{i=1}^{N} \langle z, x_i \rangle y_i$$

Note for example that if $\{x_i\}_{i=1}^\infty$ is an orthonormal basis for $L^2(\mathbf{R}^n)$ then we can define finite rank projection operators by setting:

$$\pi_N x = \sum_{i=1}^{N} \langle x, x_i \rangle x_i.$$

Clearly $\lim_{N\to\infty} \pi_N x = x \qquad \forall x \in L^2(\mathbf{R}^n)$. Also $\pi_N^* = \pi_N$. Such a projection is called an orthogonal projection. Note that

$$\|x\|^2 = \|\pi_N x\|^2 + \|(\mathrm{I} - \pi_N)x\|^2.$$

Proposition 6.2. *: If $\{x_i\}$ is an orthonormal basis and K is compact operator then the finite rank operators $K\pi_N$ converge to K in the operator norm topology. That is $||K - K\pi_N|| \to 0$ as $N \to \infty$.*

Proof. We need to show that given $\varepsilon > 0$ there is an N_0 so that $||K - K\pi_N|| < \varepsilon$ if $N > N_0$. Suppose not. Then we can find a sequence $\{y_N\}$ such that $||y_N|| = 1$

$$||(K - K\pi_N)y_N|| \geq \varepsilon$$

As $Ky_N = K(I - \pi_N)y_N + K\pi_N y_N$. We can assume that $\pi_N y_N = 0$.

Claim: A sequence $\{y_N\}$ s.t. $||y_N|| = 1, \quad \pi_N y_N = 0$ converges weakly to zero. Let $x \in L^2(\mathbf{R}^n)$ given $\eta > 0$ there exists an M s.t. $||x - \pi_M x|| < \eta$. Thus

$$\begin{aligned} |\langle x, y_n \rangle| &= Kx - \pi_N x, y_N \rangle| \\ &\leq ||x - \pi_N x|| \\ &< \eta \qquad \text{if}\, N > M. \end{aligned} \tag{27}$$

As η is arbitrary this shows $y_N \to 0$. As K is compact this implies that $Ky_N \to 0$ but $||Ky_N|| \geq \epsilon$ by assumption. □

Exercise 6.2. Show that $\pi_N K \pi_N$ converges to K in the operator norm.

Exercise 6.3. If K is a compact operator then its adjoint, K^* is as well.

Exercise 6.4. Show that if $\{K_N\}$ is a sequence of finite rank operators converging in norm to an operator K then K is a compact operator.

Example 6.2. On ℓ^2 we define the operator

$$K(a_1, a_2, \dots) = (\frac{a_1}{1}, \frac{a_2}{2}, \frac{a_3}{3}, \cdots)$$

This is easily seen to be compact.

Example 6.3. Let f be a 2π–periodic C^0-function then we define an operator on $L^2(S^1)$ by:

$$K_f g = \int_0^{2\pi} f(x-y)g(y)dy$$

We claim that this is a compact operator.

Proof.

$$|K_f g(x_1) - K_f g(x_2)| = |\int_0^{2\pi} (f(x_1 - y) - f(x_2 - y)g(y)dy|$$
$$\leq (\int_0^{2\pi} |f(x_1 - y) - f(x_2 - y)|^2 dy)^{1/2} ||y||_2.$$

Let

$$\omega(\delta) = \sup_{|x_1 - x_2| \leq \delta} (\int_0^{2\pi} |f(x_1 - y) - f(x_2 - y)|^2 dy)^{1/2}.$$

The function, $\omega(\delta)$ is clearly a monotone increasing function for which $\lim_{\delta \downarrow 0} \omega(\delta) = 0$. $K_f g(x)$ is therefore a continuous function with modulus of continuity $||g||_2 \omega(\delta)$. Using the Arzela—Ascoli theorem we can easily prove that K_f is a compact operator. □

Let $\psi \in C_c^\infty(\mathbf{R}^n)$, neither operator

$$M_\psi f = \psi f, \quad C_\psi f = \psi * f$$

is compact on $L^2(\mathbf{R}^n)$. However

$$K_\psi f = M_\psi C_\psi f \quad \text{or} \quad \tilde{K}_\psi = C_\psi M_\psi f$$

are compact operators.

If K is a compact, self adjoint operator then it has a spectral theory identical to that of a self adjoint operator on a finite dimensional space.

Theorem 6.2. *: If K is a self adjoint, compact operator then there is an orthonormal sequence $\{x_i\}_{i=1}$ and a sequence $\{\lambda_i\}$ such that*

$$Ax_i = \lambda_i x_i.$$

$$\lim_{i \to \infty} |\lambda_i| = 0. \qquad (\operatorname{Ker} K)^\perp = \overline{Span\{x_i : i = 1 \dots \infty\}}$$

To prove this theorem we can use the Courant-Fisher minmax method.

Claim: Suppose there is a vector x such that $\langle Kx, x\rangle > 0$ then there is a unit vector $x_0^{+\cdot}$ such that $\mu_0 = \langle Kx_0^+, x_0^+\rangle = \sup_{\{x|\,||x||=1\}}\langle Kx, x\rangle$. Moreover x_0^+ is an eigenvector with $Kx_0^+ = \mu_0 x_0^+$. Evidently an analogous statement is true if $\langle Kx, x\rangle < 0$ for some x. We proceed inductively by observing that

$$K : \langle x_0\rangle^\perp \to \langle x_0\rangle^\perp.$$

By using compactness we can show the last statements.

If K is a compact operator then $(I+K)$ may or may not be an invertible operator. If $||K|| < 1$ then the Neumann series gives a formula for $(I + K)^{-1}$:

$$(I + K)^{-1} = \sum_{i=0}^{\infty}(-1)^2 K^2.$$

The series converge in the norm topology. It is a simple computation to show that the limit of the sum is in fact $(I + K)^{-1}$.

In general we can write $K = K_0 + K_1$ where K_0 has finite rank and $||K_1|| < 1$. This implies that $(I + K_1)^{-1}$ exists, note that

$$(I + K)(I + K_1)^{-1} = I + K_0(I + K_1)^{-1}$$

$$(I + K_1)^{-1}(I + K) = I + (I + K_1)^{-2}K_0.$$

The operator $K_0(I + K_1)^{-1}$ and $(I + K_1)^{-1}K_0$ are finite rank operators. Notice that

$$\operatorname{Im}(I + K) = \operatorname{Im}(I + K_0(I + K_1)^{-1})$$

and the

$$\operatorname{Ker}(I + K) = \operatorname{Ker}(I + (I + K_1)^{-1}K_0).$$

In a short while we will see that this implies the following

$$\dim \operatorname{Ker}(I + K) + \dim[\operatorname{Im}(I + K)]^\perp.$$

Proposition 6.3. *: If K is a compact operator then $\operatorname{Im}(I+K)$ is a closed subspace.*

Proof. Let $y_n = (I + K)x_n$ be a convergent sequence with $\lim_{n\to\infty} y_n = y$. Evidently we can choose $\{x_n\} \subset [\operatorname{Ker}(I + K)]^\perp$.

Claim: there exists a constant $c > 0$ such that

$$\inf_{x\in(\operatorname{Ker}(I+K))^\perp} \frac{||(I + K)x||}{||x||} \geq c.$$

If not then there is a sequence $\{x_n\}$ with

$$||x_n|| = 1,\ x_n \perp \operatorname{Ker}(I + K) \text{ and } \lim_{n\to\infty}(I + K)x_n|| = 0.$$

Observe that $\{x_n\}$ has a weakly convergent subsequence $\{x_{n_j}\}$. Let $y = \text{w}-\lim_{\lambda\to\infty} x_{n_j}$. Note that as K is compact, $\{Kx_{n_j}\}$ converges strongly. As K is continuous $\lim_{j\to\infty} Kx_{n_j} = Ky$. By assumption $\lim_{j\to\infty} ||x_{n_j} + Ky|| = 0$, this implies that $\{x_{n_j}\}$ is actually a strong Cauchy sequence as well and so $\lim_{j\to\infty} x_{n_j} = y$. This implies that

$y \neq 0$, in fact $||y|| = \lim_{j\to\infty} ||x_{n_j}||$ and $(I+K)y = 0$. But $y \perp [\mathrm{Ker}(I+K)]$. This proves the claim.

From the claim we conclude that

$$||x_n - x_m|| \leq C||y_n - y_m||.$$

Thus $\{x_n\}$ is also a Cauchy sequence which therefore converges to some point, x. Clearly $(I+K)x = y$. This proves the proposition. □

Proposition 6.4. *: If K is a compact operator then*

$$\mathrm{Im}(I+K) = [\mathrm{Ker}(I+K^*)]^{\perp}.$$

Proof. If A is any bounded operator then $(\mathrm{Im}\, A)^{\perp} = \mathrm{Ker}\, A^*$. To see this we use the definition: $x \in (\mathrm{Im}\, A)^{\perp}$ if and only if $\langle x, Ay\rangle = 0 \quad \forall y \in H$ Or $\langle A^*x, y\rangle = 0$. But $\langle A^*x, y\rangle = 0 \quad \forall y \in H \Rightarrow A^*x = 0$. So $\mathrm{Ker}\, A^* \supset (\mathrm{Im}\, A)^{\perp}$, the other containment follows similarly. So

$$(\mathrm{Im}\, A)^{\perp\perp} = [\mathrm{Ker}\, A^*]^{\perp}.$$

But $(\mathrm{Im}\, A)^{\perp\perp} =$ closure of $\mathrm{Im}\, A$. As $\mathrm{Im}(I+K)$ is closed this implies that

$$\mathrm{Im}(I+K) = [\mathrm{Ker}(I+K^*)]^{\perp}.$$

□

Observe that the main point here was that $(I+K)$ has a closed range. This motivates the following definition.

Definition 6.2. A bounded operator $A : H \to H$ is called a Fredholm operator if

1) A has a closed range

2) ker A is finite dimensional

3) H/AH is finite dimensional

(Note 3⇒ 1).

Our proof shows:

Proposition 6.5. *: If A is a Fredholm operator then*

$$\mathrm{Im}\, A = (\mathrm{Ker}\, A^*)^{\perp}.$$

For a Fredholm operator we can define an integer invariant: called the index:

$$\mathrm{ind}(A) = \dim \mathrm{Ker}\, A - \dim \mathrm{Ker}\, A^*.$$

This index has many interesting properties:

1) If A is Fredholm then there is an $\varepsilon > 0$ such that for any operator B with $||B|| < \varepsilon$ we have that $A+B$ is Fredholm and

$$\mathrm{ind}(A+B) = \mathrm{ind}(A).$$

2) If A is Fredholm and K is compact then $A+K$ is Fredholm and

$$\mathrm{ind}(A+K) = \mathrm{ind}\, A.$$

Notice that we don't require a norm estimate for K.

3) If A and B are both Fredholm then AB is Fredholm and

$$\operatorname{ind}(AB) = \operatorname{ind} A + \operatorname{ind} B$$

4) $\operatorname{ind} A = -\operatorname{ind} A^*$.

We define one more class of operators: An operator K is called trace class if the non-negative self adjoint square root of K^*K has eigenvalues $\{\lambda_n\}$ such

$$||K||_{tr} = \sum_{j=1} \lambda_j < \infty.$$

$||\cdot||_{tr}$ defines a norm on the trace class operators. If $\{e_i\}$ is an orthonormal basis then we define:

$$\operatorname{tr} K = \sum_{i=1}^{\infty} \langle Ke_i, e_i \rangle.$$

Note that this independent of the choice of basis. If $Ke_i = \sum_{j=1}^{\infty} a_{ij} e_j$ then $\operatorname{tr} K = \sum a_{ii}$. If $\{f_j\}$ is another orthonormal basis then

$$e_i = \sum b_{ij} f_j$$

Where $\delta_{ij} = \langle e_i, e_j \rangle = \langle \sum b_{ik} f_k, \sum b_{j\ell} f_\ell \rangle = \sum b_{ik} \bar{b}_{jk}$ is an absolutely convergent sum.

$$\begin{aligned}
\sum \langle Ke_i, e_i \rangle &= \sum \langle K \sum b_{ik} f_k, \sum b_{i\ell} f_\ell \rangle \\
&= \sum_i \sum_\ell \langle \sum b_{ik} K f_k, b_{i\ell} f_\ell \rangle \\
&= \sum_i \sum_\ell \sum_k \langle b_{ik} K f_k, b_{i\ell} f_\ell \rangle \\
&= \sum_\ell \sum_k \sum_i b_{ik} \bar{b}_{i\ell} \langle K f_k, f_\ell \rangle \\
&= \sum \langle K f_k, f_k \rangle.
\end{aligned}$$

If $\{\lambda_j\}$ are the eigenvalues of K with multiplicity then

$$\operatorname{tr} K = \sum_{j=1}^{\infty} \lambda j.$$

If A is a bounded operator and K is trace class then AK and KA are trace class moreover:

$$\operatorname{tr} AK = \operatorname{tr} KA.$$

Lemma 6.1. *: If A is Fredholm then there is a bounded operator, B such*

$$BA = I - P$$

$$AB = I - Q$$

Where P is the orthogonal projection onto Ker A *and Q is the orthogonal projection on* Ker A^*.

Proof. We can consider $A : [\operatorname{Ker} A]^{\perp} \to \operatorname{Im} A$. As $\operatorname{Im} A$ is closed, it has the structure of a Hilbert space. This restriction of A is one to one and onto and therefore has a bounded inverse, call it B_0. We extend B to all of H by setting it equal to zero $[\operatorname{Im} A]^{\perp}$. If $x \in H$ then $x = x_0 + x_1$ where $x_0 \quad \in [\operatorname{Ker} A)$ and $x_1 \in [\operatorname{Ker} A]^{\perp}$. $BAx = BAx_1 = x_1 = (I-P)x$. On the other hand. $y = y_0 + y_1$ where $y_0 \in [\operatorname{Im} A]^{\perp}$ and $y_1 \in \operatorname{Im} A$ then $ABy = ABy_1 = y_1 = (I - Q)y$. □

Observe that if P is a finite rank orthogonal projection operator then

$$\dim \operatorname{Im} P = \operatorname{tr} P$$

From this we deduce the formula:

$$\operatorname{ind} A = \operatorname{tr}(AB - BA) = \operatorname{tr}(P - Q) = \operatorname{tr} P - \operatorname{tr} Q.$$

Far more interesting is the following

Proposition 6.6. *: If A is a Fredholm operator and B is a bounded operator such that $AB - I$ and $BA - I$ are trace class then*

$$\operatorname{ind} A = \operatorname{tr}(AB - BA)$$

Proof. Let B' be the operator from the lemma. So we have that $B'A = I - P' \qquad AB' = I - Q', \quad \operatorname{ind} A = \operatorname{tr}(AB' = B'A)$.

Claim: $B - B'$ is trace class: Let $BA = I - P$, $AB = I - Q \quad P, Q$ are trace class,

$$BAB' = B(I - Q')$$
$$(I - P)B' \quad \text{thus}$$
$$B - B' = BQ' - PB'.$$

As $A\,B$ and B' are bounded and Q' and P are trace class it follows that $B - B'$ is trace class. Observe that

$$\begin{aligned} \operatorname{tr}(AB' - B'A) - \operatorname{tr}(AB - BA) &= \operatorname{tr}(A(B' - B) - (B' - B)A) \\ &= \operatorname{tr}[A, B' - B]. \end{aligned} \tag{28}$$

As $B' - B$ is trace class and A is bounded it follows that

$$\operatorname{tr}[A, B' - B] = 0.$$

□

There are several things we should take away from this discussion:

1. There is a class of operators on an infinite dimensional space which behave very much like operators on finite dimensional spaces, in so far as solving $Ax = y$ is concerned,

2. For this class of there is an interesting integer invariant.

3. To compute this invariant it is only necessary to approximately invert the operator.

The theme of approximately inverting operators will come up again with notions of a "small error" informed by the previous discussion. In a Hilbert space compact operators are pretty small, trace class operators are even smaller. As a corollary of the proposition we have:

Corollary 6.1. *: If K is compact then* $\operatorname{ind}(I + K) = 0$.

Proof. We write $K = K_0 + K_1$ with K_0 finite rank and $||K_1|| < 1$ Then

$$(I+K)(I+K_1)^{-1}I = K_0(I+K_1)^{-1}$$
$$(I+K_1)^{-1}(I+K) - I = (I+K_1)^{-1}K_0,$$

As the errors have finite rank it follows that they are trace class. Hence

$$\operatorname{ind}(I+K) = \operatorname{tr}[(I+K_1)^{-1}, K_0] = 0.$$

□

This in turn implies that:

Corollary 6.2. *: If K is a compact operator the*

$$\dim \operatorname{Ker}(I+K) = \dim \operatorname{Ker}((I+K^*).$$

This is exactly as in the finite, equidimensional case.

An example of a Fredholm operator with non zero index is the following. Let

$$H^2_+ = \{\sum_{i=0}^{\infty} a_j e^{ij\theta} | \sum_{j=0}^{\infty} |a_j|^2 < \infty\},$$

the boundary values of L^2–holomorphic functions on the disk. We define a map by $Af = e^{i\theta} f$, it is easy to show that $\operatorname{ind} A = -1$.

Exercise 6.5. What is A^* in the previous example?

Exercise 6.6. Let $\pi : L^2(S^1) \to H^2_+$ denote the orthogonal projection. Let $a \in C^0(S^1)$; we define an operator $T_a : H^2_+ \to H^2_+$ by

$$T_a f = \pi(af).$$

1. Show that the norm of the operator T_a satisfies

$$\|T_a\| \leq \|a\|_\infty,$$

thus if $a_n \to a$ in the C^0–topology then the operators T_{a_n} converge to T_a in the norm topology.

2. If a is a trigonometric polynomial,

$$a(e^{i\theta}) = \sum_{j=-N}^{N} \alpha_j e^{ij\theta}$$

then prove that the operator $[\pi, a] : H^2_+ \to L^2$ is compact. Use part 1. to conclude this for any $a \in C^0(S^1)$.

3. Use part 2. to show that if a is non–vanishing then T_a is a Fredholm operator. Hint: What kind of operator is $T_a T_{1/a} - I$?

4. Prove that for $a \in C^1(S^1)$, a non–vanishing function

$$\operatorname{ind}(T_a) = \frac{i}{2\pi} \int_0^{2\pi} \frac{da}{a}.$$

Hint: prove this for $a = e^{ik\theta}$. Show that for any non–vanishing a there is a smooth family of nonvanishing, differentiable functions, a_t such that $a_0 = a$ and $a_1 = e^{ik\theta}$ for an appropriate value of k.

References for the material on functional analysis are

1. "Real and Complex Analysis" by Walter Rudin,

2. "Functional Analysis" by F. Riesz and B Sz–Nagy,

3. "Perturbation Theory of Linear Operators" by T. Kato,

Department of Mathematics, University of Pennsylvania

E-mail address: cle@math.upenn.edu