Functional Analysis (Math 920)
Lecture Notes for Spring '08

Jeff Schenker

Michigan State University
E-mail address: jeffrey@math.msu.edu

Contents

Comments and course information

These are lecture notes for Functional Analysis (Math 920), Spring 2008. The text for this course is *Functional Analysis* by Peter D. Lax, John Wiley & Sons (2002), referred to as "Lax" below. In some places I follow the book closely in others additional material and alternative proofs are given.

Other excellent texts include

- M. Reed and B. Simon, *Methods of Modern Mathematical Physics Vol. I: Functional Analysis*, Academic Press (1980).
- W. Rudin, *Functional Analysis*, McGraw-Hill, 2nd ed. (1991).

(As needed, these will be referred to below as "Reed and Simon" and "Rudin" respectively.)

Part 1

Hahn-Banach Theorem and Applications

LECTURE 1

Linear spaces and the Hahn Banach Theorem

Reading: Chapter 1 and §3.1 of Lax

Many objects in mathematics — particularly in analysis — are, or may be described in terms of, linear spaces (also called vector spaces). For example:

(1) $C(M)$ = space of continuous functions ($\mathbb{R}$ or $\mathbb{C}$ valued) on a manifold M.
(2) $A(U)$ = space of analytic functions in a domain $U \subset \mathbb{C}$.
(3) $L^p(\mu) = \{p$ integrable functions on a measure space $\mathcal{M}, \mu\}$.

The key features here are the axioms of linear algebra,

Definition 1.1. A linear space X over a field F (in this course $F = \mathbb{R}$ or $\mathbb{C}$) is a set on which we have defined

(1) addition: $x, y \in X \mapsto x + y \in X$
and
(2) scalar multiplication: $k \in F$, $x \in X \mapsto kx \in X$

with the following properties

(1) $(X, +)$ is an abelian group (+ is commutative and associative and $\exists$ identity and inverses.)
 - identity is called 0 (“zero”)
 - inverse of x is denoted $-x$
(2) scalar multiplication is
 - associative: $a(bx) = (ab)x$,
 - distributive: $a(x + y) = ax + by$ and $(a + b)x = ax + bx$,

 and satisfies $1x = x$.

Remark. It follows from the axioms that $0x = 0$ and $-x = (-1)x$.

Recall from linear algebra that a set of vectors $S \subset X$ is linearly independent if

$$\sum_{j=1}^{n} a_j x_j = 0 \text{ with } x_1, \ldots, x_n \in S \implies a_1 = \cdots = a_n = 0$$

and that the dimension of X is the cardinality of a maximal linearly independent set in X. The dimension is also the cardinality of a minimal spanning set, where the span of a set S is the set

$$\operatorname{span} S = \left\{ \sum_{j=1}^{n} a_j x_j \ : \ a_1, \ldots, a_n \in \mathbb{R} \text{ and } x_1, \ldots, x_n \in S \right\},$$

and S is spanning, or spans X, if $\operatorname{span} S = X$.

More or less, functional analysis is linear algebra done on spaces with infinite dimension. Stated this way it may seem odd that functional analysis is part of analysis. For finite dimensional spaces the axioms of linear algebra are very rigid: there is essentially only one interesting topology on a finite dimensional space and up to isomorphism there is only one linear space of each finite dimension. In infinite dimensions we shall see that topology matters a great deal, and the topologies of interest are related to the sort of analysis that one is trying to do.

That explains the second word in the name ”functional analysis.” Regarding “functional,” this is an archaic term for a function defined on a domain of functions. Since most of the spaces we study are function spaces, like $C(M)$, the functions defined on them are “functionals.” Thus “functional analysis.” In particular, we define a linear functional to be a linear map $\ell : X \to F$, which means

$$\ell(x+y) = \ell(x) + \ell(y) \text{ and } \ell(ax) = a\ell(x) \text{ for all } x, y \in X \text{ and } a \in F.$$

Often, one is able to define a linear functional at first only for a limited set of vectors $Y \subset X$. For example, one may define the Riemann integral on $Y = C[0,1]$, say, which is a subset of the space $B[0,1]$ of all bounded functions on $[0,1]$. In most cases, as in the example, the set Y is a subspace:

DEFINITION 1.2. A subset $Y \subset X$ of a linear space is a linear subspace if it is closed under addition and scalar multiplication: $y_1, y_2 \in Y$ and $a \in F \implies y_1 + ay_2 \in Y$.

For functionals defined, at first, on a subspace of a linear space of $\mathbb{R}$ we have

THEOREM 1.1 (Hahn (1927), Banach (1929)). *Let X be a linear space over $\mathbb{R}$ and p a real valued function on X with the properties*

(1) $p(ax) = ap(x)$ for all $x \in X$ and $a > 0$ (Positive homogeneity)
(2) $p(x+y) \le p(x) + p(y)$ for all $x, y \in X$ (subadditivity).

If ℓ is a linear functional defined on a linear subspace of Y and dominated by p, that is $\ell(y) \le p(y)$ for all $y \in Y$, then ℓ can be extended to all of X as a linear functional dominated by p, so $\ell(x) \le p(x)$ for all $x \in X$.

EXAMPLE. Let $X = B[0,1]$ and $Y = C[0,1]$. On Y, let $\ell(f) = \int_0^1 f(t)\mathrm{d}t$ (Riemann integral). Let $p : B \to \mathbb{R}$ be $p(f) = \sup\{|f(x)| \ : \ x \in [0,1]\}$. Then p satisfies (1) and (2) and $\ell(f) \le p(f)$. Thus we can extend ℓ to all of $B[0,1]$. We will return to this example and see that we can extend ℓ so that $\ell(f) \ge 0$ whenever $f \ge 0$. This defines a finitely additive set function on all(!) subset of $[0,1]$ via $\mu(S) = \ell(\chi_S)$. For Borel measurable sets it turns out the result is Lebesgue measure. That does not follow from Hahn-Banach however.

The proof of Hahn-Banach is not constructive, but relies on the following result equivalent to the axiom of choice:

THEOREM 1.2 (Zorn's Lemma). *Let S be a partially ordered set such that every totally ordered subset has an upper bound. Then S has a maximal element.*

To understand the statement, we need

DEFINITION 1.3. A partially ordered set S is a set on which an order relation $a \le b$ is defined for some (but not necessarily all) pairs $a, b \in S$ with the following properties

(1) transitivity: if $a \le b$ and $b \le c$ then $a \le c$
(2) reflexivity: if $a \le a$ for all $a \in S$.

(Note that (1) asserts two things: that a and c are comparable and that $a \le c$.) A subset T of S is totally ordered if $x, y \in T \implies x \le y$ or $y \le x$. An element $u \in S$ is an upper bound for $T \subset S$ if $x \in T \implies x \le u$. A maximal element $m \in S$ satisfies $m \le b \implies m = b$.

Proof of Hahn-Banach. To apply Zorn's Lemma, we need a poset,

$$S = \{\text{extensions of } \ell \text{ dominated by } p\}\,.$$

That is S consists of pairs (ℓ', Y') with ℓ' a linear functional defined on a subspace $Y' \supset Y$ so that

$$\ell'(y) = \ell(y),\ y \in Y \quad \text{and } \ell'(y) \le p(y),\ y \in Y'.$$

Order S as follows

$$(\ell_1, Y_1) \ \le\ (\ell_2, Y_2) \iff Y_1 \subset Y_2 \text{ and } \ell_2|_{Y_1} = \ell_1.$$

If T is a totally ordered subset of S, let $(\overline{\ell}, \overline{Y})$ be

$$\overline{Y} = \bigcup \{Y' : (\ell', Y') \in T\}$$

and

$$\overline{\ell}(y) = \ell'(y) \text{ for } y \in Y'.$$

Since T is totally ordered the definition of $\overline{\ell}$ is unambiguous. Clearly $(\overline{\ell}, \overline{Y})$ is an upper bound for T. Thus by Zorn's Lemma there exists a maximal element (ℓ^+, Y^+).

To finish, we need to see that $Y^+ = X$. It suffices to show that (ℓ', Y') has an extension whenever $Y' \ne X$. Indeed, let $x_0 \in X$. We want ℓ'' on $Y'' = \{ax_0 + y \ :\ y \in Y,\ a \in \mathbb{R}\}$. By linearity we need only define $\ell''(x_0)$. The constraint is that we need

$$a\ell''(x_0) + \ell'(y) \le p(ax_0 + y)$$

for all a, y. Dividing through by $|a|$, since Y' is a subspace, we need only show

$$\pm\, \ell''(x_0) \le p(y \pm x_0) - \ell'(y) \tag{1.1}$$

for all $y \in Y'$. We can find $\ell''(x_0)$ as long as

$$\ell'(y') - p(y' - x_0) \le p(x_0 + y) - \ell'(y) \text{ for all } y, y' \in Y', \tag{1.2}$$

or equivalently

$$\ell'(y' + y) \le p(x_0 + y) + p(y' - x_0) \text{ for all } y, y' \in Y'. \tag{1.3}$$

Since

$$\ell'(y' + y) \le p(y' + y) = p(y' - x_0 + y + x_0) \le p(x_0 + y) + p(y' - x_0),$$

(1.3), and thus (1.2), holds. So we can satisfy (1.1). □

In finite dimensions, one can give a constructive proof involving only finitely many choices. In infinite dimensions the situation is a quite a bit different, and the approach via Zorn's lemma typically involves uncountably many "choices."

LECTURE 2

Geometric Hahn-Banach Theorems

Reading: §3.2, §3.3 of Lax.

We may use Hahn-Banach to understand something of the geometry of linear spaces. We want to understand if the following picture holds in infinite dimension:

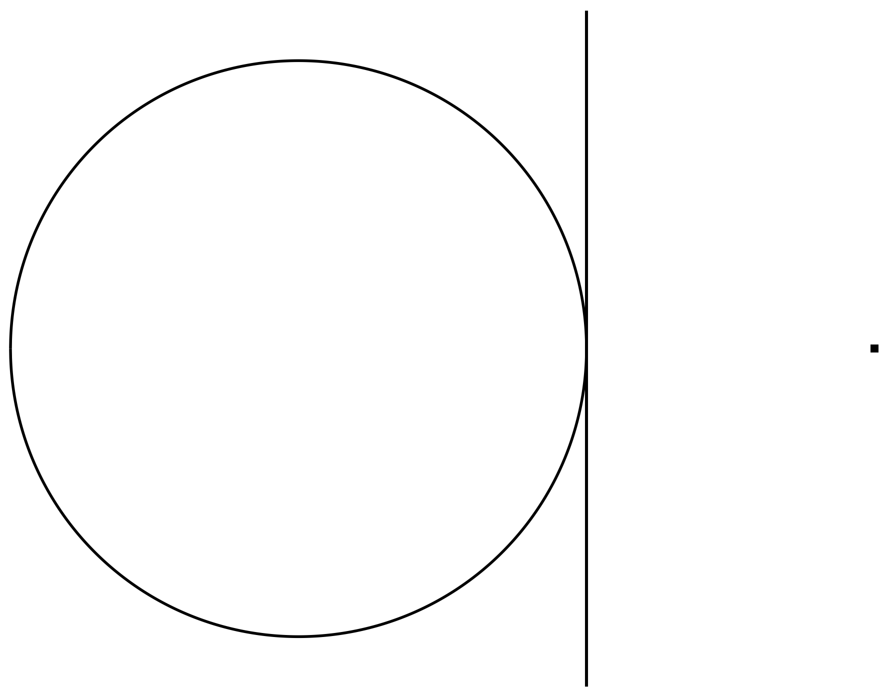

FIGURE 2.1. Separating a point from a convex set by a ~~line~~ hyperplane

DEFINITION 2.1. A set $S \subset X$ is convex if for all $x, y \in S$ and $t \in [0,1]$ we have $tx + (1-t)y \in S$.

DEFINITION 2.2. A point $x \in S \subset X$ is an interior point of S if for all $y \in X$ $\exists \varepsilon > 0$ s.t. $|t| < \varepsilon \implies x + ty \in S$.

REMARK. We can a define a topology using this notion, letting $U \subset X$ be open $\iff$ all $x \in U$ are interior. From the standpoint of abstract linear algebra this seems to be a "natural" topology on X. In practice, however, it has way too many open sets and we work with weaker topologies that are relevant to the analysis under considerations. Much of functional analysis centers around the interplay of different topologies.

We are aiming at the following

THEOREM 2.1. *Let K be a non-empty convex subset of X, a linear space over $\mathbb{R}$, and suppose K has at least one interior point. If $y \notin K$ then $\exists$ a linear functional $\ell : X \to \mathbb{R}$ s.t.*

$$\ell(x) \leq \ell(y) \text{ for all } x \in K, \tag{2.1}$$

with strict inequality for all interior points x of K.

This is the "hyperplane separation theorem," essentially validates the picture drawn above. A set of the form $\{\ell(x) = c\}$ with ℓ a linear functional is a "hyperplane" and the sets $\{\ell(x) < c\}$ are "half spaces."

To accomplish the proof we will use Hahn-Banach. We need a dominating function p.

DEFINITION 2.3. Let $K \subset X$ be convex and suppose 0 is an interior point. The gauge of K (with respect to the origin) is the function $p_K : X \to \mathbb{R}$ defined as

$$p_K(x) \;=\; \inf\{a \;:\; a > 0 \text{ and } \frac{x}{a} \in K\}.$$

(Note that $p_K(x) < \infty$ for all x since 0 is interior.)

LEMMA 2.2. *p_K is positive homogeneous and sub-additive.*

PROOF. Positive homogeneity is clear (even if K is not convex). To prove sub-additivity we use convexity. Consider $p_K(x + y)$. Let a, b be such that $x/a, y/b \in K$. Then

$$t\frac{x}{a} + (1-t)\frac{y}{b} \in K \quad \forall t \in [0,1],$$

so

$$\frac{x+y}{a+b} = \frac{a}{a+b}\frac{x}{a} + \frac{b}{a+b}\frac{y}{b} \in K.$$

Thus $p_K(x + y) \leq a + b$, and optimizing over a, b we obtain the result. $\square$

PROOF OF HYPERPLANE SEPARATION THM. Suffices to assume $0 \in K$ is interior and $c = 1$. Let p_K be the gauge of K. Note that $p_K(x) \leq 1$ for $x \in K$ and that $p_K(x) < 1$ if x is interior, as then $(1 + t)x \in K$ for small $t > 0$. Conversely if $p_K(x) < 1$ then x is an interior point (why?), so

$$p_K(x) < 1 \iff x \in K.$$

Now define $\ell(y) = 1$, so $\ell(ay) = a$ for $a \in \mathbb{R}$. Since $y \notin K$ it is not an interior point and so $p_K(y) \geq 1$. Thus $p_K(ay) \geq a$ for $a \geq 0$ and also, trivially, for $a < 0$ (since $p_K \geq 0$). Thus

$$p_K(ay) \geq \ell(ay)$$

for all $a \in \mathbb{R}$. By Hahn-Banach, with Y the one dimensional space $\{ay\}$, ℓ may be extended to all of x so that $p_K(x) \geq \ell(x)$ which implies (2.1). $\square$

An extension of this is the following

THEOREM 2.3. *Let H, M be disjoint convex subsets of X, at least one of which has an interior point. Then H and M can be separate by a hyperplane $\ell(x) = c$: there is a linear functional ℓ and $c \in \mathbb{R}$ such that*

$$\ell(u) \leq c \leq \ell(v) \forall u \in H,\; v \in M.$$

Proof. The proof rests on a trick of applying the hyperplane separation theorem with the set

$$K = H - M = \{u - v \ : \ u \in H \text{ and } v \in M\}$$

and the point $y = 0$. Note that $0 \notin K$ since $H \cup M = \emptyset$. Since K has an interior point (why?), we see that there is a linear functional such that $\ell(x) \leq 0$ for all $x \in K$. But then $\ell(u) \leq \ell(v)$ for all $u \in H$, $v \in M$. □

In many applications, one wants to consider a vector space X over $\mathbb{C}$. Of course, then X is also a vector space over $\mathbb{R}$ so the real Hahn-Banach theorem applies. Using this one can show the following

Theorem 2.4 (Complex Hahn-Banach, Bohenblust and Sobczyk (1938) and Soukhomlinoff (1938)). *Let X be a linear space over $\mathbb{C}$ and $p : X \to [0, \infty)$ such that*

(1) $p(ax) = |a|p(x) \forall a \in \mathbb{C}, x \in X$.
(2) $p(x + y) \leq p(x) + p(y)$ (sub-additivity).

Let Y be a $\mathbb{C}$ linear subspace of X and $\ell : Y \to \mathbb{C}$ a linear functional such that

$$|\ell(y)| \leq p(y) \tag{2.2}$$

for all $y \in Y$. Then ℓ can be extended to all of X so that (2.2) *holds for all $y \in X$.*

Remark. A function p that satisfies (1) and (2) is called a semi-norm. It is a norm if $p(x) = 0 \implies x = 0$.

Proof. Let $\ell_1(y) = \operatorname{Re} \ell(y)$, the real part of ℓ. Then ℓ_1 is a real linear functional and $-\ell_1(\mathrm{i}y) = -\operatorname{Re} \mathrm{i}\ell(y) = \operatorname{Im} \ell(y)$, the imaginary part of ℓ. Thus

$$\ell(y) = \ell_1(y) - \mathrm{i}\ell_1(\mathrm{i}y). \tag{2.3}$$

Clearly $|\ell_1(y)| \leq p(y)$ so by the real Hahn-Banach theorem we can extend ℓ_1 to all of X so that $\ell_1(y) \leq p(y)$ for all $y \in X$. Since $-\ell_1(y) = \ell_1(-y) \leq p(-y) = p(y)$, we have $|\ell_1(y)| \leq p(y)$ for all $y \in X$. Now define the extension of ℓ via (2.3). Given $y \in X$ let $\theta = \arg \ln \ell(y)$. Thus $\ell(y) = \mathrm{e}^{\mathrm{i}\theta}\ell_1(\mathrm{e}^{-\mathrm{i}\theta}y)$ (why?). So,

$$|\ell(y)| = |\ell_1(\mathrm{e}^{-\mathrm{i}\theta}y)| \leq p(y).$$

□

Lax gives another beautiful extension of Hahn-Banach, due to Agnew and Morse, which involves a family of commuting linear maps. We will cover a simplified version of this next time.

LECTURE 3

Applications of Hahn-Banach

Reading: §4.1, §4.2 of Lax.

To get an idea what one can do with the Hahn-Banach theorem let's consider a concrete application on the linear space $X = B(S)$ of all real valued bounded functions on a set S. $B(S)$ has a natural partial order, namely $x \leq y$ if $x(s) \leq y(s)$ for all $s \in S$. If $0 \leq x$ then x is nonnegative. On $B(S)$ a positive linear functional ℓ satisfies $\ell(y) \geq 0$ for all $y \geq 0$.

THEOREM 3.1. *Let Y be a linear subspace of $B(S)$ that contains $y_0 \geq 1$, so $y_0(s) \geq 1$ for all $s \in S$. If ℓ is a positive linear functional on Y then ℓ can be extended to all of B as a positive linear functional.*

This theorem can be formulated in an abstract context as follows. The nonnegative functions form a cone, where

DEFINITION 3.1. A subset $P \subset X$ of a linear space over $\mathbb{R}$ is a cone if $tx + sy \in P$ whenever $x, y \in P$ and $t, s \geq 0$. A linear functional on X is P-nonnegative if $\ell(x) \geq 0$ for all $x \in P$.

THEOREM 3.2. *Let $P \subset X$ be a cone with an interior point x_0. If Y is a subspace containing x_0 on which is defined a $P \cap Y$-positive linear functional ℓ, then ℓ has an extension to X which is P-positive.*

PROOF. Define a dominating function p as follows

$$p(x) = \inf\{\ell(y) \ : \ y - x \in P, \ y \in Y\}.$$

Note that $y_0 - tx \in P$ for some $t > 0$ (since y_0 is interior to P), so $\frac{1}{t}y_0 - x \in P$. This shows that $p(x)$ is well defined. It is clear that p is positive homegeneous. To see that it is sub-additive, let $x_1, x_2 \in X$ and let y_1, y_2 be so that $y_j - x_j \in P$. Then $y_1 + y_2 - (x_1 + x_2) \in P$, so

$$p(x_1 + x_2) \leq \ell(y_1) + \ell(y_2).$$

Minimizing over $y_{1,2}$ gives sub-additivity.

Since $\ell(x) = \ell(x - y) + \ell(y) \leq \ell(y)$ if $x \in Y$ and $y - x \in P$ we conclude that $\ell(x) \leq p(x)$ for all $x \in Y$. By Hahn-Banach we may extend ℓ to all of X so that $\ell(x) \leq p(x)$ for all x.

Now let $x \in P$. Then $p(-x) \leq 0$ (why?), so $-\ell(x) = \ell(-x) \leq 0$ which shows that ℓ is P-positive. $\square$

The theorem on $B(S)$ follows from the Theorem 3.2 once we observe that the condition $y_0 \geq 1$ implies that y_0 is an interior point of the cone of positive functions. The linear functional that one constructs in this way is monotone:

$$x \leq y \implies \ell(x) \leq \ell(y),$$

which is in fact equivalent to positivity. This can allow us to do a little bit of analysis, even though we haven't introduced notions of topology or convergence.

To see how this could work, let's apply the result to the Riemann integral on $C[0,1]$. We conclude the existence of a positive linear functional $\ell : B[0,1] \to \mathbb{R}$ that gives $\ell(f) = \int_0^1 f(x)\mathrm{d}x$. This gives an "integral" of arbitrary bounded functions, without a measurability condition. Since the integral is a linear functional, it is finitely additive. Of course it is not countably additive since we know from real analysis that such a thing doesn't exist. Furthermore, the integral is not uniquely defined.

Nonetheless, the extension is also not arbitrary, and the constraint of positivity actually pins down $\ell(f)$ for many functions. For example, consider $\chi_{(a,b)}$ (the characteristic function of an open interval). By positivity we know that

$$f \le \chi_{(a,b)} \le g, \ f, g \in C[0,1] \implies \int_0^1 f(x)\mathrm{d}x \le \ell(\chi_{(a,b)}) \le \int_0^1 g(x)\mathrm{d}x.$$

Taking the sup over f and inf over g, using properties only of the Riemann integral, we see that $\ell(\chi_{(a,b)}) = b - a$ (hardly a surprising result). Likewise, we can see that $\ell(\chi_U) = |U|$ for any open set, and by finite additivity $\ell(\chi_F) = 1 - \ell(F^c) = 1 - |F^c| = |F|$ for any closed set ($|\cdot|$ is Lebesgue measure). Finally if $S \subset [0,1]$ and

$$\sup \{\ell(F) \ : \ F \text{ closed and } F \subset S\} \ = \ \inf \{\ell(U) \ : \ U \text{ open and } U \supset S\}$$

then $\ell(\chi_S)$ must be equal to these two numbers. We see that $\ell(\chi_S) = |S|$ for any Lebesgue measurable set. We have just painlessly constructed Lebesgue measure from the Riemann integral without using any measure theory!

The rigidity of the extension is a bit surprising if we compare with what happens in finite dimensions. For instance, consider the linear functional $\ell(x,0) = x$ defined on $Y = \{(x,0)\} \subset \mathbb{R}^2$. Let P be the cone $\{(x,y) \ : \ |y| \le \alpha x\}$. So ℓ is $P \cap Y$-positive. To extend ℓ to all of $\mathbb{R}^2$ we need to define $\ell(0,1)$. To keep the extension positive we need only require

$$y\ell(0,1) + 1 \ge 0$$

if $|y| \le \alpha$. Thus we must have $|\ell(0,1)| \le 1/\alpha$, and any choice in this interval will work. Only when $\alpha = \infty$ and the cone degenerates to a half space does the condition pin $\ell(0,1)$ down. So some interesting things happen in ∞ dimensions.

A second example application assigning a limiting value to sequences

$$\mathbf{a} = (a_1, a_2, \ldots).$$

Let B denote the space of all bounded $\mathbb{R}$-valued sequences and let L denote the subspace of sequences with a limit. We quickly conclude that there is an positive linear extension of the positive linear functional lim to all of B. We would like to conclude a little more, however. After all, for convergent sequences,

$$\lim a_{n+k} = \lim a_n$$

for any k.

To formalize this property we define a linear map on B by

$$T(a_1, a_2, \ldots) \ = \ (a_2, a_3, \ldots).$$

(T is the "backwards shift" operator or "left translation.") Here,

Definition 3.2. Let X_1, X_2 be linear spaces over a field F. A linear map $T : X_1 \to X_2$ is a function such that

$$T(x + ay) = T(x) + aT(y) \quad \forall x, y \in X_1 \text{ and } a \in F.$$

Thus $\lim T\mathbf{a} = \lim \mathbf{a}$ for $\mathbf{a} \in L$. We would like the extension to have this property. Can this be done?

The answer is given by the following a simplified version of the Theorem of Agnew and Morse given in Chapter 3 of Lax:

THEOREM 3.3. *Let X be a linear space over $\mathbb{R}$ and let $T : X \to X$ be a linear map. Let $p : X \to \mathbb{R}$ be a positive homogeneous, sub-additive function invariant under T:*

$$p(Tx) = p(x) \quad \forall x \in X.$$

Let ℓ be a linear functional defined on a subspace $Y \subset X$ invariant under T,

$$Ty \in Y \quad \forall y \in Y.$$

If ℓ is invariant under T — $\ell(Ty) = \ell(y)$ — and dominated by p then ℓ may be extended to all of X so that the extension is also invariant under T and dominated by p.

REMARK. Applying this to the space B of bounded sequences and letting

$$p(\mathbf{a}) = \limsup \mathbf{a},$$

we find that there is an extension of the linear functional lim from L to all of B such that for all $\mathbf{a} = (a_1, a_2, \ldots)$

(1) $\lim_{n\to\infty} a_{n+k} = \lim_{n\to\infty} a_n$ for all k.
(2) $\liminf \mathbf{a} \le \lim \mathbf{a} \le \limsup \mathbf{a}$. (The first inequality follows from the second applied to $-\mathbf{a}$.)

PROOF. Let us define

$$q(x) \ = \ \inf p(Ax),$$

where the infimum is over all convex combinations of powers of T,

$$A = \sum_{j=0}^{n} a_j T^j \tag{3.1}$$

where $T^0 = \mathbf{1}$, the identity operator on X, $\mathbf{1}x = x$, and a_j is any finite, non-negative sequence with $\sum_j a_j = 1$. Clearly $q(x)$ is positive homogeneous and $q(x) \le p(x)$. Also, since ℓ and Y are invariant under any A of the form (3.1) we have

$$\ell(y) \le q(y) \quad \forall y \in Y.$$

To apply Hahn-Banach with q as the dominating function we need to show it is sub-additivie. To this end, let $x, y \in X$ and let A, B be of the form (3.1) so that

$$p(Ax) \le q(x) + \epsilon \text{ and } p(By) \le q(y) + \epsilon.$$

Then $A \circ B = B \circ A$ is of the form (3.1) and

$$q(x+y) \le p(AB(x+y)) \le p(BAx) + p(ABy) \le p(Ax) + p(By) \le q(x) + q(y) + 2\epsilon.$$

Thus q is sub-additive.

So now we know that an extension exists to all of X with $\ell(x) \le q(x)$ for all x. However,

$$q(x - Tx) \le p\left(\frac{1}{n}\sum_{j=0}^{n} T^j(\mathbf{1} - T)x\right) = \frac{1}{n} p(x - T^{n+1}x) \le \frac{1}{n}\left(p(x) + p(-x)\right) \to 0.$$

So $q(x - Tx) = 0$ for all x and thus $\ell(x) - \ell(Tx) = 0$ for all x. $\square$

Part 2

Banach Spaces

LECTURE 4

Normed and Banach Spaces

Reading: §5.1 of Lax.

The Hahn-Banach theorem made use of a dominating function $p(x)$. When this function is non-negative, it can be understood roughly as a kind of "distance" from a point x to the origin. For that to work, we should have $p(x) > 0$ whenever $x \neq 0$. Such a function is called a norm:

Definition 4.1. Let X be a linear space over $F = \mathbb{R}$ or $\mathbb{C}$. A norm on X is a function $\|\cdot\| : X \to [0, \infty)$ such that

(1) $\|x\| = 0 \iff x = 0$.
(2) $\|x + y\| \leq \|x\| + \|y\|$ (subadditivity)
(3) $\|ax\| = |a| \, \|x\|$ for all $a \in F$ and $x \in X$ (homogeneity).

A normed space is a linear space X with a norm $\|\cdot\|$.

The norm on a normed space gives a metric topology if we define the distance between two points via

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

Condition 1 guarantees that two distinct points are a finite distance apart. Sub-additivity gives the triangle inequality. The metric is

(1) translation invariant: $d(x + z, y + z) = d(x, y)$
and
(2) homogeneous $d(ax, ay) = |a| d(x, y)$.

Thus any normed space is a metric space and we have the following notions:

(1) a sequence x_n converges to x, $x_n \to x$, if $d(x_n, x) = \|x_n - x\| \to 0$.
(2) a set $U \subset X$ is open if for every $x \in U$ there is a ball $\{y \ : \ \|y - x\| < \epsilon\} \subset U$.
(3) a set $K \subset X$ is closed if $X \setminus K$ is open.
(4) a set $K \subset X$ is compact if every open cover of K has a finite sub-cover.

The norm defines the topology but not the other way around. Indeed two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are equivalent if there is $c > 0$ such that

$$c \|x\|_1 \leq \|x\|_2 \leq c^{-1} \|x\|_2 \quad \forall x \in X.$$

Equivalent norms define the same topology. (Why?)

Recall from real analysis that a metric space X is complete if every Cauchy sequence x_n converges in X. In a normed space, a Cauchy sequence x_n is one such that

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \text{ such that } n, m > N \implies \|x_n - x_m\| < \epsilon.$$

A complete normed space is called a Banach space.

Not every normed space is complete. For example $C[0,1]$ with the norm

$$\|f\|_1 = \int_0^1 |f(x)|\mathrm{d}x$$

fails to be complete. (It is, however, a complete space in the uniform norm, $\|f\|_u = \sup_{x\in[0,1]} |f(x)|$.) However, every normed space X has a completion, defined abstractly as a set of equivalence classes of Cauchy sequences in X. This space, denoted $\overline{X}$ is a Banach space.

Examples of Normed and Banach spaces

(1) For each $p \in [1,\infty)$ let

$$\ell_p = \{p \text{ summable sequences}\} = \{(a_1, a_2, \dots) \mid \sum_{j=1}^{\infty} |a_j|^p < \infty\}.$$

Define a norm on ℓ_p via

$$\|\mathbf{a}\|_p \;=\; \left[\sum_{j=1}^{\infty} |a_j|^p\right]^{\frac{1}{p}}.$$

Then ℓ_p is a Banach space.

(2) Let

$$\ell_\infty = \{\text{bounded sequences}\} \;=\; B(\mathbb{N}),$$

with norm

$$\|\mathbf{a}\|_\infty = \sup_j |a_j|. \tag{$\star$}$$

Then ℓ_∞ is a Banach space.

(3) Let

$$c_0 = \{\text{sequences converging to } 0\} \;=\; \{(a_1, a_2, \dots) \mid \lim_{j\to\infty} a_j = 0\},$$

with norm ($\star$). Then c_0 is a Banach space.

(4) Let

$$\begin{aligned}\mathcal{F} = \{&\text{sequences with finitely many non-zero terms}\}\\ &= \{(a_1, a_2, \dots) \mid \exists N \in N \text{such that } n \geq N \implies a_n = 0\}.\end{aligned}$$

Then for any $p \geq 1$, $\mathcal{F}_p = (\mathcal{F}, \|\cdot\|_p)$ is a normed space which is not complete. The completion of $\mathcal{F}_p$ is isomorphic to ℓ_p.

(5) Let $D \subset \mathbb{R}^d$ be a domain and let $p \in [1,\infty)$.

(a) Let $X = C_c(D)$ be the space of continuous functions with compact support in D, with the norm

$$\|f\|_p \;=\; \left[\int_D |f(x)|^p \mathrm{d}x\right]^{\frac{1}{p}}.$$

Then X is a normed space, which is not complete. Its completion is denoted $L^p(D)$ and may be identified with the set of equivalence classes of measurable functions $f : D \to \mathbb{C}$ such that

$$\int_D |f(x)|^p \mathrm{d}x < \infty \quad \text{(Lebesgue measure)},$$

with two functions f, g called equivalent if $f(x) = g(x)$ for almost every x.

(b) Let X denote the set of C^1 functions on D such that

$$\int_D |f(x)|^p \mathrm{d}x < \infty \quad \text{and} \quad \int_D |\partial_j f(x)|^p \mathrm{d}x < \infty, \ j = 1, \dots, n.$$

Put the following norm on X,

$$\|f\|_{1,p} \ = \ \left[\int_D |f(x)|^p \mathrm{d}x + \sum_{j=1}^n \int_D |\partial_j f(x)|^p \mathrm{d}x.\right]^{\frac{1}{p}}.$$

Then X is a normed space which is not complete. Its completion is denoted $W^{1,p}(D)$ and is called a Sobolev space and may be identified with the subspace of $L^p(D)$ consisting of (equivalence classes) of functions all of whose first derivatives are in $L^p(D)$ in the sense of distributions. (We'll come back to this.)

Note that $\mathbf{a} \in \ell_1$ is summable,

$$\sum_j a_j \ \le \ \|\mathbf{a}\|_1 .$$

THEOREM 4.1 (Hölder's Inqequality). *Let $1 < p < \infty$ and let q be such that*

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

If $\mathbf{a} \in \ell_p$ and $\mathbf{b} \in \ell_q$ then

$$\mathbf{ab} = (a_1 b_1, a_2 b_2, \dots) \in \ell_1$$

and

$$\left|\sum_{j=1}^\infty a_j b_j\right| \le \|\mathbf{a}\|_p \|\mathbf{b}\|_q .$$

PROOF. First note that for two non-negative numbers a, b it holds that

$$ab \le \frac{1}{p} a^p + \frac{1}{q} b^q.$$

For $p = q = 2$ this is the familiar "arithmetic-geometric mean" inequality which follows since $(a-b)^2 \ge 0$. For general a, b this may be seen as follows. The function $x \mapsto \exp(x)$ is convex: $\exp(tx + (1-t)y) \le t\exp(x) + (1-t)\exp(y)$. (Recall from calculus that a C^2 function f is convex if $f'' \ge 0$.) Thus,

$$ab = \exp\left(\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q\right) \le \frac{1}{p}\exp(\ln a^p) + \frac{1}{q}\exp(\ln b^q) \ = \ \frac{1}{p} a^p + \frac{1}{q} b^q.$$

Applying this bound co-ordinate wise and summing up we find that

$$\|\mathbf{ab}\|_1 \ \le \ \frac{1}{p}\|\mathbf{a}\|_p^p + \frac{1}{q}\|\mathbf{b}\|_q^q .$$

The result follows from this bound by "homogenization:" we have

$$\|\mathbf{a}\|_p\,, \|\mathbf{b}\|_q \implies \|\mathbf{ab}\| \leq 1,$$

so

$$\left\| \frac{\mathbf{a}}{\|\mathbf{a}\|_p} \frac{\mathbf{b}}{\|\mathbf{b}\|_q} \right\|_1 \leq 1$$

from which the desired estimate follows by homogeneity of the norm. $\square$

Similarly, we have

THEOREM 4.2 (Hölder's Inequality). *Let $1 < p < \infty$ and let q be the conjugate exponent. If $f \in L^p(D)$ and $g \in L^q(D)$ then $fg \in L^1(D)$ and*

$$\int_D |f(x)g(x)|\mathrm{d}x \ \leq \|f\|_p \, \|g\|_q \, .$$

LECTURE 5

Noncompactness of the Ball and Uniform Convexity

Reading: §5.1 and §5.2 of Lax.

First a few more definitions:

Definition 5.1. A normed space X over $F = \mathbb{R}$ or $\mathbb{C}$ is called separable if it has a countable, dense subset.

Most spaces we consider are separable, with a few notable exceptions.

(1) The space M of all signed (or complex) measures μ on $[0, 1]$, say, with norm

$$\|\mu\| = \int_0^1 |\mu|(\mathrm{d}x).$$

Here $|\mu|$ denotes the total variation of μ,

$$|\mu|(A) = \sup_{\text{Partitions } A_1,\ldots,A_n \text{ of } A} \sum_{j=1}^{n} |\mu(A_j)|.$$

This space is a Banach space. Since the point mass

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

is an element of A and $\|\delta_x - \delta_y\| = 2$ if $x \neq y$, we have an uncountable family of elements of M all at a fixed distance of one another. Thus there can be no countable dense subset. (Why?)

(2) ℓ_∞ is also not separable. To see this, note that to each subset of $A \subset \mathbb{N}$ we may associate the sequence χ_A, and

$$\|\chi_A - \chi_B\|_\infty = 1$$

if $A \neq B$.

(3) ℓ_p is separable for $1 \leq p < \infty$.

(4) $L^p(D)$ is separable for $1 \leq p < \infty$.

(5) $L^\infty(D)$ is not separable.

Noncompactness of the Unit Ball

Theorem 5.1 (F. Riesz). *Let X be a normed linear space. Then the closed unit ball $B_1(0) = \{x : \|X\| \leq 1\}$ is compact if and only if X is finite dimensional.*

Proof. The fact that the unit ball is compact if X is finite dimensional is the Heine-Borel Theorem from Real Analysis.

To see the converse, we use the following

LEMMA 5.2. *Let Y be a closed proper subspace of a normed space X. Then there is a unit vector $z \in X$, $\|z\| = 1$, such that*

$$\|z - y\| > \frac{1}{2} \quad \forall y \in Y.$$

PROOF OF LEMMA. Since Y is proper, $\exists x \in X \setminus Y$. Then

$$\inf_{y \in Y} \|x - y\| = d > 0.$$

(This is a property of closed sets in a metric space.) We do not know the existence of a minimizing y, but we can certainly find y_0 such that

$$0 < \|x - y_0\| < 2d.$$

Let $z = \frac{x - y_0}{\|x - y_0\|}$. Then

$$\|z - y\| \;=\; \frac{\|x - y_0 - \|x - y_0\|\, y\|}{\|x - y_0\|} \geq \frac{d}{2d} = \frac{1}{2}.$$

□

Returning to the proof of the Theorem, we will use the fact that every sequence in a compact metric space has a convergent subsequence. Thus it suffices to show that if X is infinite dimensional then there is a sequence in $B_1(0)$ with no convergent subsequence.

Let y_1 be any unit vector and construct a sequence of unit vectors, by induction, so that

$$\|y_k - y\| > \frac{1}{2} \quad \forall y \in \operatorname{span}\{y_1, \ldots, y_{k-1}\}.$$

(Note that $\operatorname{span}\{y_1, \ldots, y_{k-1}\}$ is finite dimensional, hence complete, and thus a closed subspace of X.) Since X is finite dimensional the process never stops. No subsequence of y_j can be Cauchy, much less convergent. □

Uniform convexity

The following theorem may be easily shown using compactness:

THEOREM 5.3. *Let X be a finite dimensional linear normed space. Let K be a closed convex subset of X and z any point of X. Then there is a unique point of K closer to z than any other point of K. That is there is a unique solution $y_0 \in K$ to the minimization problem*

$$\|y_0 - z\| = \inf_{y \in K} \|y - z\| . \tag{$\star$}$$

Try to prove this theorem. (The existence of a minimizer follows from compactness; the uniqueness follows from convexity.)

The conclusion of theorem does not hold in a general infinite dimensional space. Nonetheless there is a property which allows for the conclusion, even though compactness fails!

DEFINITION 5.2. A normed linear space X is uniformly convex if there is a function $\epsilon : (0, \infty) \to (0, \infty)$, such that

(1) ϵ is increasing.
(2) $\lim_{r \to 0} \epsilon(r) = 0$.
(3) $\left\| \frac{1}{2}(x + y) \right\| \leq 1 - \epsilon(\|x - y\|)$ for all $x, y \in B_1(0)$, the unit ball of X.

THEOREM 5.4 (Clarkson 1936). *Let X be a uniformly convex Banach space, K a closed convex subset of X, and z any point of X. Then the minimization problem $(\star)$ has a unique solution $y_0 \in K$.*

PROOF. If $z \in K$ then $y_0 = z$ is the solution, and is clearly unique.

When $z \notin K$, we may assume $z = 0$ (translating z and K if necessary). Let

$$s = \inf_{y \in K} \|y\| .$$

So $s > 0$. Now let $y_n \in K$ be a minimizing sequence, so

$$\|y_n\| \to s.$$

Now let $x_n = y_n / \|y_n\|$, and consider

$$\frac{1}{2}(x_n + x_m) = \frac{1}{2\|y_n\|} y_n + \frac{1}{2\|y_m\|} y_m = \left(\frac{1}{2\|y_n\|} + \frac{1}{2\|y_m\|} \right) (ty_n + (1-t)y_m)$$

for suitable t. So $ty_n + (1-t)y_m \in K$ and

$$\|ty_n + (1-t)y_m\| \geq s.$$

Thus

$$1 - \epsilon(\|x_n - x_m\|) \geq \frac{1}{2} \left(\frac{s}{\|y_n\|} + \frac{s}{\|y_m\|} \right) \to 1.$$

We conclude that x_n is a Cauchy sequence, from which it follows that y_n is Cauchy. The limit $y_0 \in \lim_n y_n$ exists in K since X is complete and K is closed. Clearly $\|y_0\| = s$. $\square$

Warning: Not every Banach space is uniformly convex. For example, the space $C(D)$ of continuous functions on a compact set D is not uniformly convex. It may even happen that

$$\left\| \frac{1}{2}(f+g) \right\|_\infty = 1$$

for unit vectors f and g. (They need only have disjoint support.) Lax gives an example of a closed convex set in $C[-1,1]$ in which the minimization problem $(\star)$ has no solution.

It can also happen that a solution exists but is not unique. For example, in $C[-1,1]$ let $K = \{$functions that vanish on $[-1,0]\}$. and let $f = 1$ on $[-1,1]$. Clearly

$$\sup_x |f(x) - g(x)| \geq 1 \quad \forall g \in K,$$

and the distance 1 is attained for any $g \in K$ that satisfies $0 \leq g(x) \leq 1$.

LECTURE 6

Linear Functionals on a Banach Space

Reading: §8.1 and §8.2 of Lax

DEFINITION 6.1. A linear functional $\ell : X \to F$ on a normed space X over $F = \mathbb{R}$ or $\mathbb{C}$ is bounded if there is $c < \infty$ such that

$$|\ell(x)| \leq c \|x\| \quad \forall x \in X.$$

The inf over all such c is the norm $\|\ell\|$ of ℓ,

$$\|\ell\| = \sup_{x \neq 0,\ x \in X} \frac{|\ell(x)|}{\|x\|}. \tag{$\star$}$$

THEOREM 6.1. *A linear functional ℓ on a normed space is bounded if and only if it is continuous.*

PROOF. It is useful to recall that

THEOREM 6.2. *Let X, Y be metric spaces. Then $f : X \to Y$ is continuous if and only if $f(x_n)$ is a convergent sequence in Y whenever x_n is convergent in X.*

REMARK. Continuity $\implies$ the sequence condition in any topological space. That the sequence condition $\implies$ continuity follows from the fact that the topology has a countable basis at a point. (In a metric space, $B_{2^{-n}}(x)$, say.) Specifically, suppose the function is not continuous. Then there is an open set $U \subset Y$ such that $f^{-1}(U)$ is not open. So there is $x \in f^{-1}(U)$ such that for all n $B_{2^{-n}}(x) \not\subset f^{-1}(U)$. Now let x_n be a sequence such that

(1) $x_n \in B_{2^{-n}}(x) \setminus f^{-1}(U)$ for n odd.
(2) $x_n = x$ for n even.

Clearly $x_n \to x$. However, $f(x_n)$ cannot converge since $\lim_k f(x_{2k}) = f(x)$ and any limit point of $f(x_{2k+1})$ lies in the closed set U^c containing all the points $f(x_{2k+1})$.

In the normed space context, note that

$$|\ell(x_n) - \ell(x)| \leq \|\ell\| \|x_n - x\|,$$

so boundedness certainly implies continuity.

Conversely, if ℓ is unbounded then we can find vectors x_n so that $\ell(x_n) \geq n \|x_n\|$. Since this inequality is homogeneous under scaling, we may suppose that $\|x_n\| = 1/\sqrt{n}$, say. Thus $x_n \to 0$ and $\ell(x_n) \to \infty$, so ℓ is not continuous. □

The set X' of all bounded linear functionals on X is called the dual of X. It is a linear space, and in fact a normed space under the norm ($\star$). (It is straightforward to show that ($\star$) defines a norm.)

THEOREM 6.3. *The dual X' of a normed space X is a Banach space.*

PROOF. We need to show X' is complete. Suppose ℓ_m is a Cauchy Sequence. Then for each $x \in X$ we have

$$|\ell_n(x) - \ell_m(x)| \le \|\ell_n - \ell_m\| \, \|x\| \,,$$

so $\ell_n(x)$ is a Cauchy sequence of scalars. Let

$$\ell(x) = \lim_{n\to\infty} \ell_n(x) \quad \forall x \in X.$$

It is easy to see that ℓ is linear. Let us show that it is bounded. Since $| \, \|\ell_n\| - \|\ell_m\| \, | \le \|\ell_n - \ell_m\|$ (this follows from sub-additivity), we see that the sequence $\|\ell_n\|$ is Cauchy, and thus bounded. So,

$$|\ell(x)| \le \sup_n \|\ell_n\| \, \|x\|$$

and ℓ is bounded. Similarly,

$$|\ell_n(x) - \ell(x)| \le \sup_{m \ge n} \|\ell_n - \ell_m\| \, \|x\| \,,$$

so

$$\|\ell_n - \ell\| \le \sup_{m\ge n} \|\ell_n - \ell_m\|$$

and it follows that $\ell_n \to \ell$. $\square$

Of course, all of this could be vacuous. How do we know that there are any bounded linear functionals? Here the Hahn-Banach theorem provides the answer.

THEOREM 6.4. *Let $y_1, \ldots, y_N$ be N linearly independent vectors in a normed space X and $\alpha_1, \ldots, \alpha_N$ arbitrary scalars. Then there is a bounded linear functional $\ell \in X'$ such that*

$$\ell(y_j) = \alpha_j, \quad j = 1, \ldots, N.$$

PROOF. Let $Y = \operatorname{span}\{y_1, \ldots, y_N\}$ and define ℓ on Y by

$$\ell(\sum_{j=1}^N b_j) = \sum_{j=1}^N b_j \alpha_j.$$

(We use linear independence here to guarantee that ℓ is well defined.) Clearly, ℓ is linear. Furthermore, since Y is finite dimensional ℓ is bounded.

(Explicitly, since any two norms on a finite dimensional space are equivalent, we can find c such that

$$\sum_j |b_j| \, \|y_j\| \le c \left\| \sum_j b_j y_j \right\| .$$

Thus, $\ell(y) \le c \sup_j |\alpha_j| \, \|y\|$ for $y \in Y$.)

Thus ℓ is a linear functional on Y dominated by the norm $\|\cdot\|$. By the Hahn-Banach theorem, it has an extension to X that is also dominated by $\|\cdot\|$, i.e., that is bounded. $\square$

A closed subspace Y of a normed space X is itself a normed space. If X is a Banach space, so is Y. A linear functional $\ell \in X'$ on X can be restricted to Y and is still bounded. That is there is a restriction map $R : X' \to Y'$ such that

$$R(\ell)(y) = \ell(y) \quad \forall y \in Y.$$

It is clear that R is a linear map and that

$$\|R(\ell)\| \le \|\ell\| \,.$$

The Hahn-Banach Theorem shows that R is surjective. On the other hand, unless $Y = X$, the kernel of R is certainly non-trivial. To see this, let x be a vector in $X \setminus Y$ and define ℓ on $\operatorname{span}\{x\} \cup Y$ by

$$\ell(ax + y) = a \quad \forall a \in F \text{ and } y \in Y.$$

Then ℓ is bounded on the closed subspace $\operatorname{span}\{x\} \cup Y$ and by Hahn-Banach has a closed extension. Clearly $R(\ell) = 0$. The kernel of R is denoted $Y^\perp$, so

$$Y^\perp = \{\ell \in X' \ : \ \ell(y) = 0 \quad \forall y \in Y\},$$

and is a Banach space.

Now the quotient space X/Y is defined to be the set of "cosets of Y,"

$$\{Y + x \ : \ x \in X\}.$$

The coset $Y + x$ is denoted $[x]$. The choice of label x is, of course, not unique as $x + y$ with $y \in Y$ would do just as well. It is a standard fact that

$$[x_1] + [x_2] = [x_1 + x_2]$$

gives a well defined addition on Y/X so that it is a linear space.

Lemma 6.5. *If Y is a closed subspace of a normed space then*

$$\|[x]\| \ = \ \inf_{y \in Y} \|x + y\|$$

is a norm on X/Y. If X is a Banach space, so is X/Y.

Proof. Exercise. □

A bounded linear functional that vanishes on Y, that is an element $\ell \in Y^\perp$, can be understood as a linear functional on X/Y since the definition

$$\ell([x]) = \ell(x)$$

is unambiguous. Thus we have a map $J : Y^\perp \to (X/Y)'$ defined by $J(\ell)([x]) = \ell(x)$. Conversely, there is a bounded linear map $\Pi : X \to X/Y$ given by $\Pi(x) = [x]$ and any linear functional $\ell \in (X/Y)'$ pulls back to a bounded linear functional $\ell \circ \Pi$ in $Y^\perp$. Clearly $J(\ell \circ \Pi) = \ell$. Thus we have, loosely, that

$$(X/Y)' = Y^\perp.$$

LECTURE 7

Isometries of a Banach Space

Reading: §5.3 of Lax.

DEFINITION 7.1. Let X, Y be normed spaces. A linear map $T : X \to Y$ is bounded if there is $c > 0$ such that

$$\|T(x)\| \leq c \|x\| .$$

The norm of T is the smallest such c, that is

$$\|T\| \ = \ \sup_{x \neq 0} \frac{\|T(x)\|}{\|x\|}.$$

THEOREM 7.1. *A linear map $T : X \to Y$ between normed spaces X and Y is continuous if and only if it is bounded.*

REMARK. The proof is a simple extension of the corresponding result for linear functionals.

An isometry of normed spaces X and Y is a map $M : X \to Y$ such that

(1) M is surjective.
(2) $\|M(x) - M(y)\| = \|x - y\|$.

Clearly translations $T_u : X \to X$, $T_u(x) = x + u$ are isometries of a normed linear space. A linear map $T : X \to Y$ is an isometry if T is surjective and

$$\|T(x)\| = \|x\| \quad \forall x \in X.$$

A map $M : X \to Y$ is affine if $M(x) - M(0)$ is linear. So, M is affine if it is the composition of a linear map and a translation.

THEOREM 7.2 (Mazur and Ulam 1932). *Let X and Y be normed spaces over $\mathbb{R}$. Any isometry $M : X \to Y$ is an affine map.*

REMARK. The theorem conclusion does not hold for normed spaces over $\mathbb{C}$. In that context any isometry is a real -affine map ($M(x) - M(0)$ is real linear), but not necessarily a complex-affine map. For example on $C([0,1], \mathbb{C})$ the map $f \mapsto \overline{f}$ (complex conjugation) is an isometry and is not complex linear.

PROOF. It suffices to show $M(0) = 0 \implies M$ is linear. To prove linearity it suffices to show

$$M\left(\frac{1}{2}(x+y)\right) = \frac{1}{2}(M(x) + M(y)) \quad \forall x, y \in X.$$

(Why?)

Let x and y be points in X and $z = \frac{1}{2}(x+y)$. Note that

$$\|x - z\| = \|y - z\| = \frac{1}{2} \|x - y\| ,$$

so z is "half-way between x and y." Let

$$x' = M(x),\ y' = M(y),\ z' = M(z).$$

We need to show $2z' = x' + y'$. Since M is an isometry,

$$\|x' - z'\| = \|y' - z'\| = \frac{1}{2}\|x' - y'\|,$$

and all of these are equal to $\frac{1}{2}\|x - y\|$. So z' is "half-way between x' and y'." It may happen that $\frac{1}{2}(x' + y')$ is the unique point of Y with this property (in which case we are done). this happens, for instance, if the norm in Y is <u>strictly sub-additive</u>, meaning

$$\beta x' \neq \alpha y' \implies \|x' + y'\| < \|x'\| + \|y'\|.$$

In general, however, there may be a number of points "half-way between x' and y'."

So, let

$$A_1 = \{u \in X \ :\ \|x - u\| = \|y - u\| = \frac{1}{2}\|x - y\|\},$$

and

$$A_1' = \{u' \in Y \ :\ \|x' - u'\| = \|y' - u'\| = \frac{1}{2}\|x' - y'\|\}.$$

Since M is an isometry, we have $A_1' = M(A_1)$. Let d_1 denote the diameter of A_1,

$$d_1 = \sup_{u,v \in A_1} \|u - v\|.$$

This is also the diameter of A_1'. Now, let

$$A_2 = \left\{u \in A_1 \ :\ v \in A_1 \implies \|u - v\| \leq \frac{1}{2}d_1\right\},$$

the set of "centers of A_1." Note that $z \in A_2$ since if $u \in A_1$ then $2z - u \in A_1$:

$$\|x - (2z - u)\| = \|u - y\| = \|x - u\| = \|y - (2z - u)\|.$$

Similarly, let

$$A_2' = \left\{u' \in A_1' \ :\ v' \in A_1' \implies \|u' - v'\| \leq \frac{1}{2}d_1\right\}.$$

Again, since M is an isometry we have $A_2' = M(A_2)$. In a similar way, define decreasing sequences of sets, A_j and A_j', inductively by

$$A_j = \{u \in A_{j-1} \ :\ v \in A_{j-1} \implies \|u - v\| \leq \frac{1}{2}\operatorname{diam}(A_{j-1})\},$$

and

$$A_j' = \{u' \in A_{j-1}' \ :v' \in A_{j-1}' \implies \|u' - v'\| \leq \frac{1}{2}\operatorname{diam}(A_{j-1}')\}.$$

Again $M(A_j) = A_j'$ and $z \in A_j$ since A_{j-1} is invariant under inversion around z: $u \in A_{j-1} \implies 2z - u \in A_{j-1}$. Since $\operatorname{diam}(A_j) \leq 2^{1-j}d_1$ we conclude that

$$\bigcap_{j=1}^{\infty} = \{z\},\ \text{and}\ \bigcap_{j=1}^{\infty} A_j' = \left\{\frac{1}{2}(x' + y')\right\}.$$

Since $z' \in A_j'$ for all j, $(\star)$ follows. $\square$

Homework I

Homework I

Due: February 15, 2008

(1) (Ex. 1, Ch.1) Let X be a linear space. Let $\{S_\alpha : \alpha \in \Omega\}$ be a collection of subspaces. (Here Ω is some index set, not necessarily finite or countable.) Prove that
   (a) The sum $\sum_\alpha S_\alpha = \{x_1 + \cdots + x_n : x_j \in S_{\alpha_j}, \ \alpha_j \in \Omega, \ j = 1, \ldots, n, \ n \in \mathbb{N}\}$ is a subspace.
   (b) The intersection $\cap_{\alpha\in\Omega} S_\alpha$ is a subspace.
   (c) The union $\cup_\alpha S_\alpha$ is a subspace provided Ω is a totally ordered set and $\alpha \leq \beta \implies S_\alpha \subset S_\beta$.

(2) (Ex. 2, Ch. 4) The Cesaro means of a bounded sequence c_n are the numbers

$$M_n(\mathbf{c}) = \frac{1}{n}\sum_{j=1}^{n} c_j.$$

A bounded sequence $\mathbf{c}$ is called <u>Cesaro summable</u> if $\lim_{n\to\infty} M_n(\mathbf{c})$ exists.
   (a) Show that any convergent sequence $\mathbf{c}$ is Cesaro summable and

$$\lim_{n\to\infty} M_n(\mathbf{c}) = \lim_{n\to\infty} c_n.$$

   (b) Give an explicit example of a bounded Cesaro summable sequence that is not convergent.
   (c) Show there is a Banach limit LIM such that for any bounded Cesaro summable sequence $\mathrm{LIM}_{n\to\infty} c_n = \lim_{n\to\infty} M_N(\mathbf{c})$. (A <u>Banach limit</u> is a linear functional on the space ℓ_∞ of bounded sequences which agrees with the limit on convergent sequence, is invariant under shifts, and is bounded above and below by $\liminf$ and $\limsup$.)

(3) (Ex. 2, Ch.5) Let X be a normed linear space and Y a subspace of X. Prove that the closure of Y is a linear subspace of X.

(4) (Ex. 3, Ch.5) Show that if X is a Banach space and Y is a closed subspace of X, then the quotient space X/Y is complete.

(5) Let $M : X \to Y$ be a continuous map between normed spacse X and Y such that

$$M(0) = 0 \quad \text{and} \quad M\left(\frac{1}{2}(x+y)\right) = \frac{1}{2}M(x) + \frac{1}{2}M(y), \ \forall x, y \in X.$$

Show that M is linear.

(6) Let $\ell \in c_0^\star$ be a bounded linear functional on c_0. Prove that there is a sequence $\mathbf{b} \in \ell_1$ such that

$$\ell(\mathbf{a}) = \sum_{j=1}^{\infty} b_j a_j.$$

Conclude that $c_0^\star$ is isometrically isomorphic to ℓ_1.

(7) Let $\ell \in \ell_1^\star$ be a bounded linear functional on ℓ_1. Prove that there is a sequence $\mathbf{b} \in \ell_\infty$ such that

$$\ell(\mathbf{a}) = \sum_{j=1}^{\infty} b_j a_j.$$

Conclude that $\ell_1^\star$ is isometrically isomorphic to ℓ_∞.

(8) Show that $\ell_p^\star$ is isometrically isomorphic to ℓ_q, $1 < p < \infty$ and $1/p + 1/q = 1$.

(9) Let $\mathcal{M} = \{\mathbf{a} \in \ell_p \ : \ a_{2n} = 0, \ \forall n\} \subset \ell_p$. Show that $\mathcal{M}$ is a closed subspace and that $\ell_p/\mathcal{M}$ is isometrically isomorphic to ℓ_p.

(10) Let X_1, X_2 be Banach spaces, with norms $\|\cdot\|_1, \|\cdot\|_2$. Define the direct sum

$$X_1 \oplus X_2 = \{(x, y) \ : \ x \in X_1 \text{ and } y \in X_2\},$$

with coordinatewise addition and scalar multiplication. For each $p \in [1, \infty]$ define the norm

$$\|(x,y)\|_p = \begin{cases} (\|x\|_1^p + \|y\|_2^p)^{1/p} & 1 \le p < \infty, \\ \max\{\|x\|_1, \|y\|_2\} & p = \infty. \end{cases}$$

Show that $X_1 \oplus X_2$ is a Banach space under $\|\cdot\|_p$ and all these norms are equivalent.

(11) Consider $X = L^1(\mathbb{R}) + L^2(\mathbb{R}) = \{g + h \ : \ g \in L^1(\mathbb{R}) \text{ and } h \in L^2(\mathbb{R})\}$. Given $p \in [1, \infty)$, define

$$\|f\|_p^{\inf} = \inf\{(\|g\|_{L^1}^p + \|h\|_{L^2}^p)^{1/p} : f = g + h, \ g \in L^1 \text{ and } h \in L^2\}.$$

Show that $\|f\|_p^{\inf}$ is a norm and that all of these norms are equivalent. (Challenge: Is X a Banach space?)

(12) Consider $Y = L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, which is a subspace of X from the previous problem.
  (a) For each $p \in [1, \infty)$, define

$$\|f\|_p^+ = \frac{1}{2}\left(\|f\|_{L^1}^p + \|f\|_{L^2}^p\right)^{1/p},$$

  for $f \in Y$. Show that $\|\cdot\|_p^+$ is a norm and all of these norms are equivalent.
  (b) Show that Y is a Banach space.
  (c) Show that Y is dense in X under any of the norms $\|\cdot\|_p^{\inf}$.
  (d) Show that $\|f\|_p^{\inf} \le \|f\|_p^+$. Is there a constant c such that $\|f\|_p^+ \le c\,\|f\|_p^{\inf}$?

Part 3

Hilbert Spaces and Applications

LECTURE 8

Scalar Products and Hilbert Spaces

Reading: §6.1 and §6.2 of Lax.

Definition 8.1. A scalar product on a linear space X over $\mathbb{R}$ is a real valued function $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R}$ with the following properties

(1) Bilinearity: $x \mapsto \langle x, y\rangle$ and $y \mapsto \langle x, y\rangle$ are linear functions.
(2) Symmetry: $\langle x, y\rangle = \langle y, x\rangle$.
(3) Positivity: $\langle x, x\rangle > 0$ if $x \neq 0$. (Note that $\langle 0, 0\rangle = 0$ by bilinearity.

A (complex) scalar product on a linear space X over $\mathbb{C}$ is a complex valued function $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C}$ with the properties

(1) Sesquilinearity: $x \mapsto \langle x, y\rangle$ is linear and $y \mapsto \langle x, y\rangle$ is skewlinear,

$$\langle x, y + ay'\rangle \;=\; \langle x, y\rangle + a^{\star} \langle x, y'\rangle \,.$$

(2) Skew symmetry: $\langle x, y\rangle = \langle y, x\rangle^{\star}$.
(3) Positivity: $\langle x, x\rangle > 0$ for $x \neq 0$.

Given a (real or complex) scalar product, the associated norm is

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Remark. A complex linear space is also a real linear space, and associated to any complex inner product is a real inner product:

$$(x, y) = \operatorname{Re} \langle x, y\rangle \,.$$

Note that the associated norms are the same, so the metric space structure is the same whether or not we consider the space as real or complex. Note that

$$(\mathrm{i}x, y) \;=\; (x, \mathrm{i}y). \tag{$\star$}$$

and the real and complex inner products are related by

$$\langle x, y\rangle \;=\; (x, y) - \mathrm{i}(\mathrm{i}x, y) \;=\; (x, y) + \mathrm{i}(x, \mathrm{i}y). \tag{$\star\star$}$$

Conversely, given any real inner product on a complex linear space which satisfies ($\star$), ($\star\star$) gives a complex inner product.

We have not shown that the definition $\|x\| = \sqrt{\langle x, x\rangle}$ actually gives a norm. Homogeneity and positivity are clear. To verify subadditivity we need the following important Theorem.

Theorem 8.1 (Cauchy-Schwarz). *A real or complex scalar product satisfies*

$$|\langle x, y\rangle| \;\le\; \|x\| \, \|y\| \,,$$

with equality only if $ax = by$.

REMARK. A corollary is that

$$\|x\| = \max_{\|u\|=1} |\langle x, u\rangle|,$$

from which follows sub-additivity

$$\|x + y\| \le \|x\| + \|y\|.$$

PROOF. It suffices to consider the real case, since given x, y we can always find θ so that $\left\langle e^{i\theta}x, y\right\rangle = e^{i\theta}\langle x, y\rangle$ is real. Also, we may assume $y \neq 0$.

So let $\langle\cdot,\cdot\rangle$ be a real inner product and $t \in \mathbb{R}$. Then

$$\|x + ty\|^2 = \|x\|^2 + 2t\langle x, y\rangle + t^2\|y\|^2.$$

Minimizing the r.h.s. over t we find that,

$$t_{\min} = -\frac{\langle x, y\rangle}{\|y\|^2},$$

and

$$0 \le \|x\|^2 - \frac{\langle x, y\rangle^2}{\|y\|^2}.$$

The Cauchy-Schwarz inequality follows. □

Another important, related result, is the parallelogram identity

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

von Neumann has shown that any norm which satisfies the parallelogram law comes from an inner product.

DEFINITION 8.2. A linear space with a scalar product that is complete in the induced norm is a Hilbert space.

Any scalar product space can be completed in the norm. It follows from the Schwarz inequality that the scalar product is cts. in each of its factors and extends uniquely to the completion, which is thus a Hilbert space.

Examples:

(1) ℓ^2 is a Hilbert space with the inner product

$$\langle a, b\rangle = \sum_j a_j b_j^\star,$$

which is finite by the Hölder inequality.

(2) $C[0, 1]$ is an inner product space with respect to the inner product

$$\langle x, y\rangle = \int_0^1 x(t)y(t)^\star \mathrm{d}t.$$

It is not complete. The completion is known as $L^2[0, 1]$ and can be associated with the set of equivalence classes of Lebesgue square integrable functions.

REMARK. There is no standard as to which factor of the inner product is skew-linear. In the physics literature, it is usually the first factor; in math it is usually the second.

DEFINITION 8.3. Two vectors in an inner product space are orthogonal if

$$\langle x, y \rangle = 0.$$

The orthogonal complement of a set s is

$$S^\perp = \{v \ : \ \langle v, y \rangle = 0 \ \forall y \in Y\}.$$

LEMMA 8.2. *Given any set S in a Hilbert space, $S^\perp$ is a closed subspace.*

PROOF. That $S^\perp$ is a subspace is clear. That it is closed follows, from continuity of the inner product in each factor, since if $v_n \to v$, $v_n \in S^\perp$, then

$$\langle v, y \rangle = \lim_n \langle v_n, y \rangle = 0 \text{ for } y \in S.$$

□

THEOREM 8.3. *Let H be a Hilbert space, Y a closed subspace of H and $Y^\perp$ the orthogonal complement of Y. Then*

(1) Any vector $x \in H$ can be written uniquely as a linear combination

$$x = y + v, \quad y \in Y \text{ and } v \in Y^\perp.$$

(2) $(Y^\perp)^\perp = Y$.

To prove this theorem, we need

LEMMA 8.4. *Given a nonempty closed, convex subset K of a Hilbert space, and a point $x \in H$, there is a unique point y in K that is closer to x than any other point of$K$.*

PROOF. This follows if we show that H is uniformly convex, by the Theorem of Clarkson from lecture 5. Let x, y be unit vectors. It follows from the parallelogram law that

$$\left\| \frac{1}{2}x + y \right\|^2 \ = \ 1 - \frac{1}{4}\|x - y\|^2,$$

so

$$\left\| \frac{1}{2}x + y \right\| \ \le \ 1 - \underbrace{\left(1 - \sqrt{1 - \frac{1}{4}\|x - y\|^2}\right)}_{\epsilon\|x-y\|}.$$

□

PROOF OF THEOREM. According to the Lemma there is a unique point $y \in Y$ closest to a given point $x \in H$. Let $v = x - y$. We claim that $\langle v, y' \rangle = 0$ for any $y' \in Y$. Indeed, we must have

$$\|v\|^2 \le \|v + ty'\|^2 = \|v\|^2 + 2t \operatorname{Re} \langle v, y' \rangle + t^2 \|y'\|^2$$

for any t. In other words the function

$$0 \le 2t \operatorname{Re} \langle v, y' \rangle + t^2 \|y'\|^2 \text{ for all } t,$$

which can occur only if $\operatorname{Re} \langle v, y' \rangle = 0$. Since this holds for all $y' \in Y$ we get $\langle v, y' \rangle = 0$ by complex linearity.

Thus the decomposition $x = y + v$ is possible. Is it unique? Suppose $x = y + v = y' + v'$. Then $y - y' = v - v' \in Y \cap Y^\perp$. But $z \in Y \cap Y^\perp \implies \langle z, z \rangle = 0$ so $z = 0$.

Part (3) is left as a simple exercise. □

LECTURE 9

Riesz-Frechet and Lax-Milgram Theorems

Reading: §6.3 of Lax.

We have already seen that for fixed $y \in H$, a Hilbert space, the map $\ell_y(x) = \langle x, y\rangle$ is a bounded linear functional — boundedness follows from Cauchy-Schwarz. In fancy language $y \mapsto \ell_y$ embeds H into $H^\star$, the dual of H. In fact, since

$$\|\ell_y\| = \sup_x \frac{|\langle x, y\rangle|}{\|x\|} \; = \; \|y\| \, ,$$

again by Cauchy-Schwarz, this map is an isometry onto it's range. In a real Hilbert space, this is a linear map; in a complex Hilbert space, it is skew-linear:

$$\ell_{y+\alpha y'} = \ell_y + \alpha^\star \ell_{y'} .$$

The question now comes up whether we get every linear functional in $H^\star$ this way? The answer turns out to be "yes."

Theorem 9.1 (Riesz-Frechet). *Let $\ell(x)$ be a bounded linear functional on a Hilbert space H. Then there is a unique $y \in H$ such that*

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

Before turning to the proof, let us state several basic facts, whose proof is left as an exercise:

Lemma 9.2.

(1) Let X be a linear space and ℓ a non-zero linear functional on X. Then the null space of ℓ is a linear subspace of co-dimension 1. That is, if$Y = \{y \; : \; \ell(y) = 0\}$ then there exists $x_0 \notin Y$ and any vector $x \in X$ may be written uniquely as

$$x \; = \; \alpha x_0 + y, \quad \alpha \in F \text{ and } y \in Y.$$

(2) If two linear functionals ℓ, m share the same null space, they are constant multiples of each other: $\ell = cm$.

(3) If X is a Banach space and ℓ is bounded, then the null-space of ℓ is closed.

Proof. If $\ell = 0$ then $y = 0$ will do, and this is the unique such point y.

If $\ell \neq 0$, then it has a null space Y, which by the lemma is a closed subspace of co-dimension 1. The orthogonal complement $Y^\perp$ must be one dimensional. Let $\hat{y}$ be a unit vector in $Y^\perp$ — it is unique up to a scalar multiple. Then $m(x) = \langle x, \hat{y}\rangle$ is a linear functional, with null-space Y. Thus $\ell = \alpha m$ and we may take $y = \alpha^\star \hat{y}$.

To see that y is unique, note that if $\langle x, y\rangle = \langle x, y'\rangle$ for all x then $\|y - y'\| = 0$, so $y = y'$. □

In applications, one is often given not a linear functional, but a quadratic form:

DEFINITION 9.1. Let H be a Hilbert space over $\mathbb{R}$. A bi-linear form on H is a function $B : H \times H \to \mathbb{R}$ such that

$$x \mapsto B(x, y) \text{ and } y \mapsto B(x, y)$$

are linear maps. A skew-linear form on a Hilbert space H over $\mathbb{C}$ is a map $B : H \times H \to \mathbb{C}$ such that

$$x \mapsto B(x, y) \text{ is linear, and } y \mapsto B(x, y) \text{ is skew-linear.}$$

A quadratic form refers to a bi-linear form or a skew-linear form depending on whether the field of scalars is $\mathbb{R}$ or $\mathbb{C}$. A quadratic form B on H is bounded if there is a constant $c > 0$ such that

$$|B(x, y)| \;\le\; c \|x\| \|y\| \, ,$$

and is bounded from below if there is a constant $b > 0$ such that

$$|B(y, y)| \;\ge b \|y\|^2 \, .$$

THEOREM 9.3 (Lax-Milgram). *Let H be a Hilbert space, over $\mathbb{R}$ or $\mathbb{C}$, and let B be a bounded quadratic form on H that is bounded from below. Then ever bounded linear functional $\ell \in H^\star$ may be written*

$$\ell(x) = B(x, y), \quad \textit{for unique } y \in H.$$

PROOF. For fixed y, $x \mapsto B(x, y)$ is a bounded linear functional. By Riesz-Frechet there exists $z : H \mapsto H$ such that

$$B(x, y) = \langle x, z(y) \rangle \, .$$

It is easy to see that the map $y \mapsto z(y)$ is linear. Thus the range of z,

$$\operatorname{ran} z \;=\; \{z(y) \;:\; y \in H\} \, ,$$

is a linear subspace of H.

Let us prove that $\operatorname{ran} z$ is a closed subspace. Here we need the fact that B is bounded from below. Indeed,

$$B(y, y) = \langle y, z(y) \rangle \, ,$$

so

$$b \|y\|^2 \le \|y\| \|z(y)\| \quad \Longrightarrow \quad b \|y\| \le \|z(y)\| \, .$$

If y_n is any sequence then

$$\|y_n - y_m\| \le b^{-1} \|z(y_n) - z(y_m)\| \, .$$

Thus $z(y_n) \to z_0 \implies y_n$ Cauchy $\implies y_n \to y_0$, and it is easy to see we must have $z_0 = z(y_0)$. Thus $z_0 \in \operatorname{ran} z$, so $\operatorname{ran} z$ is closed.

Now we show that $\operatorname{ran} z = H$. Since $\operatorname{ran} z$ is closed it suffices to show $\operatorname{ran} z^\perp = \{0\}$. Let $x \perp \operatorname{ran} z$. It follows that

$$B(x, y) \;=\; \langle x, z(y) \rangle \;=\; 0 \quad \forall y \in H.$$

Thus $B(x, x) = 0$ and so $x = 0$ since $\|x\|^2 \le b^{-1} |B(x, x)|$.

Since $\operatorname{ran} z = H$ we see by Riesz-Frechet that any linear functional ℓ may be written $\ell(x) = \langle x, z(y) \rangle = B(x, y)$ for some y. Uniqueness of y follows as above, since if $B(x, y) = B(x, y')$ for all x we conclude that $\|y - y'\| = 0$ since B is bounded from below. $\square$

Application: Radon-Nikodym Theorem

THEOREM 9.4 (Radon-Nikodym). *Let M, Σ be a measurable space on which we have defined two finite non-negative measures μ and ν. If $\mu(A) = 0 \implies \nu(A) = 0$ for all $A \in \Sigma$ then there exists a $\Sigma-$measurable function $h : M \to [0, \infty)$ such that*

$$\nu(A) = \int_A h \mathrm{d}\mu, \quad A \in \Sigma.$$

PROOF. Consider the measure $\mu + \nu$ and the Hilbert space $L^2(\mu + \nu)$,

$$\|f\|^2 = \int_M f^2 \mathrm{d}(\mu + \nu).$$

Define a linear functional

$$\ell(f) \ = \ \int_M f \mathrm{d}\mu.$$

Since

$$|\ell(f)|^2 \ \leq \ \mu(M) \int_M f^2 \mathrm{d}\mu \ \leq \ \mu(M) \|f\|^2 ,$$

it follows that ℓ is bounded (on $L^2(\mu + \nu)$). Thus $\exists g \in L^2(\mu + \nu)$ such that

$$\int_M f \mathrm{d}\mu \ = \ \int_M fg \mathrm{d}(\mu + \nu).$$

Rewrite this as,

$$\int_M f(1 - g) \mathrm{d}\mu \ = \ \int_M f \mathrm{d}\nu. \tag{$\star$}$$

Let $F = \{x \in M \ : \ g(x) \leq 0\}$ and plug $f = \chi_F$ into ($\star$). Then

$$\mu(F) \leq \int_F (1 - g) \mathrm{d}\mu \ = \ \int_F g \mathrm{d}\nu \leq 0.$$

Thus $\mu(F) = 0$. Likewise let $G = \{x \in M \ : \ g(x) > 1\}$ and plug $f = \chi_G$ into ($\star$). If $\mu(G) > 0$ then

$$0 \ > \ \int_G (1 - g) \mathrm{d}\mu \ = \ \int_G g \mathrm{d}\nu > \nu(G),$$

which is a contradiction. Thus $\mu(G) = 0$. Hence, after modifying g on the μ null set $F \cup G$ we may assume that ($\star$) holds with

$$0 < g(x) \leq 1.$$

But then given $E \in \Sigma$, plugging $f = g^{-1}\chi_E$ into ($\star$) we get

$$\nu(E) \ = \ \int_M fg \mathrm{d}\nu \ = \ \int_M \chi_E \frac{1 - g}{g} \mathrm{d}\mu \ = \ \int_E \frac{1 - g}{g} \mathrm{d}\mu.$$

Thus take $h = (1 - g)/g$. □

LECTURE 10

Geometry of a Hilbert space and Gram-Schmidt process

Reading: §6.4 of Lax.

Recall that the linear span of a set S in a linear space X is the collection of finite linear combinations of elements of S:

$$\operatorname{span} S \;=\; \left\{ \sum_{j=1}^{n} \alpha_j x_j \;:\; x_j \in S,\ \alpha_j \in F,\ j = 1, \dots, n,\ n \in \mathbb{N} \right\}.$$

This is also the smallest subspace containing S:

$$\operatorname{span} S \;=\; \cap\{Y \;:\; Y \subset X \text{ is a subspace and } S \subset Y\}.$$

If X is a Banach space, it is natural to look at the smallest closed subspace containing S:

$$\overline{\operatorname{span}}\, S \;=\; \cap\{Y \;:\; Y \subset X \text{ is a closed subspace and } S \subset Y\}.$$

Proposition 10.1. *Let X be a Banach space. Then $\overline{\operatorname{span}}\, S = \overline{\operatorname{span} S}$.*

The proof is left as an exercise.

In a Hilbert space we have a geometric characterization of $\overline{\operatorname{span}}\, S$:

Theorem 10.2. *Let $S \subset H$ be any subset of a Hilbert space H. Then*

$$\overline{\operatorname{span}}\, S = \left(S^{\perp}\right)^{\perp}.$$

That is, $y \in \overline{\operatorname{span}}\, S$ if and only if y is perpendicular to everything that is perpendicular to S:

$$\langle y, z\rangle = 0 \text{ for all } z \text{ such that } \langle x, z\rangle = 0 \text{ for all } x \in S.$$

Proof. Recall that a closed subspace Y satisfies $(Y^{\perp})^{\perp} = Y$. Thus it suffices to show $(\overline{\operatorname{span}}\, S)^{\perp} = S^{\perp}$. Since $S \subset \overline{\operatorname{span}}\, S$ we clearly have $(\overline{\operatorname{span}}\, S)^{\perp}$. On the other hand, if $z \in S^{\perp}$. Thus z is perpendicular to $\operatorname{span} S$ and by continuity of the scalar product $z \perp \overline{\operatorname{span} S} = \overline{\operatorname{span}}\, S$. Thus $S^{\perp} \subset (\overline{\operatorname{span}}\, S)^{\perp}$. □

Definition 10.1. A collection of vectors S in an inner product space H is called <u>orthonormal</u> if

$$\langle x, y\rangle = \begin{cases} 1 & x = y \in S \\ 0 & x \neq y,\ x, y \in S. \end{cases}$$

An orthonormal collection S is called an <u>orthonormal basis</u> if $\overline{\operatorname{span}}\, S = H$.

Lemma 10.3. *Let S be an orthonormal set of vectors in a Hilbert space H. Then the $\overline{\operatorname{span}}\, S$ consists of all vectors of the form*

$$x = \sum_{j=1}^{\infty} \alpha_j x_j, \quad x_j \in S,\ j = 1, \dots, \infty, \tag{$\star$}$$

where the α_j are square summable:

$$\sum_{j=1}^{\infty} |\alpha_j|^2 < \infty.$$

The sum converges in the Hilbert space:

$$\left\| x - \sum_{j=1}^{n} \alpha_j x_j \right\| \to 0,$$

and

$$\|x\|^2 = \sum_{j=1}^{\infty} |\alpha_j|^2.$$

Furthermore, the sum may be written

$$x = \sum_{y \in S} \langle x, y \rangle \, y.$$

In particular, $\langle x, y \rangle \neq 0$ for only countably many elements $y \in S$.

Remark. Most orthonormal sets encountered in practice are countable, so we would tend to write $S = \{x_1, \ldots, \}$ and

$$x = \sum_{j=1}^{\infty} \langle x, x_j \rangle \, x_j.$$

However, the lemma holds even for uncountable orthonormal sets.

Proof. It is clear that all vectors of the form ($\star$) are in $\overline{\operatorname{span} S} = \overline{\operatorname{span}}\, S$. Furthermore vectors of this form make up a subspace, which is easily seen to be closed. (Exercise: show that this subspace is closed. This rests on the fact that a subset of a complete metric space is closed iff it is sequentially complete.) By definition $\overline{\operatorname{span}}\, S$ is contained in this subspace. Thus the two subspaces are equal.

The remaining formulae are easy consequences of the form ($\star$). □

Theorem 10.4. *Every Hilbert space contains an orthonormal basis.*

Proof. We use Zorn's Lemma. Consider the collection of all orthonormal sets, with $S \leq T$ iff $S \subset T$. This collection is non-empty since any unit vector makes up a one element orthonormal set.

A totally ordered collection has an upper bound — the union of all sets in the collection. Thus there is a maximal orthonormal set. Call it $S_{\max}$.

Suppose $\overline{\operatorname{span}}\, S_{\max} \subsetneq X$. Then, $\overline{\operatorname{span}}\, S_{\max}^{\perp}$ is a non-trivial closed subspace. Let $y \in \overline{\operatorname{span}}\, S_{\max}^{\perp}$ be a unit vector. So $S_{\max} \cup \{y\}$ is an orthonormal set contradicting the fact that $S_{\max}$ is maximal. □

Corollary 10.5 (Bessel's inequality). *Let S be any orthonormal set in a Hilbert space H (not necessarily a basis), then*

$$\sum_{y \in S} |\langle x, y \rangle|^2 \ \leq \ \|x\|^2 \ \text{for all } x \in H.$$

Equality holds for every x if and only if S is a basis.

If the Hilbert space H is separable — it contains a countable dense set — then any orthonormal basis is countable. In this case we can avoid Zorn's Lemma. The foundation of this is the Gram-Schmidt process.

THEOREM 10.6 (Gram-Schmidt Process). *Let y_j be a sequence of vectors in a Hilbert space. Then there is an orthonormal sequence x_j such that*

$$\operatorname{span}\{y_1, \ldots, y_n\} \subset \operatorname{span}\{x_1, \ldots, x_n\}.$$

PROOF. The proof is constructive. We may assume, without loss, that $\{y_1, \ldots, y_n\}$ is linearly independent for each n. (Otherwise throw out vectors y_n until this is the case.) Then $y_1 \neq 0$ so let

$$x_1 = \frac{y_1}{\|y_1\|}.$$

Clearly $\operatorname{span}\{x_1\} = \operatorname{span}\{y_1\}$.

Now, suppose we are given $x_1, \ldots, x_{n-1}$ such that

$$\operatorname{span}\{x_1, \ldots, x_{n-1}\} = \operatorname{span}\{y_1, \ldots, y_{n-1}\}.$$

Let

$$x_n = \frac{y_n - \sum_{j=1}^{n-1} \langle y_n, x_{n-1}\rangle x_{n-1}}{\left\| y_n - \sum_{j=1}^{n-1} \langle y_n, x_{n-1}\rangle x_{n-1} \right\|}.$$

This is OK since $y_n \neq \sum_{j=1}^{n} \langle y_n, x_{n-1}\rangle \in \operatorname{span}\{y_1, \ldots, y_{n-1}\}$. Clearly

$$\operatorname{span}\{x_1, \ldots, x_n\} = \operatorname{span}\{y_1, \ldots, y_n\}.$$

By induction, the result follows. □

COROLLARY 10.7. *Let H be a separable Hilbert space. Then H has a countable orthonormal basis.*

Finally, let us discuss the isometries of Hilbert spaces.

THEOREM 10.8. *Let H and H' be Hilbert spaces. Given an orthonormal basis S for H, an orthonormal set $S' \subset H'$ and a one-to-one onto map $f : S \to S'$, define a linear map $H \to H'$ via*

$$\sum_{y \in S} \alpha_y y \xmapsto{T_f} \sum_{y \in S} \alpha_y f(y).$$

Then T is a linear isometry onto $\overline{\operatorname{span}}\, S' \subset H'$. Furthermore, any isometry of H with a subspace of H' is of this form.

COROLLARY 10.9. *Two Hilbert spaces are isomorphic iff their orthonormal bases have equal cardinality. In particular, every Hilbert space is isomorphic with $\ell^2(S)$ for some set S. Any separable, infinite dimensional Hilbert space is isomorphic to ℓ^2.*

REMARK. For an arbitrary set S, $\ell^2(S)$ is defined to be the set of functions $f : S \to \mathbb{R}$ or $\mathbb{C}$ such that

$$\sum_{y \in S} |f(y)|^2 < \infty.$$

Note that $f \in \ell^2(S) \implies \{y : f(y) \neq 0\}$ is countable.

The proof of these results is left as an exercise.

Part 4

Locally Convex Spaces

LECTURE 11

Locally Convex Spaces and Spaces of Test Functions

Reading: §13.-13.2 and §B.1 of Lax.

A Banach space is one example of a topological vector space (TVS), which is a linear space X together with a topology on X such that the basic operations of addition and scalar multiplication are continuous functions.

DEFINITION 11.1. A topological vector space is a linear space X with a Hausdorff topology such that

(1) $(x, y) \mapsto x + y$ is a continuous map from $X \times X$ (with the product topology) into X.
(2) $(k, x) \mapsto kx$ is a continuous map from $F \times X$ (with the product topolgoy, $F = \mathbb{R}$ or $\mathbb{C}$) into X.

REMARK. Recall that a Hausdorff space is one in which points may be separated by open sets: given $x, y \in X$, $x \neq y$ there are disjoint open sets U, V, $U \cap V = \emptyset$ such that $x \in U$ and $y \in V$.

THEOREM 11.1. *Let X be a TVS and let $U \subset X$ be open. Then*

(1) For any $x \in X$, $U - x = \{y \ : y + x \in U\}$ is open.
(2) For any scalar $k \neq 0$, $kU = \{y \ : \ k^{-1}y \in U\}$ is open
(3) Every point of U is interior: given $x \in U$ and $y \in X$ there is $\epsilon > 0$ such that for any scalar t with $|t| < \epsilon$ we have $x + ty \in U$.

PROOF. The set $U - x$ is the inverse image of U under the map $y \mapsto y + x$. Thus (1) follows from continuity of the map $y \mapsto y + x$ which follows from joint continuity of $(y, x) \mapsto y + x$. (Why?)

Likewise (2) follows from continuity of $y \mapsto k^{-1}y$.

Since $U - x$ is open it suffices to suppose $x = 0 \in U$. For fixed $y \in X$ the map $t \mapsto ty$ is continuous. (Why?) Thus $\{t \ : \ ty \in U\}$ is open. Since this set contains $t = 0$ it must contain an interval $(-\epsilon, \epsilon)$ (or an open ball at the origin if the field of scalars is $\mathbb{C}$). □

The class of TVSs is rather large. However, almost all of the TVSs important to analysis have the following property:

DEFINITION 11.2. A locally convex space (LCS) is a TVS X such that every open set containing the origin contains an open convex set containing the origin. That is, there is a basis at the origin consisting of open convex sets.

Given a LCS, X, we define the dual X' to be the set of all continuous linear functional on X. A LCS space shares the property of separation of points by linear functionals:

THEOREM 11.2. *Let X be a LCS and let $y \neq y'$ be points of X. There is a linear functional $\ell \in X'$ such that*

$$\ell(y) \neq \ell(y').$$

PROOF. Of course, we use the Hahn-Banach theorem. Specifically the hyperplane separation Theorem 2.1.

First, it suffices to suppose the field of scalars is $\mathbb{R}$, for if we construct a suitable real linear functional ℓ_r on a complex LCS we can complexify it

$$\ell(x) \ = \ \ell_r(x) - \mathrm{i}\ell_r(\mathrm{i}x).$$

Now, without loss we suppose that $y' = 0$. Since the topology on X is Hausdorff, there is an open set $U \ni y'$ with $y \notin U$. Since X is locally convex, we may suppose U to be convex. Replacing U with $U \cap (-U)$ we may assume U is symmetric about 0, so $x \in U \implies -x \in U$. Since all points of U are interior, Theorem 2.1 asserts the existence of a linear functional ℓ with $1 = \ell(y)$ and $\ell(x) < 1$ for $x \in U$. In fact, the proof shows that

$$\ell(x) \le p_U(x) \quad \forall x \in X,$$

where p_U is the gauge function of U,

$$p_U(x) = \inf\{t > 0 \ : \ t^{-1}x \in U\}.$$

We need to show that ℓ is continuous. It suffices to show $\ell^{-1}(a,b)$ is open for any $a < b \in \mathbb{R}$. Let $t \in (a,b)$. Let x_0 be any point with $\ell(x_0) = t$. Then, $\ell^{-1}(a-t, b-t) = \ell^{-1}(a,b) - x_0 \ni 0$. (Why?) Thus it suffices to suppose $a < 0 < b$ and show that $\ell^{-1}(a,b)$ contains an open neighborhood at 0. Let $t = \min\{-a, b\}$. The given $x \in U$,

$$\ell(tx) \le p_U(tx) = tp_U(x) < t \quad \text{and} \quad \ell(-tx) \le p_U(-tx) = tp_U(-x) < t.$$

Thus $tU \subset \ell^{-1}(-t,t) \subset (a,b)$. □

Converse to this construction is the following idea

THEOREM 11.3. *Let X be linear space and let L be any collection of linear functionals on X that separates points: for any $y, y' \in X$ there is $\ell \in L$ such that $\ell(x) \neq \ell(x')$. Endow X with the <u>weakest</u> topology such that all elements of L are continuous. Then X is a LCS, and the dual of X is*

$$X' = \operatorname{span} L = \{\textit{finite linear combinations of elements of } L\}.$$

REMARK. Recall that a topology on X is a collection $\mathcal{U}$ of subsets of X that includes X and $\emptyset$, and is closed under unions and finite intersections. The intersection of a family of topologies is also a topology. Thus the weakest topology with property A is the intersection of all topologies with property A.

PROOF. Exercise. □

What are some examples of LCSs? First off, any Banach space is locally convex, since the open balls at the origin are a basis of convex sets. But not every LCS has a norm which is compatible with the topology. By far the most important examples, though are so-called spaces of text functions and their duals, the so-called spaces of distributions.

LECTURE 12

Generation of a LCS by seminorms and Fréchet Spaces

Reading: §5.2 of Reed and Simon (or just these notes)

Test functions

The theory of distributions — due to Laurent Schwarz — is based on introducing a LCS of "test functions" X and it's dual X', a space of "distributions." The test functions are "nice:" we can operate on them arbitrarily with all the various operators of analysis – differentiation, integration etc. Using integration, we embed $X \hookrightarrow X'$ via a map $\phi \mapsto \ell_\phi$:

$$\ell_\phi(\psi) \;=\; \int_{\mathbb{R}^d} \phi(x)\psi(x)\mathrm{d}x.$$

Thus we think of a (test) function both as a map and as an "averaging" procedure. A key identity is integration by parts

$$\ell_\phi(\partial_i\psi) \;=\; -\ell_{\partial_i\psi}(\phi).$$

This suggests that we define

$$\partial_i\ell(\psi) \;=\; -\ell(\psi)$$

for any distribution $\ell \in X'$. One typically uses a function notation for a distribution, writing

$$\ell(\psi) \;=\; \int T(x)\psi(x),$$

even if the "function" T doesn't exist.

A common, and useful, space of test functions is

$$C_c^\infty(\mathbb{R}^d) \;=\; \{C^\infty \text{ functions on } \mathbb{R}^d \text{ with compact support.}\}.$$

We wish to topologize this set so that a sequence u_k converges u if the supports of all u_k are contained in some fixed compact $K \subset \mathbb{R}^d$ and if for any choice of mult-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ we have

$$D^\alpha u_k = \partial_1^{\alpha_1}\cdots\partial_d^{\alpha_d} u_k \;\to\; D^\alpha u \text{ uniformly in } K.$$

What is really going here is this. Define for each $n \geq 0$ and each multi-index α a <u>semi-norm</u> of $C_c^\infty(\mathbb{R}^d)$,

$$p_{n,\alpha}(u) \;=\; \sup_{|x|\leq 2^n} |D^\alpha u(x)|\,.$$

(Recall that a <u>semi-norm</u> on a linear space X is a map $p : X \to [0,\infty)$ which is positive homogeneous ($p(ax) = |a|p(x)$) and sub-additive ($p(x+y) \leq p(x) + p(y)$). It is allowed that $p(x) = 0$ for $x \neq 0$.) It may happen that $\|u\|_{n,\alpha}$ vanishes even if $u \neq 0$, however the collection separates points (Why?):

Definition 12.1. A collection of semi-norms $\mathcal{S}$ <u>separates points</u> if

$$p(u) = 0 \;\forall p \in \mathcal{C} \quad\Longrightarrow\quad u = 0.$$

Now endow $X = C_0^\infty(\mathbb{R}^d)$ with the smallest topology such that X is a TVS and each of the semi-norms $p_{n,\alpha}$ is continuous.

CLAIM. *$C_c^\infty(\mathbb{R}^d)$ with this topology is a LCS*

`Warning: this is not the standard topology on` $C_c^\infty(\mathbb{R}^d)$. `See below.`

Generation of an LCS by semi-norms

The claim follows from the following general result.

DEFINITION 12.2. Let X be a linear space and $\mathcal{S}$ a family of functions $f : X \to M$, M a topological space (usually $M = \mathbb{R}$ or $\mathbb{C}$). The TVS topology generated by $\mathcal{S}$ is the weakest topology on X such that X is a TVS and all the functions in $\mathcal{S}$ are continuous.

THEOREM 12.1. *Given a linear space X and a collection of semi-norms $\mathcal{S}$ that separates points, the TVS topology generated by $\mathcal{S}$ makes X is locally convex.*

Conversely, given a LCS X and $\mathcal{C}$ a neighborhood base at the origin consisting of convex, symmetric sets, the LCS topology on X is the TVS topology generated by $\mathcal{S} = \{p_U \ : \ U \in \mathcal{C}\}$ are continuous.

PROOF. First given an LCS space X let us show that the gauge function p_U of a convex, symmetric neighborhood of the origin U is continuous. To begin, note that

$$p_U^{-1}[0, b) \ = \ \{x \in X \ : \ x \in bU\} = bU$$

is open for each b. Next consider the sets $p_U^{-1}(b, \infty)$. Let x be in this set and let $\alpha = p_U(x)$. Consider the open neighborhood $V = x + (\alpha - b)U$ then for $y = x + (\alpha - b)y' \in V$ we have

$$p_U(y) \geq p_U(x) - (\alpha - b)p_U(y') > b.$$

So $V \subset p_U^{-1}(b, \infty)$ and thus the set is open. Continuity of p_U follows since the sets $[0, b)$, (b, ∞) as b ranges over $(0, \infty)$ generate the topology on $[0, \infty)$.

Now, any topology under which every $p \in \mathcal{S}$ is continuous certainly contains the collection

$$\mathcal{C} = \{p^{-1}[0, b) \ : \ b \in (0, \infty)\},$$

consisting of convex, symmetric sets with the origin as an interior point. Consider the smallest TVS topology containing this collection. It is easily seen to be locally convex. (We need the fact that $\mathcal{S}$ separates points to get the Hausdorff property.) Since $p = p_U$ for $U = p^{-1}[0, 1)$ we see from the above argument that all $p \in \mathcal{S}$ are continuous in this topology. Thus this is the TVS topology generated by $\mathcal{S}$.

Conversely, let $\mathcal{T}$ denote a given LCS topology on X. Thus $\mathcal{T}$ is certainly a topology under which X is a TVS and all elements of $\mathcal{S} = \{p_U : U$ a convex, symmetric neighborhood of $0\}$ are continuous. To prove it is the weakest such, we must show that any such topology contains $\mathcal{T}$. Any $U \in \mathcal{C}$ may be written as $U = p_U^{-1}([0, 1))$. Thus any topology under which X is a TVS and all p_U are continuous certainly contains $\mathcal{C}$ and all its translates, and thus $\mathcal{T}$. $\square$

Metrizable LCSs

Above we claimed that the topology on C_0^∞ could be given in terms of uniform convergence of sequences of functions. However, in a general LCS sequential convergence may not specify the topology — a set may fail to be closed even it contains the limits of all convergent sequences of its elements — because there may not be a countable neighborhood base at the origin. (Don't worry too much about this.) However, if the origin has a countable neighborhood base then it turns out that the LCS is actually metrizable, so in particular sequential convergence specifies the topology.

THEOREM 12.2. *Let X be an LCS. The following are equivalent*

(1) X is metrizable

(2) X has a countable neighborhood basis at the origin $\mathcal{C}$ consisting of convex, symmetric sets

(3) the topology on X is generated by a countable family of semi-norms.

PROOF. The equivalence of (2) and (3) is established by associating to convex, symmetric neighborhoods of the origin the corresponding gauge function and vice versa. The details are left as an exercise.

To show (1) $\implies$ (2), suppose X is metrizable. Then X has a countable neighborhood basis (this is a property of metric spaces), and in particular a countable neighborhood basis at the origin. Since X is a LCS we may find a convex, symmetric open set contained in each of the basis sets, thus obtaining a convex, symmetric, countable neighborhood basis at the origin.

To show (2) $\implies$ (1), suppose X has a countable neighborhood basis as indicated, and let $\mathcal{T}$ denote it's topology. Since $\mathcal{C}$ is countable we may assume, without loss, that it is a decreasing sequence $\mathcal{C} = \{U_1 \supset U_2 \supset \cdots\}$. (Order the elements of $\mathcal{C}$ and take finite intersections $U_k \mapsto U_1 \cap \cdots \cap U_k$.) Let $p_j(x) = p_{U_j}(x)$, so $p_k(x) \geq p_j(x)$ if $k \geq j$. Define a metric

$$d(x, y) \ = \ \sum_{j=1}^{\infty} 2^{-j} \frac{p_j(x-y)}{1 + p_j(x-y)},$$

and the metric topology $\mathcal{T}_d$. (Recognize this? Why is this a metric? Note that the collection $\{p_U \ : \ U \in \mathcal{C}\}$ separates points since $\mathcal{C}$ is a basis.)

Clearly,

$$\{x \ : \ d(x, 0) < 2^{-j-1}\} \subset U_j.$$

Thus $\mathcal{T} \subset \mathcal{T}_d$ (since any $\mathcal{T}$ open set contains a translate of some d-ball centered at each of its points). On the other hand, if $x \in tU_k$

$$d(x, 0) < \sum_{j=1}^{k} 2^{-j} \frac{t}{1+t} + \sum_{j=k+1}^{\infty} 2^{-j} \leq \frac{1}{2} \frac{t}{1+t} + 2^{-k-1}.$$

Thus

$$2^{-k} U_k \subset \{x \ : \ d(x, 0) < 2^{-k}\},$$

which shows that $\mathcal{T}_d \subset \mathcal{T}$. Thus $\mathcal{T} = \mathcal{T}_d$ and X is metrizable. □

DEFINITION 12.3. A Fréchet space is a complete, metrizable, locally convex linear space.

REMARK. Recall that a metric space is complete if every Cauchy sequence converges.

LECTURE 13

The dual of an LCS

It turns out that $C_c^\infty(\mathbb{R}^d)$, although a metric space with the metric suggested above, is not complete. The completion is the space of C_0^∞ functions that together with all their derivatives vanish at ∞, which could be topologized with the seminorms

$$p_\alpha(f) = \sup_x |D^\alpha f(x)| .$$

The correct way to think of $C_c^\infty(\mathbb{R}^d)$ is as an "inductive limit," which is as the union $\cup_n C^\infty([-2^n, 2^n])$. There is an inductive limit topology on this space, making it a complete, but non-metrizable, LCS. Each of the spaces $C^\infty([-2^n, 2^n])$ is a Fréchet space, but clearly the inductive limit topology is different that the (metrizable) LCS topology generated above. It is still true that a sequence converges according to the criteria given above, but this doesn't give the whole picture as the space isn't separable! (See Reed and Simon chapter V.)

A remedy for this, that is often sufficient, is to work with the following somewhat larger space

Definition 13.1. The Schwarz space $\mathcal{S}(\mathbb{R}^d)$ consists of every C^∞ function f on $\mathbb{R}^d$ such that

$$p_{\alpha,\beta}(f) = \sup_{x\in\mathbb{R}^d} |x^\alpha D^\beta f(x)| < \infty,$$

for every pair of multi-indices $\alpha, \beta \in \mathbb{N}^d$.

Remark.
- $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$.
- Note that $\mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ for every p.
- Likewise $f \in \mathcal{S}(\mathbb{R}^d) \implies D^\alpha f \in L^p(\mathbb{R}^d)$ for every p.

Corollary 13.1. *The space $\mathcal{S}(\mathbb{R}^d)$ with the topology generated by the seminorms $p_{\alpha,\beta}$ is a Fréchet space.*

The dual space of a LCS

It is useful to use the inner product notation to denote the pairing between elements of X and linear functionals

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

(Note that this inner product is linear in both factors even if we are dealing with complex spaces.)

Given a linear space X and a linear space of linear functionals L on X that separates points we have seen that there is a LCS topology on X such that L is the dual of X. This topology is called the L-weak topology on X and is denoted $\sigma(X, L)$.

On the other hand, given an LCS, we can think of X as a collection of linear functionals on $X^\star$, associating to $x \in X$ the map

$$\ell \mapsto \langle x, \ell \rangle \,.$$

The X-weak topology on $X^\star$, $\sigma(X^\star, X)$, is also called the <u>weak* toplogy</u>. It is generated by the family of seminorms

$$p_x(\ell) \;=\; |\langle x, \ell\rangle| \,.$$

THEOREM 13.2. *If X is an LCS then $(X^\star, \sigma(X^\star, X))^\star = X$.*

REMARK. Recall that $\sigma(X, X^\star)$ is the given LCS topology on X so we also have $(X\sigma(X, X^\star))^\star = X^\star$. Thus for any LCS $(X^\star)^\star = X$, provided we topologize $X^\star$ with the weak* toplogy. If X is a Banach space we also have a norm topology on $X^\star$, which is substantially stronger than the weak* topology and with respect to which this identity may not hold. For instance,

(1) As Banach spaces $c_0^\star = \ell_1$ and $\ell_1^\star = \ell_\infty$ and $\ell_\infty^\star$, which includes Banach limits, is strictly larger than ℓ_1.
(2) As LCS spaces $c_0^\star = \ell_1$ and $(\ell_1, \sigma(\ell_1, c_0)^\star = c_0$, etc.

The moral of the story is <u>topology matters</u>.

The following theorem is useful for determining if a linear functional is continuous.

THEOREM 13.3. *Let X be a LCS generated by a family of semi-norms $\mathcal{S}$. Then a linear functional $\ell \in X'$ if and only if there is a constant $C > 0$ and a finite collection $p_1, \ldots, p_n \in \mathcal{S}$ such that*

$$|\ell(x)| \;\le\; C \sum_{j=1}^n p_j(x) \quad \forall x \in X.$$

PROOF. ($\Rightarrow$) If ℓ is continuous then $U = \ell^{-1}(-1, 1)$ an open, convex, symmetric neighborhood of the origin in X. By virtue of the fact that $\mathcal{S}$ generates the topology on X, since U is open we have

$$\bigcap_{j=1}^n \{x \;:\; p_j(x) < \varepsilon\} \;\subset\; U$$

for some finite collection $p_1, \ldots, p_n$. Thus,

$$V = \{x \;:\; \sum_{j=1}^n p_j(x) < \varepsilon\} \;\subset\; U.$$

Now, given x let $t = \frac{2}{\varepsilon} \sum_{j=1}^n p_j(x)$. Then

$$\sum_{j=1}^n p_j(t^{-1}x) = \frac{\varepsilon}{2},$$

so $t^{-1}x \in V \subset U$. Thus

$$\left|\ell(t^{-1}x)\right| \;<\; 1 \;=\; \frac{\sum_{j=1}^n p_j(x)}{\sum_{j=1}^n p_j(x)} \;=\; \frac{2}{\varepsilon} \sum_{j=1}^n p_j(t^{-1}x).$$

Multiplying through by t we get the desired bound with $C = 2/\varepsilon$.

($\Leftarrow$) Since ℓ is linear, it suffices to show that ℓ is continuous at 0. That is we must show that $\ell^{-1}(-\varepsilon, \varepsilon)$ contains an open set containing the origin for each $\varepsilon > 0$. But clearly

$$\{x \ : \ C \sum_{j=1}^{n} p_j(x) < \varepsilon\} \ \subset \ \ell^{-1}(-\varepsilon, \varepsilon).$$

$\square$

LECTURE 14

Spaces of distributions

Tempered Distributions

The dual of $\mathcal{S}(\mathbb{R}^d)$ denoted $\mathcal{S}^\star(\mathbb{R}^d)$ is the space of tempered distributions. Here are some examples:

(1) $\mathcal{S}(\mathbb{R}^d) \subset \mathcal{S}^\star(\mathbb{R}^d)$ where we associate to a function $\phi \in \mathcal{S}(\mathbb{R}^d)$ the distribution

$$\psi \ \mapsto \ \langle \psi, \phi \rangle \ = \ \int_{\mathbb{R}^d} \psi(x)\phi(x)\mathrm{d}x.$$

(2) More generally, a function $F \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ that is polynomially bounded in the sense that $p(x)^{-1}F(x) \in L^1(\mathbb{R}^d)$ for some positive polynomial $p > 0$ may be considered as a tempered distribution

$$\psi \ \mapsto \ \langle psi, F \rangle \ = \ \int_{\mathbb{R}^d} \psi(x)F(x)\mathrm{d}x.$$

(3) Similarly, any polynomially bounded Borel measure μ, with

$$\int p(x)^{-1}\mathrm{d}|\mu|(x) \ < \ \infty.$$

is a tempered distribution:

$$\langle \psi, \mu \rangle \ = \ \int_{\mathbb{R}^d} \psi(x)\mathrm{d}\mu(x).$$

To go further we need the following generalization of Theorem 13.3 from the last lecture:

THEOREM 14.1. *Let X, Y be a LCSs generated by a families of semi-norms $\mathcal{S}, \mathcal{T}$ respectively. Then a linear map $T : X \to Y$ is continuous if and only if for any semi-norm $q \in \mathcal{S}$ there is a constant $C > 0$ and a finite collection $p_1, \ldots, p_n \in \mathcal{S}$ such that*

$$q(Tx) \ \leq \ C\sum_{j=1}^n p_j(x) \quad \forall x \in X.$$

COROLLARY 14.2. *For each $j = 1, \ldots, d$, differentiation ∂_j is a continuous map from $\mathcal{S} \to \mathcal{S}$.*

Now we define $\partial_j : \mathcal{S}^\star \to \mathcal{S}^\star$. Note that

$$\langle \partial_j \psi, F \rangle \ = \ -\langle \psi, \partial_j F \rangle \, ,$$

whenever F is a C^1 function of polynomial growth. Thus define for arbitrary $\ell \in \mathcal{S}^\star$:

$$\langle \psi, \partial_j \ell \rangle \ = \ -\langle \partial_j, \ell \rangle \, .$$

PROPOSITION 14.3. *So defined, $\partial_j : \mathcal{S}^\star \to \mathcal{S}^\star$ is a continuous map.*

PROOF. Exercise ∎

Thus we have the following generalization of the above examples:

- Let $\alpha \in \mathcal{N}^d$ be a multi-index and let F be a polynomial L^1 bounded function. Then $D^\alpha F$ is a tempered distribution:

$$\langle \psi, D^\alpha F\rangle \;=\; (-1)^\alpha \langle D^\alpha \psi, F\rangle \;==\; (-1)^\alpha \int_{\mathbb{R}^d} D^\alpha \psi(x) F(x) \mathrm{d}x.$$

THEOREM 14.4 (Structure Theorem for Tempered Distributions). *Let $\ell \in \mathcal{S}(\mathbb{R}^d)$ be a tempered distribution. Then there is a polynomially bounded* <u>*continuous*</u> *function g and a multi-index $\alpha \in \mathbb{N}^d$ such that $\ell = D^\alpha g$.*

For the proof see Reed and Simon, Ch. V.

For example, we now understand in a precise sense the identity

$$\delta(x) \;=\; \frac{\mathrm{d}^2}{\mathrm{d}x^2}|x|.$$

More generally, in $d = 2$, we have the identity

$$\delta(x) \;=\; \frac{1}{2\pi}\Delta \ln |x|.$$

This can be verified as follows. Note that $\Delta \ln |x| = 0$ if $x \neq 0$. Thus

$$\langle \phi, \Delta \ln |x|\rangle = 0 \quad \text{ if } \phi(x) = 0 \text{ for } |x| < \epsilon.$$

Let h be a compactly supported function that is 1 in the neighborhood of the origin, then $\langle \phi - h\phi, \Delta \ln |x|\rangle = 0$. Thus it suffices to suppose $\phi \in C_c^\infty(\mathbb{R}^d)$. For such ϕ

$$\langle \phi, \Delta \ln |x|\rangle \;=\; \int_{|x|<M} \Delta\phi(x) \ln |x| \mathrm{d}x \;=\; \lim_{\varepsilon \downarrow 0} \int_{\varepsilon<|x|<M} \Delta\phi(x) \ln |x| \mathrm{d}x,$$

since $\ln |x|$ is locally integrable. Now integrate by parts:

$$\int_{\varepsilon<|x|<M} \Delta\phi(x) \ln |x| \mathrm{d}x \;=\; -\int_{\varepsilon<|x|<M} \frac{x}{|x|^2}\cdot \nabla\phi \mathrm{d}x - \ln\varepsilon \int_{|x|=\varepsilon} \frac{x}{|x|}\cdot\nabla\phi(x)\mathrm{d}\sigma(x),$$

where in the first integral $\frac{x}{|x|^2}$ is $\nabla \ln |x|$ and in the second integral $-\frac{x}{|x|}$ is an outward facing normal on $\{|x| = \varepsilon\}$ and $\mathrm{d}\sigma(x)$ is the length measure on the circle. Continuing, we find that

$$\int_{\varepsilon<|x|<M} \Delta\phi(x) \ln |x| \mathrm{d}x \;==\; \frac{1}{\varepsilon}\int_{|x|=\varepsilon} \phi(x)\mathrm{d}\sigma(x) - O(\varepsilon \ln \varepsilon) \;\to\; \pi\phi(0).$$

REMARK. In a similar fashion, in dimension $d \geq 3$,

$$\delta(x) \;=\; -\Delta c_d |x|^{2-d},$$

with $c_d^{-1} = (d-2)\times$ area of the unit sphere $\{|x| = 1 \;:\; x \in \mathbb{R}^d\}$.

Other spaces of distributions

Consider the scale of spaces

$$C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset C_0^\infty(\mathbb{R}^d) \subset C^\infty(\mathbb{R}^d).$$

Each has a natural LCS topology such that it is a complete space. The middle two are Fréchet spaces and we have discussed their topologies already.

On $C_c^\infty(\mathbb{R}^d)$ we put the inductive limit topology as follows. For each n, let $\Omega_n = (-2^n, 2^n)^d$. The space $X_n = C_0^\infty(\Omega_n)$ is a Fréchet space generated by the seminorms

$$p_{n,\alpha}(u) \; = \; \sup_{x \in \omega_n} |D^\alpha u(x)| \, . \tag{$\star$}$$

Note that $X_n \subset X_{n+1}$, the embedding is continuous, and X_n is a closed subspace of X_{n+1}. Put of $C_c^\infty(\mathbb{R}^d) = X = \cup_n X_n$ the LCS topology which has the following convex neighborhood base at the origin

$$\mathcal{C} \; = \; \{U \subset X \; : \; U \text{ is convex, symmetric, all points of } U \text{ are interior, and } U \cap X_n \text{ is open for each } n\}.$$

THEOREM 14.5. *With this topology $C_c^\infty(\mathbb{R}^d)$ is a LCS and is complete. The topology does not depend on the choice of sets Ω_n and a sequence $u_n \in C_c^\infty(\mathbb{R}^d)$ converges if and only if there is N such that $u_n \in X_N$ for all n and u_n converges in X_N.*

For the proof see Reed and Simon.

A <u>distribution</u> is an element of $\mathcal{D}^\star(\mathbb{R}^d) = (C_c^\infty(\mathbb{R}^d))^\star$. Since any tempered distribution T acts on $C_c^\infty(\mathbb{R}^d)$ and furthermore if $\langle \phi, T\rangle = 0$ for all ϕ in C_c^∞ then $T = 0$ (note that $C_c^\infty(\mathbb{R}^d)$ is dense in $\mathcal{S}$), we have the embedding $\mathcal{S}^\star \subset \mathcal{D}$. Distributions in $\mathcal{D}$ need not be bounded at ∞ – any locally integrable function, like $\mathrm{e}^{\mathrm{e}^{\mathrm{e}^{|x|^2}}}$ is a distribution. Also, the structure theorem doesn't hold in this context, however we have

THEOREM 14.6. *Let T be a distribution in $\mathcal{D}(\mathbb{R}^d)$ then there is a sequence f_α of continuous functions, indexed by multi-indices, such that*

$$T = \sum_\alpha D^\alpha f_\alpha,$$

where for any $u \in C_c^\infty(\mathbb{R}^d)$ only finitely many terms contribute to the sum

$$\langle u, T\rangle \; = \; \sum_\alpha \langle u, D^\alpha f_\alpha \rangle \, .$$

Roughly speaking, the <u>order</u> of the distribution can become unbounded at ∞.

Likewise we may put a topology on $C^\infty(\mathbb{R}^d)$ given by uniform convergence on compact subsets. That is generated by the family of semi-norms ($\star$), where now u need not have compact support. This makes $\mathcal{E}(\mathbb{R}^d) = C^\infty(\mathbb{R}^d)$ into a Fréchet space. Let us denote it's dual by $\mathcal{E}^\star$. We have the inclusions

$$\mathcal{E}^\star \subset \mathcal{D}_0^\star \subset \mathcal{S}^\star \subset \mathcal{D}^\star,$$

where $\mathcal{D}_0^\star$ is the dual of $C_0^\infty(\mathbb{R}^d)$. Distributions in $\mathcal{E}^\star$ have compact support, where

DEFINITION 14.1. A distribution vanishes on an open set $U \subset \mathbb{R}^d$ if $\langle \phi, T\rangle = 0$ whenever ϕ has compact support in U. The support of a distribution is the smallest closed set F such that T vanishes on $\mathbb{R}^d \setminus F$.

Distributions in $\mathcal{D}_0$ are "bounded," in the sense that they may be written as finite sums of derivatives of finite measures. Distributions in $\mathcal{S}$ are "unbounded at ∞," but only polynomially so.

LECTURE 15

Applications: solving some PDE's

The Poisson equation

The electric potential ϕ produced by a charge distribution ρ is known to satisfy the Poisson equation:

$$-\Delta\phi(x) \;=\; \rho(x), \quad x \in \mathbb{R}^3$$

with the permitivity of space set $= 1$. Let us use the fact that

$$-\Delta \frac{1}{4\pi}|x|^{-1} = \delta(x)$$

to solve this equation. We need the following notion.

Definition 15.1. Given two functions ϕ, ψ the convolution of ϕ and ψ is

$$\phi * \psi(x) \;=\; \int_{\mathbb{R}^d} \phi(x-y)\psi(y)\mathrm{d}y,$$

whenever the integral is defined.

Proposition 15.1. *The convolution product is abelian, $\phi*\psi = \psi*\phi$ and is a continuous map of $L^1 \times L^1 \to L^1$, $\mathcal{S}\times\mathcal{S} \to \mathcal{S}$, and $\mathcal{D}\times\mathcal{D}\to\mathcal{D}$.*

Proof. Exercise. □

Theorem 15.2 (Young's Inequality). *Convolution is a continuous map from $L^p \times L^q \to L^r$ with $1/p + 1/q = 1 + 1/r$, $1 \le p, q, r \le \infty$.*

Proof. Let $f \in L^{r'}$, $1/r' + 1/r = 1$. Then

$$\left|\int f(x)\phi*\psi(x)\mathrm{d}x\right| \le \int\int |f(x)\phi(x-y)\psi(y)|\mathrm{d}x\mathrm{d}y.$$

Write the integrand as $\alpha(x,y)\beta(x,y)\gamma(x,y)$ with

$$\alpha(x,y) = |f(x)|^{\frac{r'}{p'}}|\phi(x-y)|^{\frac{q}{p'}}$$
$$\beta(x,y) = |\phi(x-y)|^{\frac{q}{r}}|\psi(y)|^{\frac{p}{r}}$$
$$\gamma(x,y) = |f(x)|^{\frac{r'}{q'}}|\psi(y)|^{\frac{p}{q'}}$$

and $1/p + 1/p' = 1$, $1/q + 1/q' = 1$. Note that $1/p' + 1/q' = 1/r' = 1 - 1/r$. So, by Holder,

$$\int\int |f(x)\phi(x-y)\psi(y)|\mathrm{d}x\mathrm{d}y \;\le\; \left[\int\int |f(x)|^{r'}|\phi(x-y)|^q \mathrm{d}x\mathrm{d}y\right]^{\frac{1}{p'}}$$
$$\times \left[\int\int |\phi(x-y)|^q|\psi(y)|^p\mathrm{d}x\mathrm{d}y\right]^{\frac{1}{r}} \left[\int\int |f(x)|^{r'}|\psi(y)|^p\mathrm{d}x\mathrm{d}y\right]^{\frac{1}{q'}}$$
$$=\; \|f\|_{r'}^{\frac{r'}{p'}+\frac{r'}{q'}}\,\|\phi\|_q^{\frac{q}{p'}+\frac{q}{r}}\,\|\psi\|_p^{\frac{p}{r}+\frac{p}{q'}}.$$

□

Now note that if $\phi \in C^\infty$ and f is integrable and compactly supported, say, then $\phi * f$ is C^∞. Indeed

$$\partial_j(\phi * f) \;=\; (\partial_j\phi) * f.$$

Likewise, if $\phi \in \mathcal{S}$ and f is integrable with compact support — or more generally polynomial decay at ∞, so $p(x)f \in L^1$ for any polynomial p — then $\phi * f \in \mathcal{S}$. If also $\psi \in \mathcal{S}$ then

$$\langle \psi, \phi * f\rangle \;=\; \left\langle \psi * \widetilde{f}, \phi \right\rangle,$$

where

$$\widetilde{f}(x) \;=\; f(-x).$$

Thus we define the convolution $T * f$ of a f, a compactly supported integrable function, and a distribution T to be the distribution

$$\langle \psi, T * f\rangle = \left\langle \psi * \widetilde{f}, T \right\rangle.$$

THEOREM 15.3. *If $f \in L^1$ has compact support and T is a tempered distribution, then $T * f \in \mathcal{S}^\star$ is a tempered distribution, and if $T \in L^1_{loc} \cap \mathcal{S}^*$ then $T * f \in L^1_{loc}$ as well and*

$$\langle \psi, T * f\rangle = \int_{\mathbb{R}^d \times \mathbb{R}^d} \psi(x) T(x-y) f(y) \mathrm{d}x\mathrm{d}y.$$

The map $\phi, T \mapsto T * \phi$ is a continuous map of $\mathcal{S} \times \mathcal{S}^\star \to C^\infty \cap \mathcal{S}^\star$.

PROOF. The first two statements are clear. To show that $T * \phi$ is C^∞ if $\phi \in \mathcal{S}$, note that this follows if T is a function since

$$\partial_j \int \phi(x-y) T(y) \mathrm{d}y \;=\; \int \partial_j \phi(x-y) T(y).$$

More generally, we may write a general distribution as $D^\alpha g$ with g continuous, and integrate by parts to see that

$$T * \phi(x) \;=\; \int D^\alpha \phi(x-y) g(y) \mathrm{d}y.$$

Continuity of the map is easy. □

Now, let $T(x) = \frac{1}{4\pi}|x|^{-1}$. This is a locally integrable function, so if $\rho \in L^1_{loc}$ we have

$$\phi(x) \;=\; \frac{1}{4\pi} \int \frac{1}{|x-y|} \rho(y) \mathrm{d}y$$

as a (weak) solution of the Poisson equation. It is not the unique solution since we can add to it any harmonic function U which satisfies $\Delta U = 0$. For example, $U(x) = x_1$ is harmonic. So is $U(x) = x_1 x_2$. In fact, one can show that any Harmonic tempered distribution is a polynomial. (The easy way to do this is to use the Fourier Transform.)

More generally we have the following theorem due to Weyl

THEOREM 15.4 (Weyl). *Let T be a distribution that satisfies $\Delta T = 0$ on an open set $D \in \mathbb{R}^n$, so $\langle \Delta\phi, T\rangle = 0$ for $\phi \in C_c^\infty(D)$. Then T is a C^∞ function in D: there is $g \in C^\infty(D)$ such that $\langle \phi, T\rangle = \langle \phi, g\rangle$ for $\phi \in C_c^\infty(D)$.*

For the proof, see appendix B.

LECTURE 16

The Dirichlet problem

Reading: §7.2 of Lax

Suppose we want to solve Poisson's equation for ϕ that vanishes outside an open set $D \subset \mathbb{R}^d$ containing the support of ρ. That is

$$\begin{cases} \Delta\phi(x) = \rho(x) & x \in D \\ \phi(x) = 0 & x \notin D. \end{cases}$$

It is useful to use Hilbert space methods. Let us set up the inner products

$$\langle f, g \rangle_0 = \int_D f(x)g(x)\mathrm{d}x \quad \text{and} \quad \langle f, g \rangle_1 = \int_D \sum_{j=1}^{d} \partial_j f(x)\partial_j g(x)\mathrm{d}x.$$

Take these over $C_c^\infty(D)$, which is incomplete but completes to $L^2(D)$ with respect to the first and to a Sobolev space in the second.

Lemma 16.1. *For any $f \in C_c^\infty(D)$*

$$\|f\|_0 \le \frac{\operatorname{diam}(D)}{\sqrt{d}} \|f\|_1$$

where $\operatorname{diam}(D) = \sup_{x,y\in D} |x - y|$ *is the width of* D.

Proof. This is a calculus exercise. Extend f to be identically 0 outside D. Note that

$$f(x) = \int_{x_1^{(0)}}^{x_1} \partial_1 f(y, x_2, \ldots, x_d)\mathrm{d}y,$$

where $x^{(0)} = (x_1^{(0)}, x_2, \ldots, x_d) \in \partial D$. Thus by Cauchy-Schwarz,

$$f(x)^2 \le \operatorname{diam}(D) \int_{\mathbb{R}} \partial_1 f(y, x_2, \ldots, x_d)\mathrm{d}y.$$

Integrating over $x \in D$ now gives

$$\|f\|_0^2 \le \operatorname{diam}(D)^2 \|\partial_1 f\|_0^2 .$$

Similarly,

$$\|f\|_0^2 \le \operatorname{diam}(D)^2 \|\partial_j f\|_0^2 .$$

Averaging these results gives the Lemma. □

Let $H_1^{(0)}$ denote the completion of $C_0^\infty(D)$ in the norm $\|\cdot\|_1$. This is a Hilbert space.

Lemma 16.2. *Every element $f \in H_1^{(0)}$ may be identified with a locally integrable function such that*

(1) $f(x) = 0$ if $x \notin D$
(2) $f \in L^2(D)$

(3) For $j = 1, \dots, d$, $\partial_j f$ (in the sense of distributions) is a locally integrable function in $L^2(D)$.

Furthermore we have

$$\langle f, g \rangle_1 \;=\; \sum_{j=1}^{D} \langle \partial_j f, \partial_j g \rangle_0$$

and

$$\langle f, \partial_j g \rangle_0 \;=\; -\langle \partial_j f, g \rangle$$

whenever $f, g \in H_1^{(0)}$.

Proof. The inequality $\|f\|_0 \le \text{const.}. \|f\|_1$ shows that any Cauchy sequence of function in the $\|\cdot\|_1$ norm is Cauchy in L^2. So any sequence $f_n \in C_c^\infty(D)$ which converges in the $H_1^{(0)}$ also converges in L^2. Identify the limit f with the corresponding element of $L^2(D)$, extended to be zero outside D. (1) and (2) follow.

To derive (3), note that for $g \in C_c^\infty$

$$\langle g, \partial_j f_n \rangle \;=\; -\langle \partial_j g, f_n \rangle$$

by integration by parts. The r.h.s. converges to $\langle \partial_j g, f \rangle$ using Cauchy-Schwarz. On the other hand, $\partial_j f_n$ converges in the L^2 norm to a function h_j. We conclude that $h_j = \partial_j f$ in the sense of distributions, so $\partial_j f \in L^2(D)$.

To show the formulas, note that they hold for elements of $C_c^\infty(D)$ and follow for the $f, g \in H_1^{(0)}$ by taking limits. $\square$

Now, fix $\rho \in L^2(D)$. This function gives rise to a linear functional ℓ on $H_1^{(0)} \subset L^2(D)$ by

$$\ell(u) \;=\; \langle u, f \rangle_0 .$$

Since

$$|\ell(u)| \;\le\; \|u\|_0 \|f\|_0 \;\le\; \frac{\operatorname{diam}(D)}{\sqrt{d}} \|f\|_0 \|u\|_1 ,$$

this is a bounded linear functional. By Riesz-Frechet there is an element $\phi \in H_1^{(0)}$ such that

$$\langle u, \rho \rangle_0 = \langle u, \phi \rangle_1 ,$$

for all $u \in H_1^{(0)}$. Specializing to $u \in C_c^\infty(D)$, we find that

$$f = -\Delta \phi$$

in the sense of distributions.

In a similar way, we may write down a bilinear form

$$B(u, v) \;=\; \int_D \left\{ \sum_{i,j} A_{i,j}(x) \partial_i u(x) \partial_j v(x) + u(x) \sum_i F_i(x) \partial_i v(x) + m(x) u(x) v(x) \right\} \mathrm{d}x.$$

Let us check the hypotheses of Lax-Milgram. Clearly B is bilinear and bounded. To estimate $B(u, u)$ from below, suppose that $A_{i,j}(x)$ is pointwise positive definite, that is

$$\sum_{i,j} A_{i,j}(x) \lambda_i \lambda_j \;\ge \sigma(x) \sum_i \lambda_i^2.$$

Then

$$B(u, u) \;\ge \int_D \left\{ \sigma(x) |\nabla u(x)|^2 - |u(x)| |\mathbf{F}(x)| |\nabla u(x)| + m(x) u(x)^2 \right\} \mathrm{d}x.$$

Now suppose the quadratic function

$$\sigma(x)t^2 - |\mathbf{F}(x)|\, ts + m(x)s^2 \ \geq \ \delta(t^2+s^2),$$

uniformly in x, for some $\delta > 0$. This amounts to the requirement that $\sigma(x) > 0$, $m(x) > 0$ and the discriminant

$$\sup_x |\mathbf{F}(x)|^2 - 4\sigma(x)m(x) < 0.$$

Then

$$B(u,u) \geq \delta(\|u\|_1^2 + \|u\|_0^2) \geq \delta \, \|u\|_1^2 .$$

Thus, by Lax-Milgram given $\rho \in L^2(D)$ we may find $\phi \in H_1^{(0)}$ such that

$$\langle v, \rho\rangle_0 \ = \ B(v,\phi) \quad \forall v \in H_1^{(0)}.$$

Such ϕ is a distributional solution to the equation

$$-\sum_{i,j} \partial_i A_{i,j}(x)\partial_j\phi(x) + \sum_i F_i(x)\partial_i\phi(x) + m(x)\phi(x) \ = \ \rho(x), \quad x \in D.$$

One might ask, in what sense do the elements of $H_1^{(0)}$ vanish outside D? The complete answer to this question involves the theory of Sobolev spaces. An elementary answer is the following:

LEMMA 16.3. *Let D be a hypercube $[-L,L]^d$ and let $f \in H_1^{(0)}$. Then*

$$\lim_{\delta\to 0} \frac{1}{|\{\mathrm{dist}(x,\partial D) < \delta\}|} \int_{\mathrm{dist}(x,\partial D)<\delta} |f(x)| = 0.$$

REMARK. $|\cdot|$ denotes Lebesgue measure, so $|\{\mathrm{dist}(x,\partial D) < \delta\}| = O(\delta)$,

PROOF. The set of points within distance δ of ∂D is a union of slabs such as $S = \{x \in D \ : \ -L < x_1 < -L+\delta\}$. Let $f \in C_c^\infty(D)$ and consider the integral of $|f(x)|$ over this slab. Integrating by parts, we have

$$\int_S |f(x)|\mathrm{d}x \ = \ \int_S \partial_1|f(x)|(-L+\delta-x_1)\mathrm{d}x,$$

since f vanishes on the boundary. Now $\partial_1|f(x)| \leq |\partial_1 f(x)|$. Thus,

$$\int_S |f(x)|\mathrm{d}x \leq \int_S |\partial_1 f(x)|(-L+\delta-x_1)\mathrm{d}x \leq \delta|S|^{\frac{1}{2}} \, \|f\|_1 \leq C\, \|f\|_1 \, \delta^{\frac{3}{2}},$$

since $|S| = O(\delta)$. Adding up the contributions from all slabs we get

$$\int_{\{\mathrm{dist}(x,\partial D)<\delta\}} |f(x)|\mathrm{d}x \leq C\, \|f\|_1 \, \delta^{\frac{3}{2}}.$$

Taking limits, this estimate follows, with constant C independent of f, for $f \in H_1^{(0)}$. Since $|\mathrm{dist}(x,\partial D) < \delta| \geq c\delta$ the result follows. □

REMARK. We were a bit wasteful. We could have averaged over just $\Sigma_\delta = \{\mathrm{dist}(x,\Sigma) < \delta\}$ with Σ a piece of the boundary with codimension 2. In that case $|\Sigma_\delta| \simeq \delta^2$ and

$$\int_{\Sigma_\delta} |f(x)|\mathrm{d}x \leq C\delta^2 \left(\int_{\Sigma_\delta} |\nabla f(x)|^2 \mathrm{d}x\right)^{\frac{1}{2}}.$$

The integral on the r.h.s. goes to zero, so the average of $|f|$ vanishes also on codimension 2 sets.

Part 5

Weak Convergence and Weak Topology

LECTURE 17

Dual of a Banach space

Reading: §8.2 -8.3 of Lax.

Recall that the dual of a Banach space X is the space X' of bounded linear functionals on X. Recall that we put a norm topology on X' by defining

$$\|\ell\| \;=\; \sup x \neq 0 \frac{|\ell(x)|}{\|x\|}.$$

Likewise, we have the following dual characterization of the norm on X:

THEOREM 17.1. *For every $x \in X$ we have*

$$\|x\| \;=\; \max \ell \neq 0,\;\; \ell \in X' \frac{|\ell(x)|}{\|\ell\|}.$$

PROOF. Since $\|\ell\|\,\|y\| \geq |\ell(y)|$ the l.h.s. is no smaller than the r.h.s. Thus we need only produce an ℓ such that $|\ell(x)| = \|x\|\,\|\ell\|$. Define ℓ first on the one dimensional subspace span$\{x\}$ by $\ell(tx) = t\,\|x\|$. Since this functional is norm bounded by 1 on this subspace it has an extension (by Hahn-Banach) to the whole space with this property. □

We also have the weak* topology on X', the weakest LCS topology such that all elements of X are continuous functionals on X'. Recall that $(X', \mathrm{wk}^*)^* = X$. One might wonder if $(X')' = X$ as Banach spaces. In fact this does not hold in general.

DEFINITION 17.1. A Banach space is called reflexive if $(X')' = X$. That is if every bounded linear functional on X' is of the form $\ell \mapsto \ell(x)$ for some $x \in X$.

Many important spaces are reflexive, but not all. For instance:

THEOREM 17.2. *$c_0' = \ell_1$, $\ell_1' = \ell_\infty$, and $\ell_\infty' \supsetneq \ell_1$.*

PROOF. Let ℓ be a linear functional on c_0. Evaluating ℓ on the sequence $\mathbf{e}_k(n) = 1$ if $k = n$ and 0 otherwise produces a sequence

$$\mathbf{b}(n) = \ell(\mathbf{e}_n).$$

If $\mathbf{a} \in c_0$ then

$$\mathbf{a} \;=\; \lim_{n\to\infty} \sum_{j=1}^{n} a(j)\mathbf{e}_j,$$

so by linearity and continuity

$$\ell(\mathbf{a}) \;=\; \lim_{n\to\infty} \sum_{j=1}^{n} a(j)\mathbf{b}(j).$$

That is for any $\mathbf{a} \in C_0$ the sequence $\mathbf{ab}$ is summable and

$$\ell(\mathbf{a}) = \sum_{j=1}^{\infty} \mathbf{a}(j)\mathbf{b}(j).$$

To see that $\mathbf{b} \in \ell_1$, take $\mathbf{a}_n(j) = \mathrm{e}^{-\mathrm{i}\arg \mathbf{b}(j)}$ if $j \leq n$ and 0 otherwise. Thus

$$\sum_{j=1}^{n} |\mathbf{b}(j)| = \ell(\mathbf{a}_n) \leq \|\ell\|.$$

It follows that $\mathbf{b} \in \ell_1$. (Conversely, it is clear that any $\mathbf{b} \in \ell_1$ gives a linear functional on c_0.)

The same idea works to prove $\ell_1' = \ell_\infty$. Finally, it is clear that $\ell_1 \subset \ell_\infty'$, however no Banach limit such as defined in the third lecture can be written as a scalar product with something in ℓ_1. □

However,

Theorem 17.3. *Every Hilbert space is reflexive.*

Proof. In this case $X' = X$ by Riesz-Fréchet. □

For L^p spaces we have.

Theorem 17.4. *Let X, μ be a finite measure space. For $1 \leq p < \infty$ the dual of $L^p(X)$ is $L^q(X)$ where $\frac{1}{p} + \frac{1}{q} = 1$.*

Remark. The result, but not the proof, extends to arbitrary measure spaces for $1 < p < \infty$. If $p = 1$ (so $q = \infty$) the result holds in σ-finite measure spaces.

Proof. The Hölder inequality shows that $L_q(X) \hookrightarrow L^p(X)'$, via the pairing

$$\langle f, g\rangle = \int_X f(x)g(x)\mathrm{d}\mu(x), \quad f \in L^p \text{ and } g \in: L^q.$$

Suppose now that ℓ is a linear functional on $L^p(X)$. For each measurable set A we have $\chi_A \in L^p(X)$. So

$$\nu(A) = \ell(\chi_A)$$

defines a finitely additive set function. In fact, it is countably additive since ℓ is continuous, and if $A_1, A_2, \ldots$ are pairwise disjoint then

$$\sum_{j=1}^{n} \chi_{A_j} \to \chi_{\cup A_j} \quad \text{in } L^p$$

as one may readily show. Furthermore $\nu(A) = 0$ if $\mu(A) = 0$ since then $\chi_A = 0$ in L^p. Thus $\nu << \mu$ and by Radon-Nikodym there is function $g \in L^1$ such that

$$\ell(\chi_A) = \nu(A) = \int_A g(x)\mathrm{d}\mu(x) = \int_X \chi_A(x)g(x)\mathrm{d}\mu(x).$$

By taking limits of simple functions we have

$$\ell(f) = \int_A f(x)g(x)\mathrm{d}\mu(x) \tag{$\star$}$$

whenever $f \in L^\infty$, using dominated convergence.

It remains to show that $g \in L^q$ for then $(\star)$ extends to all of L^p by density of L^∞. To this end, fix $t > 0$ and let

$$f_t(x) \;=\; \begin{cases} \frac{|g(x)|^q}{g(x)} & 0 < |g(x)| < t \\ 0 & |g(x)| = 0 \text{ or } |g(x)| \geq t. \end{cases}$$

Clearly $|f_t(x)| \leq t^{q-1}$ so $f_t \in L^\infty \subset L^p$. Thus

$$\int_{\{|g(x)|<t\}} |g(x)|^q \mathrm{d}\mu(x) = \int f_t(x) g(x) \mathrm{d}\mu(x) \;\leq\; \|\ell\| \left[\int |f_t(x)|^p \mathrm{d}\mu(x) \right]^{\frac{1}{p}}.$$

But $|f_t(x)|^p = |g(x)|^{pq-p} = |g(x)|^q$ on $\{|g(x)| < t\}$. Thus

$$\left[\int_{\{|g(x)|<t\}} |g(x)|^q \mathrm{d}\mu(x) \right]^{1-\frac{1}{p}} \leq \|\ell\|$$

and the result follows. □

COROLLARY 17.5. *$L^p(X)$ is reflexive for $1 < p < \infty$.*

This result also follows from

THEOREM 17.6 (Milman 1938). *Any uniformly convex Banach space is reflexive*

In general L^1 is not reflexive: $(L^1)' = L^\infty$ but L^∞ contains linear functionals that are not in L^1. Notice where the proof breaks down. Given a linear functional ℓ on L^∞ we can define a set function

$$\nu(A) = \ell(\chi_A)$$

as above. It is certainly additive, and absolutely continuous since if $\mu(A) = 0$ then $\chi_A = 0$ in L^∞. It is not, however, countably additive since

$$\sum_{j=1}^{n} \chi_{A_j} \not\to \chi_{\cup A_j} \text{ in } L^\infty.$$

The inequality $(L^\infty) \neq L^1$ also follows from:

THEOREM 17.7. *Let X be a Banach space over $\mathbb{C}$. If X' is separable so is X.*

PROOF. Let $\{\ell_n\}$ be a countable dense subset of X'. For each n there is $x_n \in X$ such that

$$\|x_n\| = 1 \quad \text{and} \quad \ell_n(x_n) \geq \frac{1}{2} \|\ell_n\|.$$

It suffices to show span$\{x_n\}$ is dense in X.

Suppose contrarily that $\overline{\text{span}}\{x_n\} \neq X$. Then there is a non-zero linear functional $\ell \in X'$ such that $\ell(x_n) = 0$ for all n. We may assume that $\|\ell\| = 1$. However, we can find n such that $\|\ell - \ell_n\| \leq \frac{1}{4}$, say. Thus $\|\ell_n\| \geq \frac{3}{4}$ and

$$0 = \ell(x_n) = \ell(x_n) - \ell_n(x_n) + \ell_n(x_n) \geq \frac{1}{2} \|\ell_n\| - \|\ell - \ell_n\| \geq \frac{1}{8}.$$

Thus no such ℓ exists and we must have $\overline{\text{span}}\{x_n\} = X$. □

LECTURE 18

Riesz-Kakutani theorem

Theorem 18.1 (Riesz-Kakutani). *Let Q be a compact Hausdorff space, $C(Q)$ the space of continuous real valued functions on Q with the* max *norm. Then $C(Q)' = M(Q) =$ set of signed Borel measures of finite total variation.*

That is to every bounded linear functional $\ell \in C(Q)'$ is associated a unique Borel measure m such that

$$\ell(f) \; = \; \int_Q f \mathrm{d}m.$$

Furthermore the norm of ℓ is the total variation $\|\ell\| = \int_Q |\mathrm{d}m|$.

Remark. i) A Borel measure is a measure on sets in the Borel σ-algebra which is the smallest σ-algebra containing the open sets on Q. ii) The dual of $C(Q;\mathbb{C})$ is the set of complex Borel measures, $m = m_r + \mathrm{i}m_i$ with $m_r, m_i \in M(Q)$. iii) Compactness is crucial here. (More on this later.)

Before proving this theorem, let us prove

Theorem 18.2. *Given $\ell \in C(Q)$ there is a unique deomposition $\ell = \ell_+ - \ell_-$ with $\ell_\pm$ positive linear functionals and $\|\ell\| = \ell_+(1) + \ell_-(1)$.*

Proof. Let $C(Q)_+ =$ set of non-negative functions in $C(Q)$. For $f \in C(Q)_+$ define

$$\ell_+(f) = \sup\{\ell(h) \ : \ 0 \leq h \leq f\}.$$

It is clear that $\ell_+(tf) = t\ell_+(f)$ for $t \geq 0$ and that $\ell_+(f) \geq \ell(0) = 0$ for $f \in C(Q)_+$. Given f, g in $C(Q)_+$ their sum is also in $C(Q)_+$. Clearly

$$\ell_+(f+g) \geq \sup\{\ell(h_1) + \ell(h_2) \ : \ 0 \leq h_1 \leq f \text{ and } 0 \leq h_2 \leq g\} \; = \; \ell_+(f) + \ell_+(g).$$

The opposite inequality clearly follows from the following result

Claim. *Given $f, g \in C(Q)$ and $0 \leq h \leq f+g$ we can write $h = h_1 + h_2$ with $h_{1,2} \in C(Q)$, $0 \leq h_1 \leq f$ and $0 \leq h_2 \leq g$.*

Proof of Claim. Let $h_1 = \min\{f, h\}$. So h_1 is continuous and $0 \leq h_1 \leq f$. Let $h_2 = h - h_1$. Since $h_1 \leq h$, $h_2 \geq 0$. When $h_1 = f$ we have $h_2 \leq f + g - f = g$. On the other hand if $h_1 = h$ we have $h_2 = 0$ so $h_2 \leq g$. □

Thus $\ell_+(tf + sg) = t\ell_+(f) + s\ell_+(g)$ whenever $t, s \geq 0$ and $f, g \in C(Q)_+$. Define ℓ_+ on all of $C(Q)$ by

$$\ell_+(f) = \ell_+(f_+) - \ell_+(f_-), \quad \text{with } f_+ = \max\{f, 0\} \text{ and } f_- = \min\{f, 0\}.$$

It is not hard to see that ℓ_+ is linear. Now set $\ell_- = \ell_+ - \ell$ and note that

$$\ell_-(f) = \sup\{\ell(h) - \ell(f) \ : \ 0 \leq h \leq f\} = \sup\{\ell(k) \ : \ -f \leq k \leq 0\} \geq 0,$$

for $f \in C(Q)_+$.

To prove $\|\ell\| = \ell_+(1) + \ell_-(1)$ note first that $\|\ell_\pm\| = \ell_\pm(1)$ since these are positive linear functionals. Thus, $\|\ell\| \le \|\ell_+\| + \|\ell_-\| = \ell_+(1) + \ell_-(1)$. On the other hand by the definition of ℓ_+ and the corresponding inequality for ℓ_- we have

$$\begin{aligned}\ell_+(1) + \ell_-(1) = \sup\{\ell(h) \ : \ 0 \le h \le 1\} + \sup\{\ell(k) \ : \ -1 \le k \le 0\} \\ = \sup\{\ell(g) \ : \ -1 \le g \le 1\} \ \le \ \|\ell\|. \quad \square\end{aligned}$$

We are now ready to prove the Riesz-Kakutani Theorem. By the splitting of a linear functional into positive and negative parts, it suffices to show

THEOREM 18.3 (Riesz-Kakutani). *Let ℓ be a positive linear functional on $C(Q)$ with Q a compact Hausdorff space. Then there is a unique positive Borel measure m on Q such that $\ell(f) = \int_Q f(x)\mathrm{d}m(x)$ and $\|\ell\| = m(Q)$. Conversely, any positive Borel measure m gives a positive linear functional on $C(Q)$ via $\langle f, m\rangle = \int_Q f(x)\mathrm{d}m(x)$.*

PROOF. First, it is clear that any positive Borel measure m gives rise to a positive linear functional by $\ell(f) = \int_Q f(x)\mathrm{d}m(x)$. (Continuous functions are Borel measurable.) Since the functional is positive $\|\ell\| = \ell(1) = m(Q)$.

It remains to show that every positive linear functional is of this form. So let $\ell \in C(Q)'$ be given. We would like to define $m(S) = \ell(\chi_S)$, for S a measurable set. However, $\chi_S \notin C(Q)$ (unless S is both open and closed, so a union of connected components). Thus we do the next best thing: given an open set U we take

$$m(U) = \sup\{\ell(f) \ : \ f \prec U\},$$

where, $f \prec U$ indicates that $f \in C(Q)$, $\operatorname{supp} f \subset U$ and $0 \le f \le 1$.

CLAIM. *Let $U_1, U_2 \subset Q$ be a disjoint pair of open sets. Then $m(U_1 \cup U_2) = m(U_1) + m(U_2)$.*

PROOF OF CLAIM. Note that given $f_j \prec U_j$, we have $f_1 + f_2 \prec U_1 \cup U_2$ and conversely given $f \prec U_1 \cup U_2$, we have $f\chi_{U_j} \prec U_j$. The identity follows. $\square$

Thus m is finitely additive on open sets. So far the construction works on an arbitrary compact space, but the open sets do not form a σ-algebra. A first step is to define m on closed sets F by

$$m(F) \ = \ m(Q) - m(F) \ = \ \inf\{\ell(f) : F \prec f\},$$

where $F \prec f$ if $0 \le f \le 1$ and $f(x) = 1$ on F.

To proceed we need the following result which uses the assumption that Q is a compact Hausdorff space:

CLAIM. *If $F \subset U$, F closed and U open, then*

$$m(F) = \inf\{\ell(f) : F \prec f \prec U\} \le \sup\{\ell(f) \ : \ F \prec f \prec U\} = m(U).$$

PROOF OF CLAIM. It is clear that

$$m(F) \le \inf\{\ell(f) : F \prec f \prec U\} \le \sup\{\ell(f) \ : \ F \prec f \prec U\} \le m(U).$$

To obtain equality on the two ends, we use Urysohn's lemma, valid here since a compact Hausdorff space is *normal* (disjoint closed sets may be separated by open sets):

THEOREM 18.4 (Uryshohn's Lemma). *Given two disjoint closed sets $F_0, F_1 \subset Q$ there is a continuous function $f : Q \to [0,1]$ such that $f \equiv 0$ on F_0 and $f \equiv 1$ on F_1.*

Given $f_1 \prec U$, let $F_1 = F \cup \operatorname{supp} f_1$, and let V be open with $F_1 \subset V \subset \overline{F} \subset U$ — such V exists because Q is normal. Now let $F_0 = V^c$. Then $f \equiv j$ on F_j, $j = 0, 1$ clearly satisfies $f_1 \leq f$ and $F \prec f \prec U$. We conclude that

$$\sup\{\ell(f) \ : \ F \prec f \prec U\} = \sup\{\ell(f) \ : \ f \prec U\} = m(U).$$

Similarly given $f_1 \succ F$, let $F_1 = F$ and $F_0 = U^c$. Pick $f_2 \in C(Q)$ with $F \prec f_2 \prec U$ and set $f = f_1 f_2$. Clearly $F \prec f \prec U$ and $f \leq f_1$. Thus $\ell(f) \leq \ell(f_1)$ so

$$\inf\{\ell(f) : F \prec f \prec U\} = \inf\{\ell(f) : F \prec f\}. \quad \square$$

It now follows that

Claim. *If F is closed then $m(F) = \inf\{m(U) : F \subset U$ and U open$\}$.*
If U is open then $m(U) = \sup\{m(F) : F \subset U$ and F closed$\}$.

Proof of claim. Clearly $m(F) \leq \inf\{m(U) : F \subset U\}$. To prove the converse, note that given $\epsilon > 0$ there is $f \succ F$ such that $\ell(f) \leq m(F) + \epsilon$. Let $U = \{x \ : \ f(x) > 1 - \epsilon\}$. Note that $g \prec U \implies g \leq \frac{1}{1-\epsilon} f$ so $\ell(g) \leq \frac{1}{1-\epsilon}\ell(f)$. Thus $m(U) \leq \frac{1}{1-\epsilon}\ell(f)$. Thus $(1-\epsilon)m(U) \leq m(F) + \epsilon$. It follows that $\inf\{m(U) : F \subset U\} \leq m(F)$.
The opposite identity, $m(U) = \sup\{m(F)\}$, follows by taking complements. $\square$

Now given arbitrary $S \subset Q$ define

$$m_+(S) = \inf\{m(U) : S \subset U \text{ and } U \text{ open}\},$$

$$m_-(S) = \sup\{m(F) : S \supset F \text{ and } F \text{ closed}\}.$$

Clearly $m_-(S) \leq m_+(S)$. If these two numbers are equal define

$$m(S) = m_\pm(S) \text{ if } m_+(S) = m_-(S).$$

Claim. *The collection $\Sigma = \{S \subset Q : m_+(S) = m_-(s)\}$ is a σ-algebra containing all Borel sets and m defines a countably additive measure on this σ-algebra.*

Proof of Claim. Clearly Σ contains all closed and all open sets. Thus once we show it is a σ-algebra it is immediate that it contains all Borell sets. Clearly $S \in \Sigma \implies S^c \in \Sigma$. Thus we need only show that Σ is closed under countable unions. This is left as an exercise as is countable additivity of m. $\square$

To complete the proof, we must show that $\int f \mathrm{d}m = \ell(f)$. It suffices to prove this for f with $0 \leq f \leq 1$. Note that

$$\int_Q f \mathrm{d}m \ = \ \int_0^1 m\{x : f(x) \geq t\}\mathrm{d}t = \int_0^1 m\{x : f(x) > t\}\mathrm{d}t.$$

Let us show that $\int f \mathrm{d}m \leq \ell(f)$. Note that

$$\int_Q f \mathrm{d}m \ \leq \ \sum_{j=1}^n \frac{1}{n} m\left\{x : f(x) \geq \frac{j-1}{n}\right\}.$$

Let $g_{j;n}$ be functions with $\left\{f(x) \geq \frac{j-1}{n}\right\} \prec g_{j;n} \prec \left\{f(x) > \frac{j-2}{n}\right\}$. Such functions exist by Urysohn's Lemma. So

$$\int_Q f \mathrm{d}m \leq \sum_{j=1}^n \frac{1}{n}\ell(g_{j;n}) = \ell(\sum_{j=1}^n \frac{1}{n} g_{j;n}).$$

Now for every $x \in Q$

$$\sum_{j=1}^{n} \frac{1}{n} g_{j;n}(x) = \sum_{j=1}^{\lfloor nf(x) \rfloor + 1} \frac{1}{n} + O(\frac{1}{n}) = \frac{1}{n} \lfloor nf(x) \rfloor + O(\frac{1}{n}).$$

We conclude that

$$\left| \sum_{j=1}^{n} \frac{1}{n} g_{j;n}(x) - f(x) \right| = O(\frac{1}{n}),$$

so

$$\ell(\sum_{j=1}^{n} \frac{1}{n} g_{j;n}) \longrightarrow \ell(f),$$

and $\int f \mathrm{d}m \leq \ell(f)$.

To show the reverse inequality, note that

$$\int_Q f \mathrm{d}m \geq \sum_{j=1}^{n} m \left\{ f(x) > \frac{j}{n} \right\}.$$

Thus,

$$\int_Q f \mathrm{d}m \geq \ell(\sum_{j=1}^{n} \frac{1}{n} h_{j;n}),$$

with $\{f(x) \geq \frac{j+1}{n}\} \prec h_{j;n} \prec \{f(x) > \frac{j}{n}\}$. Then

$$\sum_{j=1}^{n} \frac{1}{n} h_{j;n}(x) = \sum_{j=1}^{\lfloor nf(x) \rfloor - 1} \frac{1}{n} + O(\frac{1}{n}) = \frac{1}{n} \lfloor nf(x) \rfloor + O(\frac{1}{n}).$$

Again $\sum_j h_{j;n} \to f$, so $\int f \mathrm{d}m \geq \ell(f)$, completing the proof. $\square$

A word about the non-compact case. Suppose X is a locally compact Hausdorff space, so X is a Hausdorff space such that every point is contained in an open set with compact closure. Then we can consider several spaces $C_c(X) \subset C_0(X) \subset C_b(X) \subset C(X)$ where $C_b(X) = \{\text{bounded continuous functions}\}$. The middle two are Banach spaces in the sup norm. The first and last are LCS spaces: $C_c(X)$ has an inductive limit topology obtained from writing it as $C_c(X) = \cup_U C_0(U)$ where the union is over open sets U with compact closure, $C(X)$ has a topology generated by the seminorms $p_K(f) = \sup_{x \in K} |f(x)|$ for compact K. Regarding the duals of these spaces, we have

$$C_c(X)' = M(X) = \{\text{Borel measures } m \text{ such that } |m|(K) < \infty \text{ for any compact } K\}.$$
$$C_0(X)' = M_0(X) = \{\text{finite Borel measures on } X\}.$$
$$C_b(X)' = \{\text{finite Borel measures on the Stone-Čech compactification of } X\}.$$
$$C(X)' = \{\text{compactly supported Borel measures}\}.$$

LECTURE 19

Weak convergence

Reading: §10.1 and 10.2 of Lax.

We have already defined weak topologies in the general context of LCS spaces. Let us now look at them in the special case of Banach spaces. Consider a Banach space X and it's Banach space dual X'. So X' is a collection of linear functionals on X. The weakest topology on X so that every element of X' is continuous is called the weak topology on X, denoted $\sigma(X, X')$.

A sequence $\{x_n\}$ in X converges weakly to x — converges in the weak topology — if

$$\ell(x_n) \;\to\; \ell(x) \quad \text{for every } \ell \in X'.$$

Sometimes this is denoted

$$x_n \rightharpoonup x,$$

or

$$\text{wk} - \lim_{n\to\infty} x_n = x.$$

This notion is weaker than strong convergence, which is in norm:

$$\|x_n - x\| \to 0, \qquad \text{or } x_n \to x.$$

That is *more sequences converge weakly than converge strongly.* For instance

$$\chi_{[n,n+1]} \rightharpoonup 0 \quad \text{in } L^p(\mathbb{R}),\ 1 < p < \infty,$$

but

$$\left\|\chi_{[n,n+1]}\right\|_{L^p} = 1.$$

(Note that $\chi_{[n,n+1]}$ does not converge weakly in L^1.)

PROPOSITION 19.1. *Let $\{x_n\}$ be an orthonormal sequence in a Hilbert space H. Then $x_n \rightharpoonup 0$.*

REMARK. Since $\|x_n\| = 1$, x_n does not converge strongly to 0.

PROOF. Fix $y \in H$. By Bessel's inequality

$$\sum_n |\langle y, x_n\rangle|^2 \;\le\; \|y\|^2,$$

we see that $\langle y, x_n\rangle \to 0$. By the Riesz theorem on linear functionals on a Hilbet space, $x_n \rightharpoonup 0$. □

THEOREM 19.2. *Suppose $\{x_n\} \in X$, a Banach space, satisfies*

(1) x_n are uniformly bounded: $\sup_n \|x_n\| < \infty$.
(2) $\lim \ell(x_n) = \ell(x)$ for $\ell \in Y'$ with Y' dense in X'.

Then $x_n \rightharpoonup x$.

PROOF. This is an easy approximation argument and is left as an exercise. □

The interesting thing is that the converse is true: weakly convergent sequences are uniformly bounded. To prove this we will use

Theorem 19.3 (Principle of Uniform Boundedness for a complete metric space). *Let X be a complete metric space and $\mathcal{F}$ a collection of real valued continuous functions on X. If $\mathcal{F}$ is bounded at each point $x \in X$,*

$$|f(x)| \leq M(x) < \infty \quad \text{for all } f \in \mathcal{F},$$

then there is an open set $U \subset X$ and a constant $M < \infty$ such that

$$|f(x)| \leq M \quad \text{for all } x \in U \text{ and } f \in \mathcal{F}.$$

Proof. This result follows from the Baire Category Theorem of topology:

Theorem 19.4 (Baire Category Theorem). *A complete metric space is not the union of a countable number of nowhere dense sets.*

Remark. Recall that the interior of a set S is the largest open set contained in S, that is

$$\operatorname{int} S = S^o = \bigcup \{U \subset S \ : \ U \text{ open}\},$$

and that a set S is nowhere dense if it's closure $\overline{S}$ has empty interior. Thus a closed set is nowhere dense if $S \subset \overline{S^c}$. For the proof see Reed and Simon Chapter III or any book on point set topology.

To prove the PUB, note that by assumption

$$X = \bigcup_n \{x \ : \ |f(x)| \leq n \text{ for all } f \in \mathcal{F}\}.$$

Thus, at least one of the (closed) sets $\{x \ : \ |f(x)| \leq n \ \forall f \in \mathcal{F}\}$ has non-empty interior, which is to say it contains an open set U. This is the open set claimed in the theorem. □

Suppose X is a Banach space and each function $f \in \mathcal{F}$ is sub-additive ($f(x+y) \leq f(x) + f(y)$) and absolutely homogeneous ($f(ax) = |a| f(x)$). For instance each f could be of the form $f(x) = |\ell(x)|$ for some linear functional. Then

Theorem 19.5 (Principle of Uniform Boundedness for sub-additive functionals). *Let X be a Banach space and let $\mathcal{F}$ be a collection of real-valued continuous, sub-additive, absolutely homogeneous functions on X. Suppose for each $x \in X$, $|f(x)| \leq M(x) < \infty$ for all $f \in \mathcal{F}$. Then the function $f \in \mathcal{F}$ are uniformly bounded in the sense that there is $c < \infty$ such that*

$$|f(x)|| \ \leq \ c \|x\| \quad \text{for all } x \in X \text{ and } f \in \mathcal{F}.$$

Proof. Clearly the hypotheses of the PUB for metric spaces applies. Let U be the open set claimed and let $x_0 \in U$. Since U is open there is $\epsilon > 0$ such that $\|y\| < \epsilon \implies x_0 + y \in U$. Now consider y with $\|y\| < \epsilon$. We have, for $f \in \mathcal{F}$,

$$f(y) = f(y + x_0 - x_0) \leq f(y + x_0) + f(x_0) \leq 2M.$$

Thus for arbitrary $x \in X$ and $f \in \mathcal{F}$,

$$f(x) = \frac{2\|x\|}{\epsilon} f(\frac{\epsilon}{2\|x\|} x) \leq \frac{4M}{\epsilon} \|x\| . \quad \square$$

COROLLARY 19.6. *Let X be a Banach space and let $\mathcal{L}$ be a collection of bounded linear functionals that is pointwise bounded, so $\ell(x) \leq M(x)$ for all $\ell \in \mathcal{L}$, then there is a constant $c < \infty$ such that*

$$\|\ell\| \ \leq \ c \quad \text{for all } \ell \in \mathcal{L}.$$

COROLLARY 19.7. *Let X be a Banach space and let $S \subset X$ be a weakly pre-compact subset of X. Then there is a constant $c < \infty$ such that*

$$\|x\| \leq c \quad \text{for all } x \in S.$$

In particular, any weakly convergent sequence is norm bounded.

REMARK. Recall that a set S is pre-compact if $\overline{S}$ is compact. Thus any sequence in a pre-compact set has a convergent subsequence, but the limit may lie outside S.

PROOF. Think of points of X as functions on X'. Since S is weakly compact $\ell(x)$ must be bounded for each ℓ as x ranges over S (otherwise we could find a weakly divergent sequence). By the PUB there is a constant c such that $\|x\| \leq c$ for all $x \in S$. □

A function $f : X \to \mathbb{R}$, X a topological space, is called lower semi-continuous if $\{x : f(x) > t\}$ is open for each $t \in \mathbb{R}$. Such a function satisfies $f(x) \leq \liminf_n f(x_n)$ for any convergent sequence $x_n \to x$. (Note that for each $\epsilon > 0$ the set $\{y : f(y) > f(x) - \epsilon\}$ is open and thus eventually contains x_n so $\liminf x_n \geq f(x) - \epsilon$.)

THEOREM 19.8 (Weak lower semicontinuity of the norm). *Let X be a Banach space. The norm $\|\cdot\|$ is weakly lower semicontinuous. In particular, if $x_n \rightharpoonup x$ in X then*

$$\|x\| \leq \liminf \|x_n\| .$$

REMARK. 1) This should remind you of Fatou's lemma from measure theory. 2) We have already seen that the norm is not continuous, since it may "jump down" in a limit.

PROOF. Fix $t \geq 0$. Let X'_1 denote the unit ball $\{\ell : \|\ell\| \leq 1\}$ in X'. Note that $\|x\| > t$ if and only if there is a linear functional $\ell \in X'_1$ with $|\ell(x)| > t$. Thus

$$\{\|x\| > t\} = \cup_{\ell \in X'_1} \{x \ : \ \ell(x) > t\}$$

is weakly open. □

LECTURE 20

Weak sequential compactness, weak* convergence and the weak* topology

Definition 20.1. A subset C of a Banach space X is called weakly sequentially compact if any sequence of pints in C has a weakly convergent subsequence, whose weak limit is in C.

Recall that sequential compactness is, in general, a strictly weaker notion than compactness. They are equivalent, however, in metric spaces. In the present context, X is metrizable in the $\sigma(X, X')$ topology if and only if X' is separable.

Proposition 20.1. *A weakly sequentially compact set is bounded*

Proof. Use the PUB. Details left as an exercise. $\square$

Theorem 20.2. *In a reflexive Banach space X the closed unit ball is weakly sequentially compact.*

Remark. We will see that the unit ball is, in fact, weakly compact in a reflexive Banach space, and more generally in the dual of a Banach space. However, the proof of that result is far less constructive than the following.

Proof. Let $\{y_n\}$ be a sequence of points in the unit ball. Let Y be the closed linear span $\overline{\mathrm{span}}\{y_n\}$. Since X is reflexive, it follows that Y is reflexive, since

Theorem 20.3 (Thm. 15 of Chapter 8 in Lax). *Any closed subspace of a reflexive space is reflexive.*

Proof. See Lax. $\square$

Since $Y = Y''$ is separable, it follows that Y' is separable. So Y' contains a dense countable subset $\{m_j\}$. Consider the array of scalars

$$a_{i,j} \; = \; m_i(y_j).$$

The i^{th} row is bounded by $\|m_i\|$. Starting with the first row we pick a subsequence $y_{j_k^{(1)}}$ so that $m_1(y_{j_k^{(1)}})$ converges. Refine this subsequence again and again to produce subsequences $y_{j_k^{(n)}}$ for each n so that $m_1(y_{j_k^{(n)}}), \ldots, m_n(y_{j_k^{(n)}})$ all converge. Now let

$$z_n = y_{j_n^{(n)}}.$$

Clearly $m_j(z_n)$ converges as $n \to \infty$ for each j. By density of $\{m_j\}$ in Y' it follows that

$$\lim_{n\to\infty} m(z_n) = y(m)$$

exists for each $m \in Y'$. This limit is clearly a linear functional of m and since

$$\|m(z_n)\| \leq \|m\| \, \|z_n\| \leq \|m\| \, ,$$

the linear functional is bounded. Since Y is reflexive, we see that there is $y \in Y$ such that $y(m) = m(y)$. Since the restriction of $\ell \in X'$ to Y gives an element of Y', we have $z_n \rightharpoonup y$ in X. $\square$

We may also consider a weak topology on the dual X' of a Banach space. That is the weak $*$ topology $\sigma(X', X)$. A sequence u_n of linear functionals is said to be weak $*$ convergent to u if

$$\lim u_n(x) = u(x) \quad \text{ for all } x \in X,$$

also denoted

$$\text{wk}^* - \lim_{n\to\infty} u_n = u.$$

Weak* convergence of measures is also known as vague convergence. If X is reflexive then weak* convergence is the same as weak convergence, but in general the weak* topology is strictly weaker than the weak topology since the latter makes all linear functionals in X'' continuous.

THEOREM 20.4. *A weak* convergent sequence u_n is uniformly bounded and*

$$\|u\| \leq \liminf \|u_n\|,$$

if $u = \text{wk}^* - \lim u_n$.

PROOF. Exercise. $\square$

DEFINITION 20.2. A subset C of a dual Banach space X' is weak* sequentially compact if every sequence of points in C has a weak* convergent subsequence with weak* limit in C.

THEOREM 20.5 (Helly 1912). *Let X be a separable Banach space. Then the closed unit ball in X' is weak* sequentially compact.*

PROOF. Given $u_n \in X'$ with $\|u_n\| \leq 1$ and a countable dense subset $\{x_n\}$ of X, we can use the diagonal process to select a subsequence v_n of u_n so that

$$\lim_{n\to\infty} v_n(x_k)$$

exists for every x_k. By density of $\{x_k\}$ this extends to all of X:

$$\lim_{n\to\infty} v_n(x) = v(x)$$

for all $x \in X$. One readily verifies that v is linear and bounded, so it is the desired limit. $\square$

In fact, more is true. The unit ball in X' is weak* compact, even if X is no separable:

THEOREM 20.6 (Alaoglu). *Let X' be the dual of a Banach space X. The unit ball of X' is* wk* *compact.*

PROOF. Let B be the unit ball in X'. Let T be the (uncountable) product space:

$$T = \prod_{x \in X} I_x, \quad I_x = [-\|x\|, \|x\|].$$

By the Tychonov theorem T is compact in the product topology. To complete the proof, we embed B as a closed subset of T.

The infinite product space T is the collection of all functions $F : X \to \mathbb{R}$ such that $F(x) \in I_x$ for all x. Given $\ell \in B$, $|\ell(x)| \leq \|\ell\| \|x\| \leq \|x\|$ so $\ell(x) \in I_x$ for every x. Thus $B \subset T$.

Now the product topology on T is just the weakest topology such that coordinate evaluation $F \mapsto F(x)$ is continuous for every x. The restriction of this topology to B is just the wk* topology on B.

Thus we have embedded B as a subset of the compact space T. It suffices to show that B is closed. For each $x, y \in X$ and $t \in \mathbb{R}$, let

$$\Phi_{x,y;t}(F) = F(x+ty) - F(x) - F(y),$$

a continuous map of T into the field of scalars. Clearly $B \subset \Phi^{-1}_{x,y;t}(\{0\})$ and $\Phi^{-1}_{x,y;}(\{0\})$ is a closed set. Thus

$$\overline{B} \subset \bigcap_{x,y,t} \Phi^{-1}_{x,y;}(\{0\}),$$

so every element of $\overline{B}$ is linear. Since any $F \in T$ is also bounded by $\|x\|$, $|F(x)| \leq \|x\|$, we conclude that $\overline{B} = B$. □

Clearly Alaoglu's theorem implies Helly's theorem. However, the proof of Helly's theorem is much more useful. Often times what one really wants is to find a convergent sequence. The proof Helly's theorem gives you an idea how to construct it; Alaoglu's theorem just tells you it is there.

Corollary 20.7. *The unit ball in a reflexive space is weakly compact.*

Remark. In fact weak compactness of the unit ball is equivalent to reflexivity, a result due to Eberlein (1947) and Smulyan (1940).

In particular, this result applies to any Hilbert space and to L^p, $1 < p < \infty$. The unit ball in L^∞ is weak* compact since $L^\infty = (L^1)'$. The unit ball in L^1 is not weakly compact.

Here is what happens in L^1. Consider, for example, $L^1([0,1])$, and let

$$f_n(x) = n\chi_{[0,\frac{1}{n}]}(x).$$

So $\|f_n\|_{L^1} = 1$ and for any continuous function $g \in C([0,1])$ T$\int_0^1 g(x) f_n(x) \mathrm{d}x \to g(0)$.T Thus wk* $\lim f_n \mathrm{d}x = \delta(x)\mathrm{d}x$ in $M([0,1])$ but f_n has no weak limit in L^1. Of course, the sequence f_n has a weak* convergent subsequence in $L^\infty([0,1])'$, which shows the existence of a linear functional on L^∞ that restricts to $g \mapsto g(0)$ for continuous functions g. (We could have used the Hahn Banach theorem to get this.)

Here is another example. On $L^1([0,\infty))$ let

$$f_n(x) = \frac{1}{n}\chi_{[0,n]}(x).$$

Again $\|f_n\|_{L^1} = 1$. As measures $f_n \mathrm{d}x \rightharpoonup 0$ in $M_0([0,\infty))$, that is

$$\int_0^\infty f_n(x) g(x) \mathrm{d}x \;\longrightarrow\; 0 \quad g \in C_0([0,\infty)),$$

however, f_n does not converge weakly to zero in L^∞. Indeed for the constant function $g \equiv 1$,

$$\int_0^\infty f_n(x) g(x) = 1.$$

LECTURE 21

An application: positive harmonic functions

Reading: §11.6 in Lax

Positive harmonic functions

We may apply weak* compactness to prove the following:

THEOREM 21.1 (Herglotz). *Let u be a function on the open unit disk $D = \{|z| < 1\}$ such that*

(1) $u(z) \geq 0$ for all $z \in D$
(2) u is harmonic in D, that is it satisfies the mean value property

$$u(z) = \frac{1}{\pi \varepsilon^2} \int_{D_\varepsilon(z)} u(w) \mathrm{d}m(w),$$

for all $z \in D$ and $\varepsilon < 1 - |z|$. Here m is Lebesgue measure on the disk and $D_\varepsilon(z)$ is the disk of radius ε centered at z.

Then there is a unique finite, non-negative, Borel measure μ on $\partial D = \{|z| = 1\}$ such that

$$u(z) = \int_{\partial D} \frac{1 - |z|^2}{|z - w|^2} \mathrm{d}\mu(w). \tag{$\star$}$$

Conversely, any such function is a non-negative Harmonic function on the disk.

REMARK. The theorem implies $|u(z)| \leq \text{const.}/(1 - |z|)$, so non-negative harmonic functions cannot blow up arbitrarily at ∂D. One might wonder if a similar theorem holds for, say, real valued Harmonic functions. That is, given u Harmonic and real valued does there exist a signed measure μ such that u is the Poisson integral of μ? A moment's thought shows that the answer is "No!", for it is easy to construct a real valued harmonic function which violates the estimate $|u(z)| \leq \text{const.}/(1 - |z|)$. For example, $u(z) = \operatorname{Re} 1/(1 - z)^2$.

PROOF. Note that u is continuous. To see this, first observe that the mean value property implies that u is locally integrable. Next, observe that

$$u(z + h) - u(z) = \frac{1}{\pi \varepsilon^2} \left[\int_{D_\varepsilon(z+h)} u(w) \mathrm{d}m(w) - \int_{D_\varepsilon(z)} u(w) \mathrm{d}m(w) \right].$$

By dominated convergence the integral on the r.h.s. converges to zero as h converges to zero.

Since u is continuous, we may differentiate

$$\pi r^2 u(0) = \int_{D_r(0)} u(z) \mathrm{d}m(z) \;=\; \int_0^r \int_0^{2\pi} u(s\mathrm{e}^{\mathrm{i}\theta}) \mathrm{d}\theta s \mathrm{d}s$$

with respect r and conclude that for every r

$$u(0) = \frac{1}{2\pi} \int_0^{2\pi} u(r\mathrm{e}^{\mathrm{i}\theta}) \mathrm{d}\theta.$$

Thus for each $r \in (0,1)$ the measure $\mathrm{d}\mu_r(\theta) = (2\pi)^{-1} u(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\theta$ on the circle ∂D has mass $u(0)$. Think of these measures as elements of the dual to $C(\partial D)$. By Helly's theorem, we may find a weak* convergent subsequence μ_{r_n}. That is, there is a Borel measure $\mu \in C(\partial D)'$ such that

$$\int_{\partial D} f(\theta)\mathrm{d}\mu(\theta) = \lim_{n\to\infty} \frac{1}{2\pi} \int_{\partial D} f(\theta) u(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\theta$$

for every $f \in C(\partial D)$.

To complete the proof, we will use the identity

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{r^2 - |z|^2}{|r\mathrm{e}^{\mathrm{i}\theta} - z|^2} u(r\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\theta, \qquad (\star\star)$$

valid for $z \in D_r(0)$. Let us defer the proof for the moment and show how $(\star\star)$ implies the representation $(\star)$. Fix z and let

$$f_{r,z}(\mathrm{e}^{\mathrm{i}\theta}) = \frac{r^2 - |z|^2}{|r\mathrm{e}^{\mathrm{i}\theta} - z|^2}.$$

It is easy to see that $f_{r,z} \to f_z$ uniformly as $r \to 1$, where

$$f_z(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1 - |z|^2}{|\mathrm{e}^{\mathrm{i}\theta} - z|^2}.$$

Thus the weak* convergence $\mu_{r_n} \to \int \cdot \mu$ then implies

$$u(z) = \lim_n \int_{\partial D} f_{r_n;z}(\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\mu_{r_n}(\theta) = \int_0^{2\pi} f_z(\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\mu(\theta).$$

The identity $(\star)$ is a classical formula, which may be verified in a number of ways. One of these is as follows. Let $v(z)$ denote the integral on the r.h.s. It is easy to show that v is harmonic in $D_r(0)$ — for this it suffices to show that $(|w|^2 - |z|^2)/|w-z|^2$ is harmonic in $D_{|w|}(0)$ for fixed w. Furthermore, it is not too hard to show that

$$\lim_{s\uparrow r} v(r\mathrm{e}^{\mathrm{i}\theta}) = u(r\mathrm{e}^{\mathrm{i}\theta}),$$

since for any continuous function f on the circle

$$\frac{1}{2\pi} \lim_{s\uparrow r} \int_0^{2\pi} \frac{r^2 - s^2}{|r\mathrm{e}^{\mathrm{i}\theta} - s\mathrm{e}^{\mathrm{i}\phi}|^2} f(\mathrm{e}^{\mathrm{i}\phi})\mathrm{d}\phi \ = f(\mathrm{e}^{\mathrm{i}\theta}).$$

(Exercise: verify this formula.) Thus $u(z) - v(z)$ is a harmonic function on $D_{r_n}(0)$, continuous up to the boundary and identically equal to zero there. It follows from the maximum principle, applied to $u - v$ and $v - u$, that $u - v = 0$ throughout. (The maximum principle is a straightforward consequence of the mean value property and continuity.) □

Herglotz-Riesz Theorem

An important application of the above is the following:

THEOREM 21.2 (Herglotz-Riesz). *Let F be an analytic function in the unit disk D such that $\operatorname{Re} F \geq 0$ in D. Then there is a unique non-negative, finite, Borel measure μ on ∂D such that*

$$F(z) \ = \ \int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}\theta} + z}{\mathrm{e}^{\mathrm{i}\theta} - z} \mu(\mathrm{d}\theta) + \mathrm{i}\operatorname{Im} F(0).$$

Conversely every analytic function in the disk with positive real part can be written in this form.

Proof. First apply the Herglotz theorem to $\operatorname{Re} F$. Let

$$G(z) = \int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}\theta} + z}{\mathrm{e}^{\mathrm{i}\theta} - z} \mu(\mathrm{d}\theta).$$

So G and F are analytic functions on the disk whose real parts agree. It follows that $F - G$ is constant and imaginary. However $G(0) = \operatorname{Re} F(0)$ so $F(z) - G(z) = \mathrm{i} \operatorname{Im} F(0)$. □

The theorem is often used in the following form

Theorem 21.3. *Let F be an analytic map from the upper half plane $\{z \ : \ \operatorname{Im} z > 0\}$ into itself. Then there is a unique non-negative Borel measure μ on $\mathbb{R}$ and a non-negative number $A \geq 0$ such that*

$$\int_{\mathbb{R}} \frac{1}{1+x^2} \mathrm{d}\mu(x) < \infty$$

and

$$F(z) = Az + \operatorname{Re} F(\mathrm{i}) + \int_{\mathbb{R}} \frac{1+xz}{x-z} \frac{1}{1+x^2} \mathrm{d}\mu(x). \qquad (\star\star\star)$$

Furthermore

$$A = \lim_{z\to\infty} \frac{F(z)}{z},$$

and

$$\mathrm{d}\mu(x) = \mathrm{wk}^* \lim_{y \downarrow 0} \frac{1}{\pi} \operatorname{Im} F(x + \mathrm{i}y) \mathrm{d}x.$$

If $\lim_{z\to\infty}(F(z) - Az) = B$ exists and is real, and if $\lim_{z\to\infty} z(F(z) - Az - B)$ exists then μ is a finite measure and

$$F(z) = Az + B + \int_{\mathbb{R}} \frac{1}{x-z} \mathrm{d}\mu(x).$$

Remark. Note that

$$\frac{1+xz}{x-z} \frac{1}{1+x^2} = \frac{1}{x-z} - \operatorname{Re} \frac{1}{x-\mathrm{i}}.$$

Proof. Consider the function

$$G(\zeta) = -\mathrm{i}F\left(\mathrm{i}\frac{1-\zeta}{\zeta+1}\right).$$

This is an analytic map from the disk into the right half plane. By the Herglotz-Riesz theorem

$$F\left(\mathrm{i}\frac{1-\zeta}{\zeta+1}\right) = \mathrm{i} \int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}\theta} + \zeta}{\mathrm{e}^{\mathrm{i}\theta} - \zeta} \mathrm{d}\nu(\theta) + \operatorname{Re} F(\mathrm{i})$$

Now let $z = \mathrm{i}(1-\zeta)/(\zeta+1)$, so $\zeta = (1+\mathrm{i}z)/(1-\mathrm{i}z)$ and

$$F(z) = \mathrm{i} \int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}\theta}(1-\mathrm{i}z) + 1 + \mathrm{i}z}{\mathrm{e}^{\mathrm{i}\theta}(1-\mathrm{i}z) - 1 - \mathrm{i}z} \mathrm{d}\nu(\theta) + \operatorname{Re} F(\mathrm{i}).$$

Now we define a map $\phi : \partial D \setminus \{-1\} \to \mathbb{R}$ via

$$\phi(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{i}\frac{1-\mathrm{e}^{\mathrm{i}\theta}}{1+\mathrm{e}^{\mathrm{i}\theta}},$$

and let $\widetilde{\mu} = \phi\sharp\nu$, that is

$$\int f \mathrm{d}\widetilde{\mu} = \int f \circ \phi \mathrm{d}\nu$$

for functions $f \in C_0(\mathbb{R})$. Now given $g \in C(\partial D)$, $g - g(-1)1$ vanishes at -1 and may be written as

$$g - g(-1)1 = f \circ \phi$$

with $f = (g - g(-1)1) \circ \phi^{-1}$. Thus

$$\int_{\partial D} g \mathrm{d}\nu = \int (g - g(-1)1) \circ \phi^{-1} \mathrm{d}\widetilde{\mu} + g(-1)\nu(\partial D).$$

Since $\nu(\partial D) = \operatorname{Im} F(\mathrm{i})$, we conclude that

$$F(z) = \operatorname{Re} F(\mathrm{i}) + z \operatorname{Im} F(\mathrm{i}) + \int_{\mathbb{R}} \left[\mathrm{i} \frac{(1+\mathrm{i}x)(1-\mathrm{i}z) + (1-\mathrm{i}x)(1+\mathrm{i}z)}{(1+\mathrm{i}x)(1-\mathrm{i}z) - (1-\mathrm{i}x)(1+\mathrm{i}z)} - z \right] \mathrm{d}\widetilde{\mu}(x),$$

since $\phi^{-1}(x) = (1+\mathrm{i}x)/(1-\mathrm{i}x)$. After simplifying, this gives

$$F(z) = Az + \operatorname{Re} F(\mathrm{i}) + \int_{\mathbb{R}} \frac{1+xz}{x-z} \mathrm{d}\widetilde{\mu}(x),$$

with $A = \operatorname{Im} F(\mathrm{i}) - \widetilde{\mu}(\mathbb{R})$. The representation $(\star\star\star)$ follows with $\mathrm{d}\mu(x) = (1+x^2)\mathrm{d}\widetilde{\mu}(x)$.

The identity

$$A = \lim_{z\to\infty} \frac{F(z)}{z}$$

holds since

$$\frac{1}{z} \int_{\mathbb{R}} \frac{1+xz}{x-z} \mathrm{d}\widetilde{\mu}(x) \ \longrightarrow\ 0.$$

Furthermore, if $\lim_{z\to\infty} z(F(z) - Az - B)$ exists for some real number B then in particular

$$\lim_{t\to\infty} t(\operatorname{Im} F(\mathrm{i}t) - \mathrm{i}At) = \lim_{t\to\infty} t \int_{\mathbb{R}} \operatorname{Im} \frac{1+\mathrm{i}tx}{x-\mathrm{i}t} \mathrm{d}\widetilde{\mu}(x)$$

exists and is finite. The integrand on the r.h.s. is

$$\frac{t^2}{x^2+t^2} + t^2 \frac{x^2}{t^2+x^2} = \frac{t^2}{x^2+t^2}(1+x^2)$$

converges pointwise, monotonically to $(1+x^2)$. Thus μ is a finite measure, and

$$F(z) = Az + \operatorname{Re} F(\mathrm{i}) + \int_{\mathbb{R}} \frac{1}{x-z} \mathrm{d}\mu(x) - \operatorname{Re} \int_{\mathbb{R}} \frac{1}{x-\mathrm{i}} \mathrm{d}\mu(x).$$

One checks now that

$$\lim_{z\to\infty} (F(z) - Az) = \operatorname{Re} F(\mathrm{i}) - \operatorname{Re} \int_{\mathbb{R}} \frac{1}{x-\mathrm{i}} \mathrm{d}\mu(x). \quad \square$$

Presentation topics

Possible presentation topics

Please choose a presentation topic and discuss it with me by March 31. Presentations will be 20 minutes in length and given in class April 16, 18, 21, 23 and 25.

(1) §9.1 Completeness of weighted powers in $C_0(\mathbb{R})$.
(2) §9.2 Müntz Approximation Theorem.
(3) §11.2 Divergence of Fourier Series.
(4) §11.3 Approximate quadrature.
(5) §11.5 Existence of solutions to P.D.E.'s
(6) §14.3 Completely monotone functions
(7) §14.7 Theorems of Carathéodory and Bochner
(8) §16.3.2 Hilbert Transform
(9) §16.3.3 Laplace Transform
(10) §16.4 Solution operators for hyperbolic equations
(11) §16.5 Solution operator for the heat equation
(12) §22.4 Operators defined by parabolic equations

Homework II

Homework II

Due: March 31, 2008

(1) (Ex. 1, Ch.6) Show that a norm that satisfies the parallelogram law comes from an inner product.

(2) (Ex. 3, Ch. 6) Show that ℓ^2 is complete.

(3) Let X be a reflexive Banach space and $Y \subset X$ a closed subspace of X. Show that $(Y^\perp)^\perp = Y$.

(4) (Ex. 5, Ch. 8) Let X, μ be a measure space with μ a positive measure. Show that if $\mu(X) = 1$ then $\|f\|_p$ is an increasing function of p

(5) Prove Theorem 11.3 of these notes.

(6) Prove that $-\Delta|x|^{2-d} = (d-2)|S_{d-1}|\delta(x)$ in the sense of distributions on $\mathbb{R}^d$, $d \geq 3$, where $\Delta =$ Laplacian, $|S_{d-1}|$ is the area of the unit sphere $\{|x| = 1 : x \in \mathbb{R}^d\}$ and $\delta(x)$ is the Dirac delta "function."

(7) Consider the map T defined on Schwarz functions on the real line by the "principle value integral:"

$$T(\psi) \;=\; \text{P.V.} \int_{\mathbb{R}} \frac{1}{x}\psi(x)\mathrm{d}x \;=\; \lim_{\varepsilon \downarrow 0} \left[\int_\varepsilon^\infty \frac{1}{x}\psi(x)\mathrm{d}x + \int_{-\infty}^{-\varepsilon} \frac{1}{x}\psi(x)\mathrm{d}x. \right]$$

Show that T is a tempered distribution. Convolution with T is the "Hilbert transform."

(8) (Thm. 20.4 of the notes) Let X' be the dual of a Banach space. Show that $\ell \mapsto \|\ell\|$ is weak* lower semi-continuous and that

$$\|\ell\| \leq \liminf_{n\to\infty} \|\ell_n\|\,.$$

(9) Prove Prop. 20.1 of these notes.

(10) Show that the unit ball in an infinite dimensional Hilbert space contains infinitely many disjoint translates of a ball of radius $\sqrt{2}/4$. Conclude that there is no non-trivial translation invariant measure on an infinite dimensional Hilbert space.

(11) A subset S of a Banach space is weakly bounded if for all $\ell \in X'$, $\sup_{x\in S} |\ell(x)| < \infty$. Prove that S is weakly bounded if and only if it is strongly bounded, i.e., $\sup_{x\in S} \|x\| < \infty$.

(12) Let X be a locally convex space.
   (a) Let U be a compact neighborhood of 0. Show that one can find $x_1, \ldots, x_n$ so that $U \subset \cup_{i=1}^n (x_i + \frac{1}{2}U)$, and thus a finite dimensional linear space $Y \subset X$ with $U \subset Y + \frac{1}{2}U$.
   (b) Prove that $U \subset Y + (\frac{1}{2})^m U$ for any m.
   (c) Prove that $U \subset \overline{Y} = Y$.
   (d) Conclude that $X = \overline{Y} = Y$ is finite dimensional.

   Thus, any locally compact, locally convex, space is finite dimensional.

(13) Let X be a reflexive Banach space. Is the open unit ball in X open in the weak topology?

Part 6

Convexity

LECTURE 22

Convex sets in a Banach space

Reading: §8.4 and Ch. 12 of Lax

Definition 22.1. The support function $S_M : X' \to \mathbb{R}$ of a subset M of a Banach space X over $\mathbb{R}$ is the function

$$S_M(\ell) = \sup_{y\in M} \ell(y).$$

The support function S_M of a set M has the following properties:

(1) Subadditivity, $S_M(\ell + m) \le S_M(\ell) + S_M(m)$.
(2) $S_M(0) = 0$.
(3) Positive homogeneity, $S_M(a\ell) = aS_M(\ell)$ for $a > 0$.
(4) Monotonicity: for $M \subset N$, $S_M(\ell) \le S_N(\ell)$.
(5) Additivity: $S_{M+N} = S_M + S_N$. (Recall that $M + N = \{x + y | x \in M \text{ and } y \in N\}$.)
(6) $S_{-M}(\ell) = S_M(-\ell)$
(7) $S_{\overline{M}} = S_M$.

The proof of these is left as an exercise.

In addition, we have

(9) $S_{\breve{M}} = S_M$,

where

Definition 22.2. The closed convex hull of a subset M of a Banach space, denoted $\breve{M}$ is the smallest closed convex set containing M.

Remark. $\breve{M}$ is also the closure of the convex hull $\hat{M}$, where the convex hull is the smallest convex set containing M. (Exercise)

Let us prove property (9), assuming the other properties. First by (8) it suffices to show $S_{\hat{M}} = S_M$. Since $M \subset \hat{M}$, by (5) we have $S_M \le S_{\hat{M}}$. However,

$$\hat{M} = \{\sum_{j=1}^{n} a_j x_j \ : \ x_j \in M, \ a_j > 0, \ \text{ and } \sum_j a_j = 1\}.$$

(Exercise: prove that this is in fact the smallest convex set containing M.) So for any point in $\hat{M}$ we have

$$\ell(\sum_{j=1}^{n} a_j x_j) = \sum_{j=1}^{n} a_j \ell(x_j) \le S_M(\ell).$$

Thus $S_{\hat{M}} \le S_M$.

Here are some examples:

i. If $M = \{x_0\}$, S_M is just evaluation at x_0.
ii. If $M = B_R(0)$ then $S_M(\ell) = R\,\|\ell\|$.

iii. If $M = B_R(x_0)$ then $M = \{x_0\} + B_R(0)$ so $S_M(\ell) = R\,\|\ell\| + \ell(x_0)$.
iv. If M is a closed subspace then $S_M(\ell) = 0$ for $\ell \in M^\perp$ and ∞ for all other ℓ.

Note that in the last example the set M is unbounded. For bounded sets, $S_M : X' \to \mathbb{R}$, however in general we define S_M as a map from $X' \to \mathbb{R} \cup \{\infty\}$. We extend the order relation and arithmetic to $\mathbb{R} \cup \{\infty\}$ by $x \le \infty$ and $x + \infty = \infty$ for all x and $a\infty = \infty$ for $a > 0$. This extended function satisfies all of the above properties.

If M is bounded, then $S_M(\ell) \le \text{const.}\,\|\ell\|$ and is therefore continuous in the norm topology, since by sub-additivity

$$|S_M(\ell) - S_M(\ell')| \ \le\ \max\{S_M(\ell - \ell'), S_M(\ell' - \ell)\} \ \le\ \text{const.}\,\|\ell - \ell'\|\,.$$

This fails if M is unbounded, and also in the weak* topology. Nonetheless, we have the following additional property

(10) S_M is weak* lower semi-continuous

Indeed, since it is a sup of weak* continuous functions, we have

$$\{\ell \ :\ S_K(\ell) > t\} \ =\ \bigcup_{z \in K} \{\ell \ :\ \ell(z) > t\}\,.$$

(Weak* continuity of a $\mathbb{R} \cup \{\infty\}$ valued function is defined in the same way as for a $\mathbb{R}$ valued function.)

Theorem 22.1. *The closed convex hull $\breve{M}$ of a subset M of a Banach space X over $\mathbb{R}$ is equal to*

$$\breve{M} = \{z \ :\ \ell(z) \le S_M(\ell) \textit{ for all } \ell\}.$$

Proof. Since $S_M = S_{\breve{M}}$ it follows that $\ell(z) \le S_M(\ell)$ for all $z \in \breve{M}$.

Now, suppose $z \notin \breve{M}$. Since $\breve{M}$ is closed there is an open ball $B_R(z)$ with $B_R(z) \cap \breve{M} = \emptyset$. By the geometric Hahn-Banach theorem we can find a linear functional ℓ_0 and $c \in \mathbb{R}$ such that

$$\ell_0(u) \le c < \ell_0(v) \quad \text{ for all } u \in \breve{M} \text{ and } v \in B_R(z).$$

In particular, if $\|x\| \le R$ then

$$\ell_0(-x) = -\ell_0(x + z) + \ell_0(z) \le \ell(z) - c\ ,$$

so $\|\ell_0\| \le R^{-1}(\ell_0(z) - c)$. Thus ℓ_0 is bounded. From the definition of the norm of a linear functional, we have

$$\inf_{\|x\| < R} \ell_0(x) \ =\ -R\,\|\ell_0\|\,.$$

Since for $z + x \in B_R(z)$ we have $c \le \ell_0(z + x)$, we find that

$$\ell_0(z) \ge c + R\,\|\ell_0\|\,.$$

It follows that

$$\ell_0(z) \ge S_M(\ell_0) + R\,\|\ell_0\|\,.$$

Thus ℓ_0 is a linear functional such that $\ell_0(x) > S_M(\ell_0)$. □

The theorem shows that a closed, convex set K is specified as the set

$$K = \{z \ :\ \Phi_K(z) \le 0\}$$

where

$$\Phi_K(z) = \sup_{\ell : \|\ell\| \le 1} [\ell(z) - S_K(\ell)].$$

Since $S_K : X' \to \mathbb{R} \cup \{\infty\}$, the function Φ_K is initially defined as a map $X \to \mathbb{R} \cup \{-\infty\}$. However, note that $\Phi(z) = -\infty$ for some z if and only if $S_K(\ell) = \infty$ for all ℓ, in which case $\Phi(z) = -\infty$ for all z and $K = X$. For any proper closed, convex set K there is some ℓ such that $S_K(\ell) < \infty$ and $\Phi_K : X \to \mathbb{R}$.

Since S_K is weak* lower semi-continuous, it follows that, for fixed z, $\ell(z) - S_K(\ell)$ is weak* upper semi-continuous, that is for each t

$$\{\ell \ : \ \ell(z) - S_K(\ell) < t\}$$

is weak* open. This observation is relevant, since $\{\|\ell\| \leq 1\}$ is compact, and

Proposition 22.2. *Let K be a compact topological space and let $F : K \to \mathbb{R} \cup \{-\infty\}$ be upper semi-continuous. Then F is bounded from above and attains it's maximum.*

Remark. We did this for lower semi-continuous functions, but it didn't make it into the notes, so let's prove it again.

Proof. The sets $\{F(x) < t\}$ are increasing, open, and cover K. By compactness $K \subset \{F(x) < t\}$ for some t. So F is bounded from above. Now let $t_m = \sup_{x \in K} F(x)$. Suppose $F(x) < t_m$ for all x. Then the sets $\{F(x) < t\}$ for $t < t_m$ cover X. By compactness there is then some $t < t_m$ such that $K \subset \{F(x) < t\}$, contradicting $t_m = \sup_{x \in K} F(x)$. So there is a point x_m such that $t_m = F(x_m)$. $\square$

It follows that,

$$\Phi_K(z) = \max_{\|\ell\| \leq 1} [\ell(z) - S_K(\ell)].$$

The function $\Phi_K(z)$, being a sup of weakly continuous functions, is in turn weakly lower semi-continuous. In particular,

$$K = \{z \ : \ \Phi_K(z) \leq 0\}$$

is weakly closed! Thus we have obtained the following theorem:

Theorem 22.3 (Theorem 2, §12 of Lax). *A convex set K of a Banach space is closed in the norm topology if and only if it is closed in the weak topology.*

This theorem is astounding, since there are certainly strongly closed sets that are not weakly closed. (Weakly closed $\Longrightarrow$ strongly closed for any set.) For instance the complement of an open ball $\{x \ : \ \|x\| \geq 1\}$ is norm closed but not weakly closed. (Exercise: prove this.)

LECTURE 23

Convex sets in a Banach space (II)

Reading: §8.4

Recall that we showed last time that a closed convex set K in a Banach space is equal to

$$K = \{x \ : \ \Phi_K(x) \le 0\} ,$$

where

$$\Phi_K(x) = \max_{\|\ell\|\le 1}[\ell(x) - S_K(\ell)], \quad S_K(\ell) = \sup_{z\in K} \ell(z).$$

The weak lower semi-continuity of Φ_K showed that K is weakly closed. As a corollary we have

Corollary 23.1. *If X is reflexive, then a bounded, norm closed, convex subset K is weakly compact.*

Remark. This may fail if X is not reflexive. For instance in $L^1(\mathbb{R})$ the set K of non-negative functions $\rho \ge 0$ with integral $\int \rho = 1$ is convex, norm closed and bounded. However, it is not weakly compact.

Similarly, this may fail if X is a dual and we take the weak* topology on X. For instance, inside $M_0(\mathbb{R})$ – the space of finite measures on $\mathbb{R}$ – the space of probability measures is norm closed, bounded, and convex but is not weak* closed.

All of the suggests that $\Phi_K(x)$ might be a decent measure of how far a point x is from K. In fact, it is precisely the distance of x to K!

Theorem 23.2. *Let K be a closed, convex subset of a Banach space X. Then*

$$\Phi_K(x) \ = \ \inf_{u\in K} \|x-u\| .$$

Proof. Suppose $x \in K$. Then $\ell(x) - S_K(\ell) \le 0$ for all ℓ so the maximum is attained at $\ell = 0$ and $\Phi_K(x) = 0$.

If $x \notin K$ and $u \in K$ and $\|\ell\| \le 1$, then

$$S_K(\ell) \ge \ell(u) = \ell(u-x) + \ell(x) \ge \ell(x) - \|u-x\| .$$

Thus

$$\|u-x\| \ge \sup_{\|\ell\|\le 1} [\ell(x) - S_K(\ell)].$$

On the other hand, if $R < \inf_{u\in K} \|u-x\|$, then the convex set $K+B_R(0)$ still has positive distance from x. Thus by the Theorem of last lecture there is $\ell_0 \in X'$ such that

$$S_K(\ell_0) + R\,\|\ell_0\| = S_{K+B_R(0)}(\ell_0) < \ell_0(x).$$

This inequality is homogeneous under positive scaling, so we may take $\|\ell_0\| = 1$ to conclude

$$R < \ell_0(x) - S_K(\ell_0) \le \sup_{\|\ell\|\le 1} [\ell(x) - S_K(\ell)].$$

As R was any number less than $\inf_{u\in K} \|u-x\|$ the reverse inequality follows. ☐

Let us turn all of this around. Suppose we are given a function $S : X' \to \mathbb{R} \cup \{\infty\}$. Consider the set

$$K = \{z \in X \ : \ \ell(z) \le S(\ell) \text{ for all } \ell \in X'\}.$$

Then K is clearly convex and weakly closed, and it's support function

$$S_K(\ell) = \sup_{x \in K} \ell(z) \le S(\ell).$$

Can it happen that the inequality is strict? Of course it can, as the function S is arbitrary. However, if we assume that S is, like S_K, positive homogeneous, sub-additive, maps 0 to 0 and is weak* lower semi-continuous then the answer is "no," at least if X is reflexive.

THEOREM 23.3. *Let X be a reflexive Banach space and let $S : X' \to \mathbb{R} \cup \{\infty\}$ be a weak* lower semi-continuous function which is positive homogeneous, sub-additive, and maps 0 to 0. Then*

$$S(\ell) = \sup_{x \in K} \ell(x),$$

where

$$K = \{x \ : \ \ell(x) \le S(\ell) \textit{ for all } \ell\}$$

.

PROOF. To begin, let us prove the theorem under the additional restriction that S is bounded: $|S(\ell)| \le \beta \|\ell\|$. Then K is clearly bounded, since $x \in K \implies \|x\| \le \beta$. Since $S(0) = 0$ the identity holds for $\ell = 0$. So fix a non-zero linear functional ℓ_0. Let us define a linear functional $L \in X''$, the double dual, via Hahn-Banach. Begin on the one-dimensional subspace span$\{\ell_0\}$ and let

$$L(t\ell_0) = tS(\ell_0).$$

By positive homogeneity and sub-additivity $L(t\ell_0) \le S(t\ell_0)$ for all $t \in \mathbb{R}$. (Note that $S(-\ell_0) \ge -S(\ell_0)$.) Thus by Hahn-Banach there is a linear functional L on X' which satisfies

$$L(\ell) \le S(\ell)$$

for every $\ell \in X'$. Since $S(\ell) \le \beta \|\ell\|$ this functional is bounded. As X is reflexive there is $z \in X$ such that $L(\ell) = \ell(z)$. Since $\ell(z) \le S(\ell)$ for all ℓ this point $z \in K$, and since

$$S(\ell_0) = L(\ell_0) = \ell_0(z),$$

we have

$$S(\ell_0) = \max_{x \in K} \ell_0(x).$$

As ℓ_0 was arbitrary, this completes the proof for S bounded.

To extend this to unbounded S, note that

$$-\infty < \inf_{\|\ell\| \le 1} S(\ell) \le 0,$$

by wk* lower semi-continuity. Let $-\beta = \inf_{\|\ell\| \le 1} S(\ell)$. Then by positive homogeneity,

$$S(\ell) \ge -\beta \|\ell\|,$$

that is S is bounded from below.

Now, for each ϵ define

$$K_\epsilon = \{x \ : \ \ell(x) \le S_\epsilon(\ell) \text{ for all } \ell\},$$

where

$$S_\epsilon(\ell) = \inf_{\ell_1,\ell_2\in X' \,:\, \ell_1+\ell_2=\ell} \left[S(\ell_1) + \frac{1}{\epsilon}\left\|\ell_2\right\| \right].$$

It is left as an exercise to show, for each $\epsilon > 0$, that S_ϵ is positive homogeneous, sub-additive and maps 0 to 0. Note also that

$$-\beta\left\|\ell\right\| \leq S_\epsilon(\ell) \leq \frac{1}{\epsilon}\left\|\ell\right\|, \quad \text{and } S_\epsilon(\ell) \leq S(\ell),$$

so S_ϵ is bounded and smaller than S.

Now, in fact,

$$K_\epsilon = K \cap B_{\frac{1}{\epsilon}}(0).$$

Indeed, given $x \in K \cap B_{1/\epsilon}(0)$ we have

$$\ell(x) = \ell_1(x) + \ell_2(x) \leq S(\ell_1) + \frac{1}{\epsilon}\left\|\ell_2\right\|$$

if $\ell = \ell_1 + \ell_2$, so $\ell(x) \leq S_\epsilon(\ell)$ and $x \in K_\epsilon$. On the other hand if $x \in K_\epsilon$ then $\ell(x) \leq S(\ell)$ and $\frac{1}{\epsilon}\left\|\ell\right\|$ for all ℓ so $x \in K \cap B_{1/\epsilon}(0)$.

It follows that

$$K = \cup_\epsilon K_\epsilon.,$$

so

$$\sup_{x\in K} \ell(x) \;=\; \sup_\epsilon \sup_{x\in K_\epsilon} \ell(x) = \sup_\epsilon S_\epsilon(\ell).$$

Thus, it suffices to show, for fixed ℓ, that

$$S(\ell) = \sup_\epsilon S_\epsilon(\ell).$$

To show this, note that S_ϵ increases as ϵ decreases, so

$$S_0(\ell) := \lim_{\epsilon\to 0} S_\epsilon(\ell) = \sup_{\epsilon>0} S_\epsilon(\ell)$$

exists and (since $S_\epsilon \leq S$) satisfies

$$0 \leq S_0(\ell) \leq S(\ell).$$

Furthermore, for each ϵ we can find ℓ_ϵ such that

$$S_\epsilon(\ell) \leq S(\ell - \ell_\epsilon) + \frac{1}{\epsilon}\left\|\ell_\epsilon\right\| \leq S_\epsilon(\ell) + \epsilon.$$

Since $S(\ell - \ell_\epsilon) \geq -\beta\left\|\ell\right\| - \beta\left\|\ell_\epsilon\right\|$ we see that

$$-\beta\left\|\ell\right\| + \left(\frac{1}{\epsilon} - \beta\right)\left\|\ell_\epsilon\right\| \leq S_\epsilon(\ell) + \epsilon.$$

Consider the following cases: (1) $\left\|\ell_\epsilon\right\|/\epsilon$ is bounded as $\epsilon \to 0$ or (2) $\left\|\ell_\epsilon\right\|/\epsilon$ is unbounded as $\epsilon \to 0$. In case (2) $\lim_{\epsilon\to 0} S_\epsilon(\ell) = S_0(\ell) = S(\ell) = \infty$. On the other hand, in case (1) $\ell_\epsilon \to 0$ so by weak* lower semi-continuity of S we find that

$$S(\ell) \leq \liminf_{\epsilon\to 0} S(\ell - \ell_\epsilon) \leq S_0(\ell) - \liminf_{\epsilon\to 0} \frac{1}{\epsilon}\left\|\ell_\epsilon\right\| \leq S_0(\ell),$$

which completes the proof. $\square$

LECTURE 24

Krein-Milman and Stone-Weierstrass

Reading: §13.3

DEFINITION 24.1. An extreme subset S of a convex set K is a subset $S \subset K$ such that

(1) S is non-empty and convex
(2) If $x \in S$ and $x = ty + (1-t)z$ with $y, z \in K$ then $y, z \in S$.

An extreme point is a point $x \in K$ such that $\{x\}$ is an extreme subset.

The following is a classical result due to Carathéodory:

THEOREM 24.1. *Every compact convex subset K of $\mathbb{R}^N$ has extreme points, and every point of K can be written as a convex combination of (at most) $N+1$ extreme points.*

REMARK. The proof is left as an exercise. Use induction on N. The case $N = 1$ is easy!

THEOREM 24.2 (Krein and Milman). *Let X be a locally convex space, K a non-empty, compact, convex subset of X. Then*

(1) K has at least one extreme point
(2) K is the closure of the convex hull of its extreme points.

PROOF. Consider the collection $\mathcal{E}$ of all nonempty closed extreme subsets of K. Since $K \in \mathcal{E}$ this collection is nonempty. Partially order $\mathcal{E}$ by inclusion. We wish to apply Zorn's lemma to see that $\mathcal{E}$ has a "maximal" element, i.e., a set that is minimal with respect to conclusion.

Let $\mathcal{T} \subset \mathcal{E}$ be a totally ordered sub-collection. That is $\mathcal{T} = \{E_\omega : \omega \in \Omega\}$ with Ω some totally ordered index set and $E_\alpha \subset E_\beta$ if $\alpha \geq \beta$. Clearly $\cap\mathcal{T}$ is a candidate for an "upper bound." To see this we must show that $\cap\mathcal{T}$ is nonempty, closed, and extreme.

Clearly $\cap\mathcal{T}$ is closed. Furthermore,one easily shows that the intersection of an arbitrary family of extreme sets is extreme provided it is non-empty. Thus we need only show $\mathcal{T}$ is non-empty. Here we use compactness of K in a crucial way.

Recall that a family $\mathcal{F}$ of closed sets is said to have the finite intersection property (FIP) if any finite collection $F_1, \ldots F_n \in \mathcal{F}$ has non-empty intersection: $F_1 \cap \cdots \cap F_n \neq \emptyset$, and there is the following result connecting the FIP and compactness:

THEOREM 24.3. *A topological space M is compact if and only if every collection $\mathcal{F}$ of closed sets with the FIP satisfies*

$$\cap\mathcal{F} \neq \emptyset.$$

PROOF. Suppose M is compact and $\cap\mathcal{F} = \emptyset$. Then $\cup\mathcal{U} = M$ with $\mathcal{M} = \{F^c : F \in \mathcal{F}\}$. Thus $M = \cup_{j=1}^n F_j^c$ for some finite collection. Thus $\cap_{j=1}^n F_j = \emptyset$ and $\mathcal{F}$ does not have the FIP.

Conversely, if $\mathcal{F}$ is a collection of closed sets with the FIP and nonetheless $\cap\mathcal{F} = \emptyset$, then $\mathcal{U} = \{F^c : F \in \{\}$ is an open cover of M with no finite subcover so M is not compact. $\square$

Regarding our collection $\mathcal{T}$, it clearly has the FIP since it is totally ordered, so any finite collection $E_1, \dots E_n \in \mathcal{T}$ has a minimal element. Thus $\cap\mathcal{T} \neq \emptyset$.

Sp $\mathcal{E}$ has minimal elements. We claim that any such minimal element is a one point set. Indeed, suppose $E \subset \mathcal{E}$ contains two distinct points (it must also contain the line segment joining them). Then there is a continuous linear functional ℓ on X that separates these points. Let $M \subset E$ be the set

$$M = \{x \in E \;:\; \ell(x) = \max_{z \in E} \ell(x)\}.$$

Then M is a non-empty, proper, closed subset of E. It is clearly convex and one easily shows it is an extreme subset of E. It follows that M is an extreme subset of K (why?) and $M \subsetneq E$ so E is not minimal.

This proves (1): K has at least one extreme point. (It might have only one: K could be the set {x}.) Since any closed extreme subset E of K is itself a closed convex set we find that every extreme subset of E has an extreme point x. It is easy to see that an extreme point of E is also an extreme point of K (since E is an extreme subset). Thus

Every closed, extreme subset of K contains an extreme point of K.

Let K_e denote the set of extreme points, $\widehat{K}_e$ its convex hull, and $\breve{K}_e$ its closed convex hull which is the closure of $\widehat{K}_e$. One easily shows that $\breve{K}_e$ is convex. (Exercise)

Clearly $\widehat{K}_e \subset K$, so since K is closed $\breve{K}_e \subset K$. On the other hand if $z \notin \breve{K}_e$ then there is an open set U with $z \in U \subset K_e$. We may take U to be convex. By the geometric Hahn-Banach there is a linear functional ℓ and $c \in \mathbb{R}$ such that

$$\ell(x) < c \leq \ell(y) \quad \text{ for all } x \in U \text{ and } y \in \breve{K}_e.$$

(As U is open, all points of U are interior, so the first inequality is strict.) The gauge function of $U - z$,

$$p_{U-z}(x) = \inf\{t \;:\; x/t \in U - z\}.$$

is a continuous seminorm on X (why?). If $x/t \in U - z$ we have

$$\frac{1}{t}\ell(x) = \ell(\frac{x}{t} + z) - \ell(z) < c - \ell(z).$$

Thus

$$\ell(x) \leq (c - \ell(z))p_{U-z}(x),$$

and ℓ is a continuous linear functional because

LEMMA 24.4. *A linear functional on locally convex space is continuous if and only if it is bounded with respect to some continuous seminorm.*

PROOF. Exercise. □

Since K is compact ℓ achieves its minimum on K. Let E be the set of minimizers. Then E is closed, convex and extreme. (Why?) By the above derived result E contains an extreme point. Thus

$$\min_{x \in K} \ell(x) = \min_{x \in K_e} \ell(x) > \ell(z).$$

Thus $z \notin K$. □

An interesting application of this theorem is:

THEOREM 24.5 (Stone-Wierstrass). *Let S be a compact Hausdorff space and $C(S)$ the set of real valued continuous functions on S. Let $E \subset C(S)$ be an sub-algebra, that is E is linear subspace and $f, g \in E \implies fg \in E$. If the constant function $1 \in E$ and if for any pair of points $p, q \in S$ there is a function $f \in E$ with $f(p) \neq f(q)$, then E is dense in $C(S)$ in the max norm.*

This theorem is due to Stone and generalizes the classical result of Weierstrass on approximation of continuous functions on an interval with polynomials: $S = [0,1]$ and $E =$ polynomials.

SKETCH OF PROOF. Consider the collection $\mathcal{N}$ of all finite measures μ on S such that $\int f \mathrm{d}\mu = 0$ for all $f \in E$. Then S is dense if and only if $\mathcal{M} = \{0\}$. (Why?)

By construction $\mathcal{N}$ is weak* closed. So $K = \mathcal{N} \cap B_1(0)$ is weak* compact. Furthermore it is convex. Suppose K contains a non-zero measure μ. Then K must contain a non-zero extreme point μ. Since μ is extreme we must have $\|\mu\| = 1$ (otherwise we could write μ as a linear combination of 0 and some multiple of μ).

Suppose such a μ exists. Since E is an algebra $\int fg\mathrm{d}\mu = 0$ for all $f, g \in E$. Thus $g\mathrm{d}\mu \in \mathcal{N}$ for all $g \in E$. Now let g be a continuous function on S with $0 < g(p) < 1$ for all $p \in S$. Let

$$a = \int g\mathrm{d}\,|\mu| \ , \quad b = \int (1-g)\mathrm{d}\,|\mu| \,.$$

So $a, b > 0$ and $a + b = 1$ and $g\mathrm{d}\mu/a$, $(1-g)\mathrm{d}\mu/b \in K$. Since

$$\mathrm{d}\mu = a\frac{g}{a}\mathrm{d}\mu + b\frac{(1-g)}{b}\mathrm{d}\mu$$

we must have $g\mathrm{d}\mu/a = \mathrm{d}\mu$ (recall that μ is an extreme point).

Consider the support of μ:

$$\operatorname{supp}\mu \;=\; \{p \;:\; |\mu|(U) > 0 \text{ for any open neighborhood of } p\}.$$

Since $\mathrm{d}\mu = g\mathrm{d}\mu/a$ for any $g \in E$ that satisfies $0 < g(p) < 1$ we must have $g \equiv a$ on $\operatorname{supp}\mu$. (Why?) Suppose p and q are distinct points in S. Then there is a function $g \in E$ such that $0 < g < 1$ and $g(p) \neq g(q)$. (Just add a large constant to a function in E that separates p and q). Thus at most one of the points p, q lies in the support of μ. That is the support of μ is a single point $\operatorname{supp}\mu\{p_0\}$! Since $|\mu|(1) = \|\mu\| = 1$ we have

$$\int f(p)\mathrm{d}\mu(p) \;=\; f(p_0) \quad \text{or} \quad \int f(p)\mathrm{d}\mu(p) \;=\; -f(p_0).$$

However, we have $\int 1\mathrm{d}\mu = 0$ which is a contradiction, so $\mathcal{N} = \{0\}$ and E is dense. $\square$

Following the above proof, we find

THEOREM 24.6. *The extreme points of the unit ball $\{\mu \;:\; \int \mathrm{d}|\mu| \leq 1\} \subset \mathcal{M}(S)$ are the point masses $\pm\delta(p - p_0)$.*

THEOREM 24.7. *If $A \subset C(S)$ is a proper closed sub-algebra that separates points of S then $A = \{f : f(p_0) = 0\}$ for some $p_0 \in S$.*

The algebras $A_p = \{f \;:\; f(p) = 0\}$ are exactly the maximal ideals of $C(S)$. It is no accident that the set of maximal ideals is the set of proper sub-algebras that separate points and is in one to one correspondence with extreme points of the unit ball in $C(S)'$ which is in one to one correspondence with S.

LECTURE 25

Choquet type theorems

Reading: §13.4 and 14.10

In finite dimensions, Caratheodory's Theorem shows that points of a compact set may be expressed as a convex combination of extreme points: no more than $N+1$ points are needed in dimension N. (A simplex shows that this number is optimal.) The following generalizes this idea to LCS's:

THEOREM 25.1. *Let X be an LCS and K a non-empty compact, convex subset of X. For any $u \in K$ there is a Borel probability measure μ_u on $\overline{K_e}$ the closure of the set of extreme points such that*

$$\ell(u) \;=\; \int_{\overline{K_e}} \ell(x) \mathrm{d}\mu_u(x) \tag{$\star$}$$

for all $\ell \in X'$.

REMARK. The integral in the theorem is understood as the identity

$$u = \int_{\overline{K_e}} x \mathrm{d}\mu_u(x),$$

"in the weak sense." This expresses u as a generalized convex combination of points of $\overline{K_e}$.

Lax presents without proof a sharper theorem , due to Choquet, in which the measure μ_u is defined on K_e provided K is metrizable. Any representation of type ($\star$) is called a "Choquet decomposition." Such representations need not be unique, even in finite dimensions, for example if K is a disk in $\mathbb{R}^2$.

PROOF. The general idea of the proof is to show that given a function f on K_e we can "evaluate it" at a pont $u \in K$ and produce a bounded linear functional: $f \mapsto f(u)$. We then represent this linear functional as a measure on K_e via Riesz-Kakutani.

There are two complications: (1) not every function f on K_e can be extended in a reasonable, or unique way to $u \in K$ and (2) Riesz-Kakutani does not apply since K_e may not be compact. The second complication is easily dealt with by replacing K_e with it's closure $\overline{K_e}$ which, being a closed subset of a compact space (K) is compact.

As for the first: any linear functional $\ell \in X'$ can be evaluated at u. Furthermore, being continuous and linear, ℓ achieves its max and min over K on the set of extreme points K_e. (We saw this in the proof of Krein-Millman.) That is

$$\min_{x \in \overline{K_e}} \ell(x) \leq \ell(u) \leq \max_{x \in \overline{K_e}} \ell(x) \quad \text{for all } u \in K.$$

It follows that if $\ell_1(x) = \ell_2(x)$ for $x \in K_e$ then $\ell_1 = \ell_2$ on K, so any linear functional is determined on K by it's restriction to K_e. Let $Y \subset C(\overline{K_e})$ be the subspace of maps $f : \overline{K_e} \to \mathbb{R}$ with

$$f(x) = \ell(x)$$

for some $\ell \in X'$, or $f(x) = c$ for some $c \in \mathbb{R}$. Define a linear functional L on Y by

$$L(f) \;=\; \begin{cases} \ell(u) & f(x) = \ell(x) \\ c & f(x) = c. \end{cases}$$

Then L is a positive linear functional on Y and Y contains the constant functions. By Theorem 3.1 of these notes, we may extend L as a positive linear functional on the space of all functions. Restricting this linear functional to $C(\overline{K_e})$, we obtain by Riesz-Kakutani that

$$L(f) = \int_{\overline{K_e}} f(x)\mathrm{d}\mu_u(x) \quad \text{for all } f \in C(\overline{K_e})$$

for some μ_u. In particular $(\star)$ holds. Since $\mu_u(1) = L(1) = 1$ μ_u is a probability measure. $\square$

The Riesz-Kakutani theorem, which was used in the above proof, is it-self an example of this theorem. Indeed, let $K \subset C(S)'$ be the closed unit ball of the dual of $C(S)$ with S a compact Hausdorff space. We saw last time that the extreme points K_e are point evaluations:

$$\ell_{p,\pm 1}(f) \;=\; \pm 1 \times f(p)$$

for some $p \in S$. Thus K_e is in one to one correspondence with $S \times \{-1,+1\}$. Exercise: show that $K_e = S \times \{-1,+1\}$ as topological spaces, where K_e has the weak* topology and $S \times \{-1,+1\}$ has the product topology using the discrete topology on $\{-1,+1\}$ (all sets – all 4 of them – are open). The above theorem shows that any linear functional $\ell \in K$, with norm no larger than 1, can be written as a combination

$$\ell(f) = \int_{S\times\{\pm 1\}} \ell_{p,\sigma}(f)\mathrm{d}m(p,\sigma) = \int_S f(p)\mathrm{d}m(p,+) - f(p)\mathrm{d}m(p,-1) = \int_S f(p)\mathrm{d}\mu(p)$$

with $\mathrm{d}\mu(p) = \mathrm{d}m(p,+) - \mathrm{d}m(p,-1)$.

Measure preserving maps

Let Ω be a compact metric space and let $T : \Omega \to \Omega$ be a homeomorphism. Consider the set K of all probability measures on Ω that are invariant with respect to T:

$$K = \{\mu \in \mathcal{M}(\Omega) \;:\; \mu \geq 0,\ \mu(\Omega) = 1,\ \text{and } \mu(S) = \mu(T(S)) \text{ for all measurable } S\}.$$

We say that μ is ergodic under T provided $T(S) \subset S \implies \mu(S) = 0$ or 1.

The following is one of the most important applications of Choquet type theorems.

Theorem 25.2. *The set K is non-empty, convex and compact in the weak* topology. The extreme points of K (which exist by Krein Millman) are the ergodic measures. Every invariant measure can be represented uniquely as an average of ergodic measures.*

Sketch of proof. Convexity of K is easy. Since K is clearly a subset of the unit ball in $\mathcal{M}(\Omega)$ it is compact once it is closed. Weak* closure follows since

$$\mu \in K \iff \int_\Omega f(T^{-1}(\omega))\mathrm{d}\mu(\omega) = \int_\Omega f(\omega)\mathrm{d}\mu(\omega)$$

for all $f \in C(\Omega)$. To see that K is non-empty, pick a point $\omega_0 \in \Omega$ and consider the sequence of measures ν_n given by

$$\int f(\omega)\mathrm{d}\nu_n(\omega) \;=\; \frac{1}{n}\sum_{j=0}^{n} f(T^{-j}(\omega_0)).$$

This is a sequence of probability measures on Ω. Since the unit ball is weak* compact there is a weak* convergent subsequence. The limit $\overline{\nu}$ of this subsequence satisfies $\nu \geq 0$, $\nu(\Omega) = 1$, and

$$\int f(T^{-1}(\omega))\mathrm{d}\overline{\nu}(\omega) \;=\; \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k} f(T^{-j-1}(\omega_0))$$

$$= \lim_{k\to\infty} \frac{1}{n_k} \sum_{j=0}^{n_k} f(T^{-j}(\omega_0) - \frac{1}{n_k} f(\omega_0) + \frac{1}{n_k} f(T^{-n_k-1}(\omega_0)) \;=\; \int f(\omega)\mathrm{d}\overline{\nu}(\omega).$$

Suppose $\mu \in K$ is not ergodic. Then there must be a disjoint union $\Omega_1 \cup \Omega_2 = \Omega$ with $T(\Omega_j) \subset \Omega_j$ and $\mu(\Omega_j) > 0$. Let μ_j be the restriction of μ to Ω_j renormalized to be a probability measure:

$$\mu_j(S) = \frac{\mu(S \cap \Omega_j)}{\mu(\Omega_j)}.$$

Then each μ_j is invariant under T and

$$\mu = \mu(\Omega_1)\mu_1 + \mu(\Omega_2)\mu_2$$

so μ is not extreme.

Conversely, suppose μ is not extreme, so

$$\mu = am_1 + (1-a)m_2$$

with $m_1 \neq m_2 \in K$ and $0 < a < 1$. Suppose $m_2 << m_1$. Then by Radon-Nikodym there is $f \in L^1(m_1)$, $f \not\equiv 1$, such that

$$\mathrm{d}\mu = (a + (1-a)f)\mathrm{d}m_1.$$

Since μ and m_1 are invariant under T we must have $f(T(\omega)) = f(\omega)$ for m_1 almost every ω. Since $f \not\equiv 1$ the sets $\Omega_1 = \{\omega : f(\omega) \leq 1\}$ and $\Omega_2 = \{\omega : f(\omega) > 1\}$ are separately invariant under T and both have positive μ measure. So μ is not ergodic.

If $m_2 \not< < m_1$ then the situation is even happier. For then there is some set S such that $m_1(S) = 0$ but $m_2(S) > 0$. Let $\Omega_1 = \cup_{j=0}^{\infty} T^j(S)$. Then $m_1(\Omega_1) = 0$, $m_2(\Omega_1) \geq m_2(S) > 0$ and Ω_1 is invariant under T. Since $m_1(\Omega_1) = 0$ we must have $\mu(\Omega_1) \leq 1 - a$ so $\mu(\Omega_2) > 0$ with $\Omega_2 = \Omega \setminus \Omega_1$. Thus μ is not ergodic.

The representation of μ as a unique integral over ergodic measures is clearly a Choquet type representation – we need not take the closure of extreme points since the space of measures is metrizable in the weak* topology as it is the dual of a separable Banach space (this was "Choquet's theorem" which we did not prove). The uniqueness requires a separate argument, for which Lax refers a paper of Oxtoby in Bull. AMS 58 (1952). $\square$

Part 7

Bounded Linear Maps

LECTURE 26

Bounded Linear Maps

Definition 26.1. Let X and Y be Banach spaces. A linear map $M : X \to Y$ is bounded if there is some $0 \leq c < \infty$ such that

$$\|Mx\|_Y \leq c \|x\|_X .$$

The smallest such c is called the norm, or operator norm, of M, denoted $\|M\|$.

Theorem 26.1. *A linear map $M : X \to Y$ is continuous if and only if it is bounded.*

Proof. Bounded implies Lipschitz continuous since

$$\operatorname{dist}(Mx, My) = \|Mx - My\| \leq \|M\| \|x - y\| = \|M\| \operatorname{dist}(x, y).$$

On the other hand if M is not bounded then there is a sequence x_n such that

$$\|Mx_n\| > n \|x_n\| .$$

As this inequality is invariant under scaling we may take $\|x_n\| = 1/\sqrt{n}$ so $x_n \to 0$ but $\|Mx_n\| \to \infty$. □

Theorem 26.2. *The operator norm has the following properties*

- *$\|aM\| = |a| \|M\|$ for all $a \in F$ (homogeneity)*
- *$\|M\| \geq 0$ and $\|M\| = 0$ if and only $M \equiv 0$ (positivity)*
- *$\|M + K\| \leq \|M\| + \|K\|$ (sub-additivity)*

That is the operator norm is a norm.

Note that the norm is defined to be

$$\|M\| = \sup_{x \neq 0} \frac{\|Mx\|}{\|x\|}.$$

By homogeneity this is the same as

$$\|M\| = \sup_{\|x\|=1} \|Mx\| .$$

If M is bounded on a normed space which is incomplete then M has an extension to the completion (as does any continuous function) which is also bounded, with the same norm.

Definition 26.2. Let $\mathcal{L}(X, Y)$ denote the set of all bounded linear maps from X to Y.

Theorem 26.3. *Let X be a normed space and let Y be a Banach space. Then $\mathcal{L}(X, Y)$ is a Banach space under the operator norm.*

Proof. Since $\|\cdot\|$ is a norm, we need to show completeness. Suppose M_n is a Cauchy sequence. For each $x \in X$ it follows that $M_n x$ is a Cauchy sequence in Y, since $\|M_n x - M_m x\| \leq$

$\|M_n - M_m\| \|x\|$. By completeness of Y there is a limit. Call this limit Mx. It is easy to see that M is linear. Moreover, if $\|x\| = 1$,

$$\|Mx\| = \lim_n \|M_n x\| \leq \limsup_{n\to\infty} \|M_n\| < \infty,$$

so $\|M\|$ is finite and M is bounded. □

The operator norm topology on $\mathcal{L}(X,Y)$ is called the uniform topology. There are two other natural topologies on this space

Definition 26.3. The strong topology on $\mathcal{L}(X,Y)$ is the weakest TVS topology such that all functions $M \mapsto Mx$ are continuous from $\mathcal{L}(X,Y) \to Y$. The weak topology (or weak operator topology) on $\mathcal{L}(X,Y)$ is the weakest TVS topology such that all maps $M \mapsto \ell(M(x))$, for $\ell \in Y'$ and $x \in X$, are continuous.

Weak and strong sequential convergence are defined similarly. A sequence M_n converges strongly if

$$s - \lim_n M_n x \quad \text{exists in } Y \text{ for every } x,$$

and converges weakly if

$$\mathrm{wk} - \lim_n M_n x \quad \text{exists in } Y \text{ for every } x.$$

Note that the weak operator topology on $\mathcal{L}(X,Y)$ is potentially much weaker than the weak topology on $\mathcal{L}(X,Y)$ as a Banach space: the linear functionals $M \mapsto \ell(M(x))$ are just one kind of linear functional on $\mathcal{L}(X,Y)$.

Definition 26.4. The transpose of a linear operator $M \in \mathcal{L}(X,Y)$ is the map $M' \in \mathcal{Y}', \mathcal{X}'$ defined by

$$\langle x, M'\ell\rangle = \langle Mx, \ell\rangle\,,$$

where we use the dual pairing notation $\langle y, \ell\rangle = \ell(y)$.

Definition 26.5. The null space of a linear operator M is the set

$$N_M = \{x \in X \ : \ Mx = 0\}.$$

The range of M is the set

$$R_M = \{Mx \ : \ x \in X\}.$$

Theorem 26.4. *Let $M \in \mathcal{L}(X,Y)$ with X and Y normed spaces. Then*

(1) M' is bounded and $\|M'\| = \|M\|$

(2) The nullspace of M' is the annihilator of the range of M:

$$N_{M'} = R_M^{\perp}.$$

(3) The nullspace of M is the annihilator of the range of M':

$$N_M = R_{M'}^{\perp} = \{x \ : \ \langle x, \ell\rangle = 0 \textit{ if } \ell \in R_{M'}\}.$$

(4) $(aM + bN)' = aM' + bN'$

Proof. Exercise, or see Lax. □

As a corollary we see that N_M and $N_{M'}$ are weak and weak* closed respectively. Since these sets are subspaces this is the same as being closed. This is the first part of

Theorem 26.5. *Let $M \in \mathcal{L}(X,Y)$ with X and Y normed spaces. Then*

(1) N_M is a closed linear subspace of X.
(2) M, regarded as a map

$$M_0 : \frac{X}{N_M} \to Y$$

is one-to-one, bounded, with $\|M_0\| = \|M\|$ and $R_{M_0} = R_M$.

PROOF. The facts that M_0 is one-to-one and $R_{M_0} = R_M$ are general facts about linear maps. To see that M_0 is bounded, recall that X/N_M is the set of equivalence classes $[x]$ with $x \sim y$ if $x - y \in N_M$ and we put a norm on this space by

$$\|[x]\| = \inf_{y \in N_M} \|x + y\| .$$

Since $M_0[x] = Mx = M(x + y)$ for any $y \in N_M$ we have

$$\|M\| = \sup_{x \neq 0} \frac{\|Mx\|}{\|x\|} = \sup_{x \neq 0} \sup_{y \in N_M} \frac{\|M(x+y)\|}{\|x+y\|} = \sup_{[x] \neq 0} \frac{\|M_0[x]\|}{\|[x]\|} = \|M_0\| .$$

□

Regarding convergence of adjoints we have the following

PROPOSITION 26.6. *If* $\mathrm{wk} - \lim M_n = M$ *then* $\mathrm{wk} - \lim M_n' = M'$.

PROOF. Exercise. □

However, this does not hold for strong limits or uniform limits. For example, on any ℓ_p space, $1 < p < \infty$, let S_j be the j^{th} forward shift, the map $S_j \in \mathcal{L}(\ell_p, \ell_p)$ given by

$$S(a_0, a_1, \ldots) = (\underbrace{0, \ldots, 0}_{j \text{ zeroes}}, a_0, a_1, \ldots).$$

Then S_j converges to zero weakly, but

$$\|S_j \mathbf{a}\|_{\ell_p} = \|\mathbf{a}\|_{\ell_p} ,$$

so S_j does not converge to zero strongly. On the other hand the adjoint S_j' is the backwards j shift,

$$S_j'(a_0, a_1, \ldots) = (a_j, a_{j+1}, \ldots).$$

This map converges to zero strongly since

$$\left\|S_j' \mathbf{a}\right\|_{\ell_{p'}} \to 0$$

for $1 < p < \infty$. (It does not converge to zero strongly for $p = 1$.)

LECTURE 27

Principle of Uniform Boundedness and Open Mapping Theorem

Reading: §15.3–15.5

As for linear functionals, we have the following useful criteria for strong/weak convergence:

Proposition 27.1. *Let $M_n \in \mathcal{L}(X,Y)$ be a sequence of bounded maps between Banach spaces X and Y. Suppose that M_n are uniformly bounded:*

$$\sup_n \|M_n\| < \infty.$$

(1) If $M_n x$ converges in norm for all x in a dense subset of X then M_n converges in the strong operator topology.
(2) If $M_n x$ converges weakly for all in a dense subset of X then M_n converges in the weak operator topology.

Proof. Exercise. □

As in the case of linear functionals, boundedness turns out to be necessary for convergence as well:

Theorem 27.2 (Principle of Uniform Boundedness). *Let X and Y be Banach spaces and let $\mathcal{M} \subset \mathcal{L}(X,Y)$ be a collection of bounded linear maps. If for each $x \in X$ and $\ell \in Y'$ there is a constant $c(x,\ell)$ such that*

$$|\langle Mx, \ell\rangle| \le c(x,\ell) \quad \textit{for all } M \in \mathcal{M},$$

then $\mathcal{M}$ is uniformly bounded, i.e., there is $c < \infty$ such that

$$\|M\| \le c \quad \textit{for all } M \in \mathcal{M}.$$

Sketch of proof. This follows very closely the proof of the PUB for linear functionals, Corollary 19.6 of these notes. First apply that result to conclude that for each $x \in X$ there is $c(x)$ with $\|Mx\| \le c(x)$ for all $M \in \mathcal{M}$. Then consider the collection of real values, continuous, positive homogeneous functions $f(x) = \|Mx\|$ on the Banach space X. Apply Theorem 19.5 to conclude that $\|Mx\| \le c\,\|x\|$ for all $M \in \mathcal{M}$. □

The PUB follows from the Baire category theorem. This theorem also implies several results which show that bounded linear maps have a number of useful properties in addition to continuity.

Theorem 27.3. *Let X and Y be Banach spaces and $M : X \to Y$ a bounded linear map of X <u>onto</u> Y. Then the image of the unit ball $MB_1(0)$ contains an open ball around the origin in Y.*

Corollary 27.4 (Open mapping theorem). *A bounded linear map M from X <u>onto</u> Y, with X and Y Banach spaces, is <u>open</u> that is $M(U)$ is open in Y for any open subset $U \subset X$.*

Proof. Let U be open in X. Then given $y \in M(U)$ we have $y = Mx$ for some $x \in U$. Since U is open $x + \epsilon B_1(0) \subset U$ for some $\epsilon > 0$. Thus $Mx + \epsilon MB_1(0) \subset MU$ so MU contains an open ball centered at $y = Mx$. □

Proof of Theorem. Since M maps X onto Y we have

$$Y = \cup_{n=1}^{\infty} MB_n(0).$$

By the Baire category theorem at least one of the sets $MB_n(0)$ is dense in some open set. That is there are $\delta > 0$ and $y \in Y$ such that $MB_n(0) \cap (y + B_\delta(0))$ is dense in $B_\delta(0)$. Since M is onto, there is $x \in X$ such that $Mx = y$. So $M(B_n(0) - x)$ is dense in $B_\delta(0)$. By the triangle inequality,

$$B_n(0) - x \subset B_{n+\|x\|}(0).$$

By homogeneity we conclude that for every $r > 0$, $MB_r(0)$ is dense in $B_{\alpha r}(0)$ with

$$\alpha = \delta/(n + \|x\|).$$

Now let $y \in B_\alpha(0) \subset Y$. Then there is $x_1 \in B_1(0)$ such that

$$\|y - Mx_1\| \le \frac{1}{2}\alpha$$

Let $y_1 = y - Mx_1$. So $y_1 \in B_{\frac{1}{2}\alpha}(0)$. Thus there is $x_2 \in B_{\frac{1}{2}}(0)$ such that $\|y_1 - Mx_2\| \le \frac{1}{4}\alpha$. That is

$$\|y - M(x_1 + x_2)\| \le \frac{1}{4}\alpha.$$

Following this procedure construct, by induction, a sequence $x_1, x_2, \ldots$ with

(1) $x_j \in B_{\frac{1}{2^{j-1}}}(0) \subset X$
(2) $\left\| y - M\sum_{j=1}^n x_j \right\| \le \frac{1}{2^j}\alpha$.

Now the partial sums $\sum_{j=1}^n x_j$ converge as $n \to \infty$ to some point $x \in X$ with

$$\|x\| \le \sum_{j=1}^{\infty} 2^{1-j} = 2.$$

Clearly $y = Mx$. Thus $B_\alpha(0) \subset MB_2(0)$ and the result follows by homogeneity. □

Theorem 27.5. *Let X and Y be Banach spaces and $M : X \to Y$ a bounded, one-to-one map of X onto Y. Then the inverse map M^{-1} is bounded from $Y \to X$.*

Note in particular that

$$\|Mx\| \ge \frac{1}{\|M^{-1}\|} \|x\| .$$

Thus bounded, one-to-one and onto $\implies$ M is bounded from below! This is somehow amazing because we were not given any quantitative information on the inverse just it's existence.

Proof. Since $MB_1(0) \supset B_d(0)$ for some $d > 0$ we have

$$M^{-1}y \in B_1(0)$$

for all $y \in B_d(0)$. By homogeneity $\|M^{-1}\| \le 1/d$. □

Definition 27.1. A map $M : X \to Y$ is <u>closed</u> if whenever $x_n \to x$ in X and $Mx_n \to y$ in Y then $Mx = y$.

THEOREM 27.6 (Closed Graph Theorem). *Let X and Y be Banach spaces and $M : X \to Y$ a closed linear map. Then M is bounded.*

PROOF. Let the graph G of M by the set of all pairs (x, Mx) with $x \in X$. By linearity of M, G is a linear space under coordinate-wise addition and scalar multiplication. Give G a norm by defining

$$\|(x, Mx)\| = \|x\| + \|Mx\| .$$

Since M is closed, the space G is complete in this norm. Now let $P : G \to X$ be the map $P(x, Mx) = x$. So P is bounded ($\|P\| \leq 1$), one-to-one and onto. It follows that $P^{-1}x = (x, Mx)$ is bounded. Thus there is a constant $c < \infty$ such that

$$\|(x, Mx)\| \leq c \|x\| .$$

So

$$\|Mx\| \leq (c - 1) \|x\| . \quad \square$$

DEFINITION 27.2. Let X be a linear space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on X. The norms $\|\cdot\|_j$, $j = 1, 2$ are called compatible if whenever a sequence converges in both norms the limits are equal.

For example, on $L^1 \cap L^2$ the L^1 and L^2 norms are compatible.

COROLLARY 27.7. *Let X be a linear space and let $\|\cdot\|_j$,$j = 1, 2$ be compatible norms. If X is complete in both norms then the two norms are equivalent: there is $c \in (0, \infty)$ such that $c^{-1} \|x\|_1 \leq \|x\|_2 \leq c \|x\|_1$.*

PROOF. Let X_j, $j = 1, 2$, be the Banach space which is X under norm $\|\cdot\|_j$, $j = 1, 2$. The identity map $I : X_1 \to X_2$ is closed by compatibility of the norms. Thus I is bounded. Likewise $I : X_2 \to X_1$ is bounded. $\square$

LECTURE 28

The Spectrum of a Linear Map

Reading: §15.5 and 20.1

First some observations about composition of linear maps. If $M : X \to Y$ and $N : Y \to Z$ are linear maps their composition $(NM)x = N(Mx)$ is a map $NM : X \to Z$.

THEOREM 28.1. *If M and N are bounded then so is NM. Furthermore*
(1) $\|NM\| \le \|N\| \, \|M\|$
(2) $(NM)' = M'N'$.

PROOF. Exercise, or see Lax. □

We now consider some general properties of linear maps from a Banach space X into itself. Let $\mathcal{L}(X) = \mathcal{L}(X, X)$. A map $M \in \mathcal{L}(X)$ is invertible if M maps X onto X and is $1-1$. We saw last time that it follows from the open mapping theorem that M^{-1} is bounded as well. Let $\mathcal{GL}(X)$ denote the set of all invertible maps in $\mathcal{L}(X)$.

PROPOSITION 28.2. *If $K, L \in \mathcal{GL}(X)$ then so is KL and $(KL)^{-1} = L^{-1}K^{-1}$.*

THEOREM 28.3. *$\mathcal{GL}(X)$ is an open set in the uniform topology.*

PROOF. We must show that if K is invertible then so is every map of the form $K + A$ with $\|A\| < \epsilon$ for some $\epsilon > 0$. Since $K + A = K(1 + K^{-1}A)$ and $\|K^{-1}A\| \le \|K^{-1}\| \, \|A\|$ it suffices to prove this for $K = \mathbb{1}$ the identity map.

The basic idea is to make use of the geometric series to write $(\mathbb{1} + A)^{-1}$ as

$$\sum_{n=0}^{\infty} (-1)^n A^n.$$

Since

$$\|A^n\| \le \|A\|^n$$

this series converges once $\|A\| < 1$. But then

$$(\mathbb{1} + A) \sum_{n=0}^{N} (-1)^n A^n = 1 + (-1)^N A^{N+1} \to 1$$

as $N \to \infty$ so $(\mathbb{1} + A)$ is invertible. □

DEFINITION 28.1. The resolvent set $\rho(M)$ of a linear operator M is the set of $\lambda \in \mathbb{C}$ such that $(\lambda\mathbb{1} - M) \in \mathcal{GL}(X)$. The spectrum $\sigma(M)$ of M is the complement of $\rho(M)$, that is the set of λ such that $\lambda\mathbb{1} - M$ is not invertible.

THEOREM 28.4. *The resolvent set $\rho(M)$ is open and contains the set $\{\lambda : |\lambda| > \|M\|\}$. Furthermore, if $|\lambda| > M$ then*

$$\left\|(\lambda\mathbb{1} - M)^{-1}\right\| \le \frac{1}{|\lambda| - \|M\|}. \qquad (\star)$$

PROOF. It follows from the previous theorem that the resolvent set $\rho(M)$ is open. Also following the proof of the previous theorem we see that if $\lambda > \|M\|$ then

$$(\lambda \mathbb{1} - M)^{-1} = \sum_{n=0}^{\infty} \lambda^{-(n+1)} M^n,$$

so $\lambda \in \rho(M)$. The estimate $(\star)$ follows by summing the r.h.s. of

$$\left\|(\lambda \mathbb{1} - M)^{-1}\right\| \le \sum_{n=0}^{\infty} |\lambda|^{-(n+1)} \|M\|^n .$$

□

THEOREM 28.5. *For any* $\lambda \in \rho(M)$ *we have*

$$\left\|(\lambda \mathbb{1} - M)^{-1}\right\| \ge \frac{1}{\operatorname{dist}(\lambda, \sigma(M))}.$$

PROOF. For $|z| < 1/\left\|(\lambda \mathbb{1} - M)^{-1}\right\|$ the geometric series

$$((\lambda + z)\mathbb{1} - M)^{-1} = \sum_{n=0}^{\infty} (-z)^n (\lambda \mathbb{1} - M)^{-1-n}$$

is absolutely convergent, so $\lambda + z \in \rho(M)$. □

It follows that for any $x \in X$ and $\ell \in X'$ the function

$$f_{x,\ell}(\lambda) = \langle (\lambda \mathbb{1} - M)^{-1} x, \ell \rangle,$$

has a convergent power series expansion at each point of $\rho(M)$. Thus $f_{x,\ell}$ is analytic in $\rho(M)$. From $(\star)$ we have

$$|f_{x,\ell}(\lambda)| \le \|x\| \, \|\ell\| \, \frac{1}{|\lambda| - \|M\|}$$

for large $|\lambda|$. Thus $f_{x,\ell}$ vanishes at infinity. By Liouville's theorem a bounded entire function is constant. Thus, we have either $f_{x,\ell} \equiv 0$ or $\rho(M) \ne \mathbb{C}$.

THEOREM 28.6. $\sigma(M)$ *is a non-empty closed subset of the disk* $\{\lambda \,:\, |\lambda| \le \|M\|\}$.

PROOF. It follows from the previous proof that $\sigma(M)$ is closed and contained in $\{|\lambda| \le \|M\|\}$. Suppose $\sigma(M)$ is empty, so $\rho(M) = \mathbb{C}$. It follows that $f_{x,\ell} \equiv 0$ for all $x \in X$ and $\ell \in X'$. This however is a contradiction as then $(\lambda \mathbb{1} - M)^{-1} = 0$, so 0 is an invertible operator. □

THEOREM 28.7. $\sigma(M) = \sigma(M')$.

Since $(\lambda \mathbb{1} - M)' = \lambda \mathbb{1} - M'$, this follows easily from

LEMMA 28.8. $K \in \mathcal{GL}(X)$ *if and only if* $K' \in \mathcal{GL}(X')$.

PROOF. If K is invertible then $KL = LK = \mathbb{1}_X$ with $L = K^{-1}$. By taking adjoints

$$L'K' = K'L' = \mathbb{1}_{X'},$$

so $K'^{-1} = [K^{-1}]'$.

Supposes now that K' is invertible. Then $\operatorname{ran} K' = X'$, so the null space of K (which is the annihlator of $\operatorname{ran} K'$) is $\{0\}$. Thus K is one-to-one. To see that K is onto, note that since $N_{K'} = \{0\}$ it follows that $\operatorname{ran} K$ is dense. Thus it suffices to show $\operatorname{ran} K$ is closed. To prove this, note that $K'' \in \mathcal{GL}(X'')$ by the first part of the theorem and that $K''x = Kx$ for $x \in X$. However, X is a norm closed subspace of its double dual, since the norm on the double dual is the same as the norm on X (by the dual characterization of the norm $\|x\| = \sup_\ell |\langle x, \ell\rangle|$). Applying the following lemma with $Y = X''$ and $L = [K'']^{-1}$ it follows that K has closed range.

LEMMA 28.9. *Let $K \in \mathcal{L}(X, Y)$ and suppose there is $L \in \mathcal{L}(Y, X)$ such that $LK = 1_X$. Then $N_X = \{0\}$ and* $\operatorname{ran} K$ *is closed.*

PROOF. Clearly $N_X = \{0\}$. On the other hand if $Kx_n \to y$ is a convergent sequence in $\operatorname{ran} K$, then $x_n \to Ly$ so $Kx_n \to KLy = y$ and $y \in \operatorname{ran} K$. □

□

In finite dimensions the spectrum of a linear operator is the same as the set of eigenvalues. This follows since if X is finite dimensional then $K \in \mathcal{L}(X)$ is invertible if and only if it is one-to-one $N_X = \{0\}$. In infinite dimension, an operator K may fail to be invertible for a number of reasons:

(1) N_K may be non-trivial.
(2) ran_K is contained in a proper closed subspace of X, in which case $N_{K'}$ is non-trivial.
(3) ran_K is a proper dense subspace of X

It follows that $\lambda \in \sigma(M)$ must satisfy (at least) one of the following:

(1) λ is an eigenvalue. That is there is non-zero x such that $Mx = \lambda x$.
(2) The range of $\lambda 1 - M$ is contained in a proper closed subspace of X, in which case λ is an eigenvalue of M'.
(3) The range of $\lambda 1 - M$ may be a proper dense subspace of X.

All three possibilities occur, so spectrum — and spectral theory — in infinite dimensions is a good deal more complicated than in finite dimensions.

LECTURE 29

Some examples

Reading: §20.2-§20.3

Shifts

Let R and L denote the right and left shifts on c_0, sequences that vanish at ∞,

$$R(a_1, a_2, \dots) = (0, a_1, a_2, \dots)$$

$$L(a_1, a_2, \dots) = (a_2, a_3, \dots).$$

Let R_p, L_p denote the restriction of these operators to ℓ_p, $1 \le p < \infty$ and let R_∞, L_∞ denote the natural extension of these to ℓ_∞. Note that we have

$$R' = L_1 \quad \text{and} \quad L' = R_1,$$

$$R_p' = L_{p'} \quad \text{and} \quad L_p' = R_{p'},$$

$1/p + 1/p' = 1$, $1 \le p < \infty$.

Proposition 29.1. *$\sigma(R) = \sigma(L) = \sigma(R_p) = \sigma(L_p) = \{|\lambda| \le 1\}$, for each $1 \le p \le \infty$.*

Proof. First note that $\|R\| = \|L\| = \|R_p\| = \|L_p\| = 1$, so all of the various spectra are contained in the unit disk. On the other hand for each $|\lambda| < 1$ the sequence

$$(\lambda, \lambda^2, \lambda^3, \dots)$$

is an eigenvector of the left shift with eigenvalue λ. Since this vector lies in $\cap_p \ell_p = \ell_1$ and the spectrum is closed we have $\sigma(L) = \sigma(L_p) = \{|\lambda| \le 1\}$. Since $R_1 = L'$ and $R_p = L_{p'}'$, $1 < p \le \infty$, it follows that $\sigma(R_p) = \{|\lambda| \le 1\}$ for $1 \le p \le \infty$. Likewise, since $L_1 = R'$ it follows that $\sigma(R) = \sigma(L_1) = \{|\lambda| \le 1\}$. $\square$

Now consider the space $c_o(\mathbb{Z})$ of two sided sequences. We may define right and left shifts here as well

$$\widetilde{R}(\dots, a_{-1}, a_0, a_1, \dots) = (\dots, a_{-2}, a_{-1}, a_0, \dots)$$

$$\widetilde{L}(\dots, a_{-1}, a_0, a_1, \dots) = (\dots, a_0, a_1, a_2, \dots).$$

Note that these operators are isometries and inverses of each other $\widetilde{R}\widetilde{L} = \widetilde{L}\widetilde{R} = 1$.

Proposition 29.2. *$\sigma(\widetilde{R}) = \sigma(\widetilde{L}) = \{|\lambda| = 1\}$.*

We use

Lemma 29.3. *Let $M \in \mathcal{L}(X)$ be an isometry onto X. Then $\sigma(M) \subset \{|\lambda| = 1\}$.*

Proof. Since $\|Mx\| = \|x\|$ for all x, M is $1-1$. As M is also onto, $0 \in \rho(M)$. As $\|M\| = 1$, it follows that $\sigma(M) \subset \{|\lambda| \leq 1\}$ and that $\{|\lambda| < 1\} \subset \rho(M)$, as the geometric series

$$(\lambda - M)^{-1} = -\sum_{n=0}^{\infty} \lambda^n M^{-n-1}$$

converges there. $\square$

Proof of Proposition. It follows from the lemma that the two spectra are subsets of the circle. We show that $\operatorname{ran}(\lambda - \widetilde{R}) \neq c_0(\mathbb{Z})$ for each λ of modulus one. The proof for $\widetilde{L}$ is similar.

Suppose $(\lambda \mathsf{1} - \widetilde{R})\mathbf{a} = \mathbf{b}$. Then the coefficients satisfy

$$\lambda a_j - a_{j-1} = b_j.$$

Since $a_{j-1} - \lambda^{-1} a_{j-2} = \lambda^{-1} b_{j-1}$ it follows that

$$\lambda a_j - \lambda^{-1} a_{j-2} = (b_j + \lambda^{-1} b_{j-1}).$$

Continuing we conclude that

$$\lambda a_j - \lambda^{-n} a_{j-n-1} = \sum_{m=0}^{n} \lambda^{-m} b_{j-m}.$$

Since $a_{j-n-1} \to 0$ as $n \to \infty$ we find that

$$a_j = \lim_{n\to\infty} \sum_{m=0}^{n} \lambda^{-m-1} b_{j-m}.$$

Thus, for instance, the sequence $b_j = \lambda^{-j} \frac{1}{|j|+1}$ cannot be in $\operatorname{ran}(\lambda \mathsf{1} - \widetilde{R})$ since

$$\sum_{m=0}^{n} \lambda^{-m-1} b_{j-m} = \lambda^{-j-1} \sum_{m=0}^{n} \frac{1}{|j-m|+1}$$

diverges as $n \to \infty$. $\square$

Note an amusing point of the proof: we actually derived a formula for the inverse of $(\lambda \mathsf{1} - \widetilde{R})$. The point is that this inverse is defined only on a dense domain in X.

Similarly, we have

Proposition 29.4. *Let $\widetilde{R}_p$, $\widetilde{L}_p$ be the right and left shifts on $\ell_p(\mathbb{Z})$ for $1 \leq p \leq \infty$. Then $\sigma(\widetilde{R}_p) = \sigma(\widetilde{L}_p) = \{|\lambda| = 1\}$.*

The case $p = \infty$ is easy since then every point of the circle is an eigenvalue of each shift. The remaining cases are left as an exercise.

Volterra Integral Operators

Consider the operator of integration

$$Vf(x) = \int_0^x f(y)\mathrm{d}y$$

on the Banach space $X = C[0,1]$.

Theorem 29.5. *V is a bounded operator, $\|V\| \le 1$, and $\sigma(V) = \{0\}$. 0 is <u>not</u> an eigenvalue of V.*

Proof. It is easy to verify that $\|V\| \le 1$. Clearly $Vf = 0$ if and only if $f = 0$ so 0 is not an eigenvalue of V. Lax's proof that $\sigma(V) = \{0\}$ relies on results from the theory of Banach algebras which we have not proved yet, so let us proceed directly. We will give a formula for the inverse of $\lambda\mathsf{1} - V$. Note that

$$V^n f(x) = \int_0^x \int_0^{y_1} \cdots \int_0^{y_{n-1}} f(y_n) \mathrm{d}y_n \mathrm{d}y_{n-1} \cdots \mathrm{d}y_1.$$

It follows that

$$\|V^n f\| \le \|f\| \int_0^1 \int_0^{y_1} \cdots \int_0^{y_{n-1}} f(y_n) \mathrm{d}y_n \mathrm{d}y_{n-1} \cdots \mathrm{d}y_n = \frac{1}{n!} \|f\| .$$

Thus the geometric series

$$\sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} V^n$$

is norm convergent with norm bounded by $|\lambda|^{-1} \mathrm{e}^{|\lambda|^{-1}}$. Clearly

$$(\lambda\mathsf{1} - V) \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} V^n \;=\; \mathsf{1}.$$

Thus $\lambda\mathsf{1} - V$ is boundedly invertible for all $\lambda \neq 0$. Since the spectrum is non-empty we must have $\sigma(V) = \{0\}$. □

One can also verify directly that V is not invertible. After all $Vf(0) = 0$ so $\operatorname{ran} V$ is contained in the space of functions that vanish at 0. Note that V' is the map on $\mathcal{M}[0,1]$ which maps a measure $\mathrm{d}\mu$ to the measure $\mu(x,1]\mathrm{d}x$. That is $V'(\mu)(S) = \int_S \mu((x,1])\mathrm{d}x$ for any measurable set. Thus 0 is an eigenvalue of V' with eigenvector $\delta_0 \mathrm{d}x$.

Part 8

Compact Linear Maps

LECTURE 30

Compact Maps

Motivation: integral operators

Let T be a compact Hausdorff space and μ a Borel measure on T. Suppose $K(s,t)$ is a continuous function on $T \times T$ and define a map $\mathbf{K} \in \mathcal{L}(C(T))$ by

$$\mathbf{K}f(s) \ = \ \int_T K(s,t)f(t)\mathrm{d}\mu(t).$$

Since K is continuous on a compact space, hence uniformly continuous and bounded, the image $\mathbf{K}f$ is continuous and satisfies

$$\max_s |\mathbf{K}f(s)| \ \leq \ \text{const.} \max_t |f(t)|.$$

Thus $\mathbf{K}$ is indeed a bounded map from $C(T)$ into itself.

Such maps are certainly on of the most important examples of bounded linear maps. (There are L^1, L^2, etc. analogues.) The map is <u>not</u> typical, however, because it has one additional important property. Suppose $|f(t)| \leq 1$ for all t. Then

$$|\mathbf{K}f(s) - \mathbf{K}f(r)| \leq |\mu|(T) \sup_t |K(s,t) - K(r,t)|.$$

Since the right hand side is finite and converges to zero as $r \to s$ we find that

Lemma 30.1. *The collection $\{\mathbf{K}f \ : \ \|f\|_{C(T)} \leq 1\}$ is <u>equicontinuous</u>. That is, given $s \in T$ and $\epsilon > 0$ there is a neighborhood U of s such that $|\mathbf{K}f(s) - \mathbf{K}f(r)| \leq \epsilon$ whenever $r \in U$ and $\|f\|_{C(T)} \leq 1$.*

Since the collection is also uniformly bounded, by the Arzela-Ascoli theorem

Theorem 30.2. *The operator $\mathbf{K}$ maps the unit ball in $C(T)$ to a pre-compact set.*

Thus integral operators as above have the rather special property of mapping the unit ball of the Banach space onto a "small subset" of the Banach space. This has many useful and important consequences.

Compact operators

Definition 30.1. A map $\mathbf{C} \in \mathcal{L}(X,Y)$ is <u>compact</u> if the image $\mathbf{C}B_1(0)$ of the unit ball in X is pre-compact in Y (in the norm topology).

Remark. Recall that a subset S of a complete metric space is pre-compact if its closure is compact. Equivalently, every sequence in S has a Cauchy subsequence.

Proposition 30.3. *If C_1 and C_2 are pre-compact subsets of a Banach space then*

(1) $C_1 + C_2$ is pre-compact.
(2) MC_1 is pre-compact for any $M \in \mathcal{L}(X,Y)$ with Y a Banach space.
(3) the convex hull of C_1 is pre-compact.

Let $\mathcal{C}(X,Y) = \{$ compact maps from $X \to Y\}$.

THEOREM 30.4. *$\mathcal{C}(X,Y)$ is closed sub-space of $\mathcal{L}(X,Y)$. Furthermore if $\mathbf{M} \in \mathcal{L}(Y,Z)$ and $\mathbf{N} \in \mathcal{L}(V,X)$ then $\mathbf{MCN} \in \mathcal{C}(V,Z)$ for any $\mathbf{C} \in \mathcal{C}(X,Y)$.*

PROOF. The proof that $\mathcal{C}(X,Y)$ is a subspace is left as an exercise, based on the proposition above, as is the proof that $\mathbf{MC}$ is compact if $\mathbf{C}$ is. To see that $\mathbf{CN}$ is compact, note that $\mathbf{CN}B_1 \subset \mathbf{C}B_{\|N\|}$, so $\frac{1}{\|N\|}\mathbf{CN}B_1$ is pre-compact and hence so is $\mathbf{CN}$ (by (2) of the proposition).

Finally, suppose $\mathbf{C}_n \to \mathbf{C}$ in $\mathcal{L}(X,Y)$. Let x_j be a sequence in the unit ball of X. Let $x_{j;1}$ be a subsequence such that $\mathbf{C}_1 x_{j;1}$ converges as $j \to \infty$. By induction construct $x_{j;n}$ for each n such that

(1) $x_{j;n}$ is a subsequence of $x_{j;n-1}$
(2) $\lim_j \mathbf{C}_n x_{j;n} = y_n$.

It follows that
$$y_m = \lim_j \mathbf{C}_m x_{j;n}$$
for any $m \le n$. Note that $y_m - y_n = \lim_j (\mathbf{C}_m - \mathbf{C}_n) x_{j;n}$, so
$$\|y_m - y_n\| \le \|\mathbf{C}_m - \mathbf{C}_n\| .$$
Thus y_n are Cauchy and have a limit y. For each n, let j_n be such that
$$\|y_n - \mathbf{C}_n x_{j_n;n}\| \le \|\mathbf{C}_n - \mathbf{C}\| .$$
Consider the diagonal sequence $x_{j_n;n}$. Then
$$\|\mathbf{C}x_{j_n;n} - y\| \le \|\mathbf{C} - \mathbf{C}_n\| + \|\mathbf{C}_n x_{j_n;n} - y_n\| + \|y_n - y\| \le 3\|\mathbf{C}_n - \mathbf{C}\| \to 0.$$
Thus $\mathbf{C}x_{j_n;n}$ is a convergent subsequence of $\mathbf{C}x_j$ and so $\mathbf{C}B_1(0)$ is compact. □

Let $\mathcal{C}(X) = \mathcal{C}(X,X)$.

THEOREM 30.5. *Let $\mathbf{C} \in \mathcal{C}(X)$, let $\mathbf{I}$ be the identity map on X, and let $\mathbf{T} = \mathbf{I} - \mathbf{C}$. Then*

(1) $N_{\mathbf{T}}$ is finite dimensional.
(2) There is an integer i such that
$$N_{\mathbf{T}^k} = N_{\mathbf{T}^i} \quad \text{for } k \ge i.$$
(3) $\operatorname{ran} \mathbf{T}$ is closed.

PROOF. Note that $x \in N_{\mathbf{T}}$ iff $x = \mathbf{T}x$. Thus the unit ball of $N_{\mathbf{T}}$ is contained in $\mathbf{T}B_1(0)$. Thus the closed unit ball of $N_{\mathbf{T}}$ is compact. So $N_{\mathbf{T}}$ is finite dimensional. (See Lecture 5.)

Note that $N_{\mathbf{T}^{k+1}} \supset N_{\mathbf{T}^k}$, and that each subspace is finite dimensional (since $\mathbf{T}^k$ is compact for any k). Suppose this sequence never stabilizes. By Lemma 5.2 there is then a vector x_k for each k such that
$$x_k \in N_{\mathbf{T}^k}\ , \quad \|x_k\| = 1 \quad \text{and} \quad \|x_k - x\| > \frac{1}{2} \quad \text{for all } x \in N_{\mathbf{T}^{k-1}}.$$
Then, if $m < n$,
$$\mathbf{C}(x_n - x_m) = x_n - \mathbf{T}x_n - x_m + \mathbf{T}x_m.$$
Now $\mathbf{T}x_n, x_m, \mathbf{T}x_m \in N_{\mathbf{T}^{n-1}}$ (since $m < n$), so
$$\|\mathbf{C}x_n - \mathbf{C}x_m\| \ge \frac{1}{2}.$$

So $\mathbf{C}x_n$ has no Cauchy subsequence, which is a contradiction.

To prove $\operatorname{ran}\mathbf{T}$ is closed, let $y_k = \mathbf{T}x_k$ and suppose $y_k \to y$. We may shift x_k as we like by elements of the $N_{\mathbf{T}}$. So x_k may get very large. However, if we consider the distance

$$d_k = \inf_{u \in N_{\mathbf{T}}} \|x_k - u\|,$$

from x_k to the $N_{\mathbf{T}}$, then I claim we have

$$\sup_k d_k < \infty.$$

Indeed, suppose not. Then, passing to a subsequence so $d_k \to \infty$ and choosing u_k so $\|x_k - u_k\| \le 2d_k$ we have

$$0 = \lim_k \frac{y_k}{d_k} = \lim_k \mathbf{T}\frac{x_k - u_k}{d_k},$$

since y_k are bounded. As $\|x_k - u_k\| / d_k \le 2$ we may pass again to a subsequence to conclude that

$$\mathbf{C}\frac{x_k - u_k}{d_k}$$

converges. Thus

$$\lim_k \frac{x_k - u_k}{d_k} = z = \lim_k \mathbf{C}\frac{x_k - u_k}{d_k} = \mathbf{C}z,$$

so $\mathbf{T}z = 0$, that is $z \in N_{\mathbf{T}}$. But then $\|x_k - u_k - d_k z\| \ge d_k$ so $\left\|\frac{x_k - u_k}{d_k} - z\right\| \ge 1$ contradicting the convergence derived above. Thus $\sup_k d_k < \infty$.

Now since d_k are bounded, we may choose $u_k \in N_{\mathbf{T}}$ so that $\|x_k - u_k\| \le 2\sup_k d_k$. Now pass to a subsequence so that

$$\mathbf{C}(x_k - u_k)$$

converges. But then

$$\lim_k (x_k - u_k) \;=\; y + \lim_k \mathbf{C}(x_k - u_k).$$

Thus $\lim_k (x_k - u_k) = x$ exists and $y = \mathbf{T}x \in \operatorname{ran}\mathbf{T}$. $\square$

LECTURE 31

Fredholm alternative

Reading: §21.1

THEOREM 31.1 (Fredholm Alternative). *Let* $\mathbf{C} \in \mathcal{C}(X)$. *Then* $\mathbf{T} = \mathbf{I} - \mathbf{C}$ *satisfies*

$$\dim N_{\mathbf{T}} = \dim(X/\operatorname{ran}\mathbf{T}).$$

REMARK. The result is known as the "Fredholm Alternative" because Fredholm proved this result in the context of integral operators. The "alternative" is that either one is an eigenvalue of $\mathbf{C}$, or the equation

$$y = x - \mathbf{C}x$$

is uniquely solvable for x for any y.

COROLLARY 31.2. *Let* $\mathbf{C} \in \mathcal{C}(X)$. *Then a non-zero point* $\lambda \in \mathbb{C}$ *is in the spectrum of* $\mathbf{C}$ *if and only if* λ *is an eigenvalue of finite algebraic and geometric multiplicity, where*

$$\textit{geometric multiplicity} = \dim N_{\lambda\mathbf{I}-\mathbf{C}}$$

and

$$\textit{algebraic multiplicity } = \sup_j \dim N_{(\lambda\mathbf{I}-\mathbf{C})^j}.$$

PROOF. Apply the previous two theorems to the compact map $\lambda^{-1}\mathbf{C}$. □

Caution: <u>a compact map need not have eigenvalues</u>. The Volterra operator of the last lecture is an example.

Recall that $\dim X/Y$ is the <u>codimension</u> of Y, denoted The identity

$$\dim N_M = \operatorname{codim}\operatorname{ran} M$$

is valid for any operator $M \in \mathcal{L}(X)$ if X is finite dimensional. In an infinite dimensional setting, this identity need not hold even if both sides are finite. If both sides are finite, the <u>index</u> of M is the difference

$$\operatorname{ind}\mathbf{T} = \dim N_M - \operatorname{codim}\operatorname{ran} M.$$

Thus the alternative says that the index of a $\mathbf{I} - \mathbf{C}$ is zero for $\mathbf{C}$ compact.

To prove the theorem we will use the following

LEMMA 31.3. *Let* $\mathbf{C} \in \mathcal{C}(X)$ *and suppose* $Y \subset X$ *is an invariant subspace for* $\mathbf{C}$, *so* $\mathbf{C}Y \subset Y$. *Then*

$$\widetilde{\mathbf{C}}[x] = [\mathbf{c}x]$$

is a compact map from $X/Y \to X/Y$.

PROOF. Exercise. □

Proof of Theorem. First suppose $\dim N_{\mathbf{T}} = 0$. We need to show that $\operatorname{ran}\mathbf{T} = X$. If, on the contrary, $X_1 = \operatorname{ran}\mathbf{T}$ is a proper subspace of X. Since $\mathbf{T}$ is one-to-one, $X_2 = \mathbf{T}X_1$ is a proper subspace of X_1. By induction, with $X_k = \mathbf{T}^k X$, we find that $X_1 \supset X_2 \supset X_3 \cdots$ with proper inclusion at every step. Now X_1 is closed by Thm 30.5 from last time. Likewise $X_k = \mathbf{T}^k X$ and $\mathbf{T}^k = \mathbf{I} - \sum_{j=1}^{k} \binom{k}{j} (-1)^{j-1} \mathbf{C}^j = \mathbf{I} -$ compact, so X_k is closed. Thus, we may find a sequence of vectors x_j such that

$$x_j \in X_j \ , \quad \|x\| = 1 \ , \quad \text{and } \operatorname{dist}(x_j, X_{j+1}) > \frac{1}{2}.$$

It follows, if $m < n$, that

$$\|\mathbf{C}x_m - \mathbf{C}x_n\| = \|\mathbf{x_m} - \mathbf{T}x_m - x_n + \mathbf{T}x_n\| > \frac{1}{2},$$

since $\mathbf{T}x_m + x_n - \mathbf{T}x_n \in X_{m+1}$. Thus $\mathbf{C}x_n$ has no Cauchy subsequence, which is a contradiction.

If $\dim N_{\mathbf{T}} > 0$, then by Thm 30.5 from last time, we have $N_{\mathbf{T}^{i+1}} = N_{\mathbf{T}^i}$ for i large enough. Let $N = N_{\mathbf{T}^i}$. So N is an invariant subspace for $\mathbf{T}$, and since $\mathbf{C} = \mathbf{I} - \mathbf{T}$ we see that N is invariant for $\mathbf{C}$ as well. Thus $\widetilde{\mathbf{C}}$ as defined above is a compact map of $X/N \to X/N$. Consider $\widetilde{\mathbf{T}} = \widetilde{\mathbf{I}} - \widetilde{\mathbf{C}}$. That, is

$$\widetilde{\mathbf{T}}[x] = [Tx],$$

which is well-defined since N is invariant under T. Note that $N_{\widetilde{\mathbf{T}}} = \{0\}$, since if $[x] \in N_{\widetilde{\mathbf{T}}}$ then $Tx \in N \implies x \in N_{\mathbf{T}^{i+1}} = N_{\mathbf{T}^i} = N$ so $[x] = 0$. Thus by $\operatorname{ran}\widetilde{\mathbf{T}} = X/N$. Thus for every $y \in X$ there is $x \in X$ and $z \in N$ such that

$$\mathbf{T}x = y + z.$$

Thus $X = \operatorname{ran}\mathbf{T} + N$. Thus

$$\dim(X/\operatorname{ran}\mathbf{T}) = \dim(N/(N \cap \operatorname{ran}\mathbf{T})).$$

By finite dimensional linear algebra

$$\dim(N/(N \cap \operatorname{ran}\mathbf{T})) = \dim N_{\mathbf{T}}. \quad \square$$

Theorem 31.4 (Schauder). $\mathbf{C} \in \mathcal{L}(X)$ *is compact if and only if* $\mathbf{C}'$ *is compact.*

Proof. We show $\mathbf{C}$ compact $\implies$ $\mathbf{C}'$ is compact. The reverse implication follows by noting that $\mathbf{C}$ is the restriction of $\mathbf{C}''$ to the closed subspace X of X'' and applying the following

Lemma 31.5. *Let* $\mathbf{C}$ *be compact on* X *and let* $Y \subset X$ *be a closed subspace. Then the restriction of* $\mathbf{C}$ *to* Y *is a compact map.*

Proof. Exerise. $\square$

We must show, if $\ell_n \in B_1(0) \subset X'$, that $\mathbf{C}'\ell_n$ has a convergent subsequence. Consider the closure of $K = \mathbf{C}B_1(0) \subset X$. This is a compact set and on this set

$$|\ell_n(x) - \ell_n(y)| \le \|x - y\|$$

since $\|\ell_n\| \le 1$. Thus $\ell_n \mid_K$ are equicontinuous, and bounded, so have a convergent subsequence by Arzela-Ascoli. That is, passing to a subsequence, we have

$$\lim_{n\to\infty} \sup_{y \in K} |\langle y, \ell_n - f(y)| = 0$$

for some cts function f on K. Thus

$$\sup_{x \in B_1(0)} |\langle x, \mathbf{C}'\ell_n - f(\mathbf{C}x)| \longrightarrow 0.$$

It follows that $\ell(x) = f(\mathbf{C}x)$ is a linear functional and the norm limit of $\mathbf{C}'\ell_n$. So $\mathbf{C}'$ is compact. $\square$

THEOREM 31.6 (Fredholm Alternative). *Let* $\mathbf{C} \in \mathcal{C}(X)$, $\mathbf{T} = \mathbf{I} - \mathbf{C}$. *Then*

(1) $x \in \operatorname{ran} \mathbf{T}$ *if and only if* $\ell(x) = 0$ *for all* $\ell \in N_{\mathbf{T}'}$.
(2) $\dim N_{\mathbf{T}} = \dim N_{\mathbf{T}'}$.

PROOF. (1) For a general operator $\mathbf{T}$ we have

$$\overline{\operatorname{ran} \mathbf{T}} = \{x \ : \ \ell(x) = 0 \text{ fo all } \ell \in N_{\mathbf{T}'}\}.$$

Since $\operatorname{ran} \mathbf{T}$ is closed (1) follows.

(2) The null space $N_{\mathbf{T}'}$ is isomorphic to the dual of $X/\operatorname{ran} \mathbf{T}$ via the pairing

$$\langle [x], \ell \rangle = \langle x, \ell \rangle$$

which is well defined since ℓ annihlates the $\operatorname{ran} \mathbf{T}$. Thus

$$\dim N_{\mathbf{T}'} = \operatorname{codim} \operatorname{ran} \mathbf{T} = \dim N_{\mathbf{T}}. \quad \square$$

LECTURE 32

Spectral Theory of Compact Maps

Theorem 32.1 (F. Riesz). *Let X be a complex Banach space and $\mathbf{C} \in \mathcal{C}(X)$. The spectrum of $\mathbf{C}$ is a denumerable or finite set of points whose only accumulation point, if any, is 0. If $\dim X = \infty$ then $0 \in \sigma(\mathbf{C})$:*

$$\sigma(\mathbf{C}) = \{\lambda_j\} \cup \{0\}.$$

Furthermore,

(1) Each non-zero $\lambda_j \in \sigma(\mathbf{C})$ is an eigenvalue of finite algebraic and geometric multiplicity.

(2) The resolvent $(z - \mathbf{C})^{-1}$ has a pole at each non-zero λ_j: that is, there is $n_j \geq 1$ such that

$$(z - \lambda_j)^{n_j} \langle (z - \mathbf{C})^{-1} x, \ell \rangle$$

is analytic in a neighborhood of λ_j for every $\ell \in X'$ and $x \in X$.

Proof. We have already seen that any non-zero point of the spectrum is an eigenvalue of finite multiplicity.

Suppose we have a sequence λ_n of eigenvalues and eigenvectors

$$\mathbf{C} x_n = \lambda_n x_n,$$

with $\lambda_n \neq \lambda_m$. Let Y_n be the linear space of $x_1, \ldots, x_n$. Suppose

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

Then

$$\sum_{j=1}^{n} a_j \lambda_j^k x_j = 0$$

for all k, so

$$\sum_{j=1}^{n} a_j p(\lambda_j) x_j = 0$$

for all polynomials p. Since the λ_j are distinct, we may pick a polynomial that vanishes for all λ_j except λ_{j_0} for example $p(\lambda) = \prod_{j \neq j_0} (\lambda - \lambda_j)$. Thus $a_j = 0$ for all j, so x_j are linearly independent. It follows that Y_{n-1} is a proper subspace of Y_n.

The implication of this is the same as above: we find a sequence of vectors $y_n \in Y_n$ such that

$$\|y_n\| = 1 \quad \text{and} \quad \|y_n - y\| > \frac{1}{2} \quad \text{for all } y \in Y_{n-1}.$$

Since

$$y_n = \sum_{j=1}^{n} a_j^{(n)} x_j$$

we have

$$\mathbf{C}y_n = \sum_{j=1}^{n} a_j^{(n)} \lambda_j x_j$$

so

$$\mathbf{C}y_n - \lambda_n x_n = \sum_{j=1}^{n-1} a_j^{(n)} (\lambda_j - \lambda_n) x_j \in Y_{n-1}.$$

Thus, for $n > m$,

$$\mathbf{C}y_n - \mathbf{C}y_m = \lambda_n y_n - y \quad \text{for some } y \in Y_{n-1}.$$

Thus

$$\|\mathbf{C}y_n - \mathbf{C}y_m\| \geq \frac{|\lambda_n|}{2}.$$

Since any subsequence of $\mathbf{C}y_n$ has a Cauchy subsequence, we must have $\lambda_n \to 0$. Since λ_n was an arbitrary sequence of eigenvalues, it follows that 0 is the only (possible) accumulation point of eigenvalues. Thus there are only finitely many eigenvalues outside any disk around the origin. Hence there at most countably many eigenvalues.

It remains to show that the resolvent has poles at λ_j. Let z be a complex number, then finding the resolvent $u = (z - \mathbf{C})^{-1}x$ amounts to solving

$$x = zu - \mathbf{C}u$$

for x. Suppose z is close to λ_j. Let $N = N_{(\lambda - \mathbf{C})^i} = N_{(\lambda - \mathbf{C})^{i+1}}$ with i sufficiently large. Let $\widetilde{\mathbf{C}}$ be the quotient map on X/N. Since λ_j is not an eigenvalue of $\widetilde{\mathbf{C}}$ (why?), we conclude that $\lambda_j - \widetilde{\mathbf{C}}$ is invertible. It follows that $z - \widetilde{\mathbf{C}}$ is invertible for $z - \lambda_j < \epsilon$ for some $\epsilon > 0$ (recall that the set of invertible maps is open) and

$$\left\|(z - \mathbf{C})^{-1}\right\| \leq \text{const.} \quad \text{for } |z - \lambda_j| < \epsilon.$$

Thus

$$z[v] - \widetilde{\mathbf{C}}[v] = [x]$$

has a unique solution for an equivalence class $[v] = v + N$ with

$$\|[v]\| \leq \text{const.} \|[x]\| .$$

Thus given x we may find $n \in N$ and $v(z) \in X$ such that

$$(z - \mathbf{C})v(z) = x - n(z).$$

We may choose $v(z)$ so that $\|v(z)\|$ is bounded for $|z - \lambda_j| < \epsilon$. It follows that $\|n(z)\|$ is bounded as well:

$$\|n(z)\| \leq (|z| + \|\mathbf{C}\|) \|v(z)\| + \|x\| .$$

Let us now solve

$$(z - \mathbf{C})y(z) = n(z)$$

for $y(z) \in N$ to obtain $u(z) = v(z) + y(z)$. This is a linear algebra problem. We know that N is an invariant subspace for $\mathbf{C}$ and that λ_j is in the spectrum $\sigma(\mathbf{C}\mid_N)$. In fact, λ_j is the unique point of the spectrum of $\mathbf{C}\mid_N$. Indeed, $(\lambda_j - \mathbf{C})^i = 0$ on N. It follows that

$$\left\|(z - \mathbf{C}\mid_N)^{-1}\right\| \leq \text{const.} \frac{1}{|\lambda - z|^i}.$$

(It doesn't really matter for the proof, but yes it is the same i in both spots.) Hence,

$$\|u(z)\| \leq \|v(z)\| + \|y(z)\| \leq \text{const.}(1 + |z - \lambda_j|^{-i})$$

for z close to λ_j.

Thus $(z-\lambda_j)^i\langle(z-\mathbf{C})^{-1}x,\ell\rangle$ is analytic in $\{0<|z-\lambda_j|<\epsilon\}$ with a removable singularity at $z=\lambda_j$. That is $\langle(z-\mathbf{C})^{-1}x,\ell\rangle$ has a pole at λ_j □

Note that the resolvent may not have a pole at 0. For instance, the Volterra integral operator has resolvent

$$(\lambda - V)^{-1} = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} V^n$$

for $\lambda \neq 0$. This analytic operator valued function has an essential singularity at 0. If one computes

$$\begin{aligned}(\lambda - V)^{-1}x^n = \sum_{m=0}^{\infty} \frac{1}{\lambda^{m+1}} V^m x^n = \sum_{m=0}^{\infty} \frac{1}{\lambda^{m+1}} \frac{n!}{(n+m)!} x^{n+m} \\ = n!\lambda^{n-1} \sum_{m=n}^{\infty} \frac{1}{m!}\left(\frac{x}{\lambda}\right)^m = n!\lambda^{n-1}\mathrm{e}^{x/\lambda} - \sum_{m=0}^{n-1} \frac{n!}{m!} x^m \lambda^{n-1-m},\end{aligned}$$

then one can see the essential singularity very clearly in the first term. The second term is regular and has limit

$$nx^{n-1} \quad \text{as } \lambda \longrightarrow 0.$$

Interestingly, this is $-V^{-1}x^n$ where V^{-1} is the densely defined left inverse for V, namely differentiation:

$$\partial_x V f(x) \;=\; \partial_x \int_0^x f(y)\mathrm{d}y \;=\; f(x).$$

Homework III

Homework III

Due: April 30, 2008

This homework assignment deals with elliptic PDE's. Let $\Omega \subset \mathbb{R}^d$ be an open set with compact closure. Suppose we are given a measurable function $\sigma(x)$ on Ω which is bounded above and below:

$$0 < a \le \sigma(x) \le a^{-1} < \infty \quad \text{for all } x \in \Omega.$$

Consider the PDE

$$\nabla \cdot \sigma(x) \nabla u(x) = f(x) \quad x \in \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{$\star$}$$

Our goal is to show that this equation has a unique solution $u = Sf$ with S a compact symmetric operator on L^2 and to consider the spectral theory of S.

(1) Show for any $f \in L^2(\Omega)$ that ($\star$) has a unique solution (in the sense of distributions) u in $L^2(\Omega)$. (Hint: use Lax Milgram. The solution is outlined in Lecture 16.)

(2) Let $\Phi(x)$ be a smooth, non-negative, compactly supported, function on $\mathbb{R}^d$ with $\int_{\mathbb{R}^d} \Phi(x) \mathrm{d}x = 1$. Let F_t map $L^2(\Omega)$ into itself by

$$F_t f(x) = \chi_\Omega(x) \int_{\mathbb{R}^d} t^{-d} \Phi(\frac{x-y}{t}) f(y) \mathrm{d}y = \chi_\Omega(x) \int_{\mathbb{R}^d} t^{-d} \Phi(\frac{y}{t}) f(x-y) \mathrm{d}y,$$

where we take $f \equiv 0$ outside Ω.

(a) Use Minkowski's inequality:

$$\left(\int \mathrm{d}\mu(x) \left| \int f(x,y) \mathrm{d}\nu(y) \right|^2 \right)^{\frac{1}{2}} \le \int \mathrm{d}\nu(y) \left(\int \mathrm{d}\mu(x) \left| f(x,y) \right|^2 \right)^{\frac{1}{2}}$$

to show that $F_t \in \mathcal{L}(L^2(\Omega)$ with $\|F_t\|_{L^2(\Omega) \to L^2(\Omega)} \le 1$.

(b) Show also that

$$\sup_x |F_t f(x)| \le C t^{-d} \|f\|_{L^2(\Omega)} .$$

(Hint: the constant will be proportional to $\mathrm{Vol}(\Omega)$.)

(c) Show that for $f \in L^2(\Omega)$ and $x \in \Omega$

$$f(x) - F_t f(x) = \int_{\mathbb{R}^d} t^{-d} \Phi(\frac{y}{t})(f(x) - f(x-y)) \mathrm{d}y.$$

(d) Apply Minkowski's inequality to conclude, if $f \in C_c^2(\Omega)$ then

$$\|f - F_t f\|_{L^2(\Omega)} \le Ct \|f\|_{H_0^1(\Omega)} , \tag{$\star\star$}$$

where

$$\|f\|_{H_0^1(\Omega)}^2 = \sum_{j=1}^d \|\partial_j f\|_{L^2(\Omega)}^2 .$$

(Hint: write $f(x) - f(x-y) = \int_0^1 y \cdot \nabla f(x - sy)) \mathrm{d}s$. Make sure you get the factor of t on the r.h.s!)

(e) Conclude that $(\star\star)$ holds for $f \in H_0^1(\Omega)$ with

$$H_0^1(\Omega) = \{\text{completion of } C_c^\infty(\Omega) \text{ in the norm } \|f\|_{H^1}^2 = \sum_{j=1}^d \|\partial_j f\|_{L^2}^2\}.$$

(3) Let $f_n \in H_0^1(\Omega)$ be a sequence so that $\partial_j f_n \rightharpoonup v_j$ weakly in $L^2(\Omega)$ for each j.

(a) Use Lemma 16.1 of Lecture 16 (Lemma 2 of Ch. 7 in Lax) to conclude that $f_n \rightharpoonup f$ weakly in $L^2(\Omega)$ for some $f \in H_0^1(\Omega)$ and that $\partial_j f = v_j$. (Hint: first use Alaogulu's theorem to conclude that f_n has a weakly convergent subsequence with a limit f that satisfies $\partial_j f = v_j$ for each j. Next show that any other weakly convergent subsequence must also converge to f. Conclude that $f_n \rightharpoonup f$.)

(b) Show that

$$F_t f_n(x) \longrightarrow F_t f(x) \text{ as } n \to \infty.$$

for every $x \in \Omega$. Since $|F_t f_n(x)| \leq Ct^{-d} \|f_n\|_{L^2(\Omega)}$ by (2b) use dominated convergence and the principle of uniform boundedness to conclude that

$$\|F_t f_n - F_t f\|_{L^2} \longrightarrow 0 \text{ as } n \to \infty.$$

(c) Use this result and $(\star\star)$ to conclude that

$$\|f_n - f\|_{L^2} \longrightarrow 0 \text{ as } n \to \infty.$$

(d) Combine these to prove

THEOREM 32.2 (Rellich). *If $M \subset H_0^1(\Omega)$ with uniformly bounded $H_0^1(\Omega)$ norm then M is pre-compact in $L^2(\Omega)$.*

(4) Returning to the PDE $(\star)$, let $S : L^2(\Omega) \to L^2(\Omega)$ be the map $Sf = u$ where u solves $(\star)$. Show that S is compact. (Hint: see §22.3 where this is done for $\sigma \equiv 1$. However, we have not assumed that Ω has a smooth boundary. But you have just proved the results needed.)

(5) Show that S is symmetric on $L^2(\Omega)$. Conclude that there is an ortho-normal basis $\{\phi_j\}$ for $L^2(\Omega)$ consisting of eigenfunctions to the linear operator $\nabla \cdot \sigma(x)\nabla$. That is, there is a set $\{\phi_j\}$ such that

$$\overline{\text{span}}\{\phi_j\} = L^2(\Omega) \ , \quad \langle \phi_j, \phi_k \rangle_{L^2(\Omega)} = \delta_{i,j},$$

and

$$\nabla \cdot \sigma(x)\nabla\phi_j(x) = \lambda_j \phi_j(x).$$

Show that the eigenvalues λ_j are real and diverge to ∞. How are they related to the eigenvalues of S?

Part 9

Compact Linear Maps in Hilbert Space

LECTURE 33

Compact Symmetric Operators

We now specialize to the case in which the Banach space X is a Hilbert space. For spectral analysis it is useful to consider a complex Hilbert space H, in which case in place of the transpose it is natural to consider

DEFINITION 33.1. The adjoint of a bounded linear map $T \in \mathcal{L}(H, K)$ with H and K Hilbert spaces is the linear map $T^\dagger \in \mathcal{L}(K, H)$ given by

$$\langle T^\dagger u, v \rangle_H = \langle u, Tv \rangle_K$$

for all $u \in K$ and $v \in H$.

Recall that, by the Riesz-Frechet theorem the dual of a Hilbert space H is isomorphic to the Hilbert space itself via the conjugate linear map

$$u \mapsto \langle \ell_u(\cdot) = \cdot, u \rangle.$$

Thus we have

$$T' \ell_u = \ell_{T^\dagger u}.$$

Since the correspondence $u \mapsto \ell_u$ is conjugate linear, it follows that

$$(T + aS)^\dagger = T + a^\star S^\dagger.$$

We also have

$$(ST)^\dagger = T^\dagger S^\dagger$$

and

$$\sigma(T^\dagger) = \sigma(T)^\star = \{\lambda^\star \ : \ \lambda \in \sigma(T)\}.$$

DEFINITION 33.2. A bounded operator $T \in \mathcal{L}(H)$ is called Hermitian (or symmetric, or self-adjoint) if $T = T^\dagger$.

PROPOSITION 33.1. *If T is Hermitian then*

(1) $\langle Tu, u \rangle$ is real for all $u \in H$.
(2) If $\langle Tu, u \rangle = 0$ for all $u \in H$ then $T = 0$.
(3) $\sigma(T) \subset \mathbb{R}$.

PROOF. (1) Note that $\langle Tu, u \rangle = \langle u, Tu \rangle = \langle Tu, u \rangle^\star$. For (2), note that if $\langle Tu, u \rangle = 0$ for all u then

$$0 = \langle T(u \pm \mathrm{i}v), u \pm \mathrm{i}v \rangle = \pm \mathrm{i} \langle Tv, u \rangle \mp \mathrm{i} \langle Tu, v \rangle.$$

Combining the results we get

$$0 = \langle Tu, v \rangle$$

for all $u, v \in H$. Thus $T \equiv 0$.

For (3), we need to show $N_{zu-T} = \{0\}$ and $\mathrm{ran}(z - T) = H$, for $z \in \mathbb{C} \setminus \mathbb{R}$. Suppose

$$zu - Tu = 0.$$

It follows that

$$z\|u\|^2 - \langle Tu, u\rangle = 0.$$

Since the second term is real we find $\operatorname{Im} z \|u\|^2 = 0$. Thus $\|u\| = 0$ and $u = 0$. That is, $N_{zu-T} = \{0\}$.

Similarly, suppose u is perpendicular to $\operatorname{ran}(z - T)$. Then

$$0 = \langle (z - T)u, u\rangle,$$

so $\|u\| = 0$. It follows that $\operatorname{ran}(z - T)$ is dense. To see that $\operatorname{ran}(z - T)$ is closed, note

$$\begin{aligned}\|(z-T)u\|^2 &= \langle (z-T)u, (z-T)u\rangle \\ &= |z|^2\|u\|^2 + \|Tu\|^2 - z\langle u, Tu\rangle - z^\star\langle Tu, u\rangle \\ &= |z|^2\|u\|^2 + \|Tu\|^2 - 2\operatorname{Re} z\langle Tu, u\rangle \\ &= (\operatorname{Im} z)^2\|u\|^2 + \|(\operatorname{Re} z - T)u\|^2 \geq (\operatorname{Im} z)^2\|u\|^2.\end{aligned}$$

Thus, if $\operatorname{ran}(z - T)u_n \to x$ we find that u_n is Cauchy. since $(z - T)$ is bounded we must have $(z - T)u = x$ for $u = \lim_n u_n$. □

THEOREM 33.2. *Let H be a infinite dimensional separable Hilbert space. If $T \in \mathcal{L}(H)$ is Hermitian and compact then there is an orthonormal basis for H consisting of eigenvectors of T. That is, there is an orthonormal basis $\{\phi_n\}_{n=1}^\infty$ and a sequence of real numbers $\lambda_n \to 0$ such that*

$$T\phi_n = \lambda_n\phi_n.$$

The spectrum $\sigma(T) = \{\lambda_n\}_{n=1}^\infty$.

REMARK. If H is not-separable and T is compact Hermitian then there is a closed separable subspace $Y \subset H$ such that $T : Y \to Y$ and

$$T|_{Y^\perp} = 0.$$

Note that we already have $\sigma(H) = \{\lambda_n\} \subset \mathbb{R}$. To proceed let us show first

LEMMA 33.3. *Let $T \in \mathcal{C}(H)$. Let $\lambda_+ = \max\sigma(T)$ and $\lambda_- = \inf\sigma(T)$. Then*

$$\lambda_+ = \sup_{\|u\|=1}\langle Tu, u\rangle = \max_{\|u\|=1}\langle Tu, u\rangle$$

and

$$\lambda_- = \inf_{\|u\|=1}\langle Tu, u\rangle = \min_{\|u\|=1}\langle Tu, u\rangle.$$

In particular, if $\sigma(T) = \{0\}$ then $T \equiv 0$.

PROOF. Let us prove the identity for λ_+. The identity for λ_- follows from the first applied to $-T$. Let $R = \sup_{\|u\|=1}\langle Tu, u\rangle$. Note that

$$\langle Tu, u\rangle \leq R\|u\|^2.$$

It follows that $R \geq 0$ as otherwise $|\langle Tu, u\rangle| \geq |R|\,\|u\|^2$, so $|R|\,\|u\| \leq \|Tu\|$, which contradicts compactness. (Look at Tu_n with u_n an orthonormal basis.)

Let u_n be a sequence of vectors with $\|u_n\| \leq 1$ so that $\langle Tu_n, u_n\rangle \to R$. Passing to a subsequence, we suppose that $u_n \rightharpoonup u_+$ and $Tu_n \to Tu_+$ (strongly). It follows that

$$\langle Tu_n, u_n\rangle \to \langle Tu_+, u_+\rangle,$$

so $\langle Tu_+, u_+\rangle = R$. Thus the sup is actually a max.

If $\lambda > R$ then, for $\|u\| \leq 1$,

$$\|(\lambda - T)u\| \geq \lambda - \langle Tu, u\rangle\lambda^2 \geq \lambda - R.$$

It follows that

$$\|(\lambda - T)u\| \geq (\lambda - R)\,\|u\|^2$$

for all u. Thus $\lambda > R \geq 0$ is not an eigenvalue of T. It follows that $\lambda_+ \leq R$.

If $R = 0$, we are done as $\lambda_+ \geq 0$. On the other hand, if $R > 0$ then $u_+ \neq 0$, since $R = \langle Tu_+, u_+$. Clearly u_+ is a unit vector. Let $w \in H$ and consider the function

$$F_w(t) = \frac{\langle T(u_+ + tw), u_+ + tw\rangle}{\|u_+ + tw\|^2}.$$

So F_w is smooth and takes its maximum at 0. Differentiating with respect to t we get

$$0 = F_w'(0) = \langle Tw, u_+\rangle + \langle Tu_+, w\rangle - \langle Tu_+, u_+\rangle\left[\langle w, u_+\rangle + \langle u_+, w\rangle\right].$$

Thus

$$0 = F_w'(0) + \mathrm{i}F_{\mathrm{i}w}'(0) = \langle Tu_+, w\rangle - R\langle u_+, w\rangle.$$

Since this holds for any w, we conclude that $Tu_+ = Ru_+$, so $R \in \sigma(T)$.

Finally, note if $\sigma(T) = \{0\}$ then $\langle Tu, u\rangle = 0$ for all u so $T \equiv 0$ by the proposition proved above. □

Proof of Theorem. If $T \equiv 0$ there is nothing to show as any orthonormal basis will do. If T is not zero, then it has a non-zero eigenvalue λ_1 by the lemma. Let the corresponding eigenvector be u_1. If $u \perp u_1$ then

$$\langle Tu, u_1\rangle = \lambda_1\langle u, u_1\rangle = 0,$$

so $Tu \perp u_1$. Thus $\{u_1\}^\perp$ is an invariant subspace for T. Either T is identically zero on $\{u_1\}^\perp$ or it has an eigenvector there. Proceed by induction.

More formally, consider the collection $\mathcal{S}$ of closed subspaces Y of H invariant under T and such that T restricted to Y has an orthonormal basis of eigenvectors. As above $Y^\perp$ is invariant under T so either $T \equiv 0$ on $Y^\perp$ or T has a non-zero eigenvector there. Partially order $\mathcal{S}$ by inclusion. It follows that any maximal element Y of $\mathcal{S}$ has $Y^\perp = \{0\}$. □

LECTURE 34

Min-Max

Let us look at the variational characterization of eigenvalues again. We have

$$\lambda_+ = \max \sigma(T) = \max_{\|u\|\le 1} \langle Tu, u\rangle.$$

and a similar identity at the bottom of the spectrum. As it turns out, one can construct in this way all of the eigenvalues. Indeed, suppose we are given the largest $N-1$ positive eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{N-1} > 0$ of T (listed according to multiplicity) together with the corresponding eigenvectors $x_1, \cdots, x_{N-1}$. Let

$$\lambda_N = \max_{u\perp x_1,\dots x_{N-1},\ \|u\|\le 1} \langle Tu, u\rangle. \tag{$\star$}$$

If $\lambda_N > 0$ then λ_N is an eigenvalue of T, $\lambda_N \le \lambda_{N-1}$, and the corresponding maximizer u_N is an eigenvector. If $\lambda_N = 0$ then there are no more positive eigenvalues. (Zero may or may not be an eigenvalue.) Similarly we can find the negative eigenvalues and corresponding eigenvectors by minimizing $\langle Tu, u\rangle$.

Theorem 34.1 (Min-Max Principle). *Let T be a compact Hermitian operator on an infinite dimensional Hilbert space.*

(1) Fischer's principle: If

$$\alpha_N = \max_{S_N} \min_{u\in S_N,\ \|u\|=1} \langle Tu, u\rangle,$$

where the max *is taken over all N dimensional subspaces S_N, then $\alpha_1 \ge \alpha_2 \cdots \ge 0$ and the non-zero entries of this sequence are the positive eigenvalues of T listed according to multiplicity.*

(2) Courant's principle: The above sequence is also given by

$$\alpha_N = \min_{S_{N-1}} \max_{u\perp S_{n-1}.\ \|u\|\le 1} \langle Tu, u\rangle,$$

where the max *is taken over all $N-1$ dimensional subspaces S_{N-1}.*

Remark. (1) In finite dimensions, the sequence is no longer non-positive and gives all the eigenvalues. In infinite dimensions, the sequence α_j, being non-positive and decreasing, is either positive for all j or eventually 0. The zero entries may not represent eigenvalues.

(2) The negative eigenvalues $\mu_1 \le \mu_2 \le \cdot < 0$ can of course be found as

$$\mu_N = \min_{S_N} \max_{u\in S_N,\ \|u\|=1} \langle Tu, u\rangle = \max_{S_{N-1}} \min_{u\perp S_{N-1},\ \|u\|\le 1} \langle Tu, u\rangle.$$

Proof. First note that $\alpha_N \ge 0$. Indeed, for any $\delta > 0$ there are only finitely many eigenvalues $\lambda < -\delta$ (counted according to multiplicity). Thus the orthogonal complement of the corresponding eigenvectors is infinite dimensional. On this subspace we have $\langle Tu, u\rangle \ge -\delta \|u\|^2$, so $\alpha_N \ge -\delta$.

Let $Ax_n = \lambda_n x_n$ be the positive eigenvalues and eigen-vectors listed according to multiplicity with $\lambda_1 \geq \lambda_2 \geq \cdots > 0$. (The sequence may terminate.) If there are M positive eigenvalues and $N > M$, let S_N be an N-dimensional subspace. By finite dimensional linear algebra, we can find a unit vector $u \in S_N$ perpendicular to $x_1, \ldots, x_M$:

$$\langle u, x_n \rangle = 0, \quad n = 1, \ldots, M.$$

By ($\star$) we must have $\langle Tu, u \rangle \leq 0$ so $\alpha_N \leq 0$. On the other hand $\alpha_N \geq 0$ so $\alpha_N = 0$.

If there are at least N positive eigenvalues, then with $S_N = \operatorname{span}(x_1, \ldots x_N)$ we have

$$\min_{u \in S_N, \|u\|=1} \langle Tu, u \rangle = \lambda_N,$$

so $\alpha_N \geq \lambda_N$. On the other hand if S_N is any N dimensional subspace then as above we may find a unit vector $u \in S_N$ which is perpendicular to the first $N-1$ eigenvectors. By ($\star$), $\langle Tu, u \rangle \leq \lambda_N$. Thus $\alpha_N \leq \lambda_N$.

Turning now to Courant's principle, let

$$\beta_N = \min_{S_{N-1}} \max_{u \perp S_{N-1},\ \|u\| \leq 1} \langle Tu, u \rangle.$$

Since $S_{N-1}^{\perp}$ is infinite dimensional and T is compact it follows that

$$\max_{u \perp S_{N-1},\ \|u\| \leq 1} \langle Tu, u \rangle \geq 0.$$

Thus $\beta_N \geq 0$. Also, note that

$$\beta_N \leq \min_{S_{N-1}} \min_{S_{N-2} \subset S_{N-1}} \max_{u \perp S_{N-2},\ \|u\| \leq 1} \langle Tu, u \rangle = \beta_{N-1}.$$

First suppose there are at least $N-1$ positive eigenvalues. It follows from ($\star$) that $\beta_N \leq \lambda_N = \alpha_N$, since $x_1, \ldots, x_{N-1}$ span an $N-1$ dimensional subspace.

If $\lambda_N = 0$, it follows that $\beta_M = 0 = \alpha_M$ for $M \geq N$.

On the other hand, suppose there are at least N positive eigenvalues and let S_{N-1} be an arbitrary $N-1$ dimensional subspace. I claim we may find a vector in the subspace spanned by $x_1, \ldots x_N$ perpendicular to S_{N-1}. Indeed, if y_j, $j = 1, \ldots, N-1$, is an orthonormal basis for S_{N-1} then we must solve

$$\sum_{k=1}^{N} \langle x_k, y_j \rangle a_k = 0,$$

for all j. Since $N-1 < N$ the matrix $\langle x_k, y_j \rangle$ has a non-trivial null space, and the resulting solution $u = \sum_k a_k x_k$ is perpendicular to S_{N-1}. But then

$$\langle Tu, u \rangle = \sum_{k=1}^{N} |a_k|^2 \lambda_k \geq \lambda_N \|u\|^2.$$

Thus $\beta_N \geq \lambda_N = \alpha_N$. □

The min-max principle is incredibly powerful. It is certainly one of the most important results in applications of functional analysis. Here is an example of what we can do with it:

Definition 34.1. Let $A, B\mathcal{L}(H)$ be Hermitian. We say that $A \geq B$ if

$$\langle Au, u \rangle \geq \langle Bu, u \rangle \quad \forall u \in H.$$

Remark. This is a partial order on the set of Hermitian operators.

Theorem 34.2. *Let $T, S \in \mathcal{C}(H)$ be Hermitian compact operators with $T \geq S$. If S has at least N positive eigenvalues $\lambda_1^+ \geq \cdots \geq \lambda_N^+ > 0$ then T has at least N positive eigenvalues $\mu_1^+ \geq \cdots \geq \mu_N^+ > 0$ and*

$$\mu_N^+ \geq \lambda_N^+.$$

Similarly, if T has at least N negative eigenvalues $\mu_1^- \leq \cdots \leq \mu_N^- < 0$ then S has at least N negative eigenvalues $\lambda_1^- \leq \cdots \leq \lambda_N^- < 0$ and

$$\lambda_N^- \leq \mu_N^-.$$

This theorem is an easy consequence of min-max. It is immediate if S and T have the same eigenvectors, but that is not at all necessary for the relation $S \leq T$. For instance

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \leq \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Similarly we have

Theorem 34.3. *Let $T \in \mathcal{C}(H)$ be compact Hermitian. If $T \geq 0$ then all eigenvalues of T are ≥ 0.*

In fact, we have

Theorem 34.4. *Let $T \in \mathcal{L}(H)$ be Hermitian. If $T \geq 0$ then $\sigma(T) \subset [0, \infty)$.*

Proof. Let $\lambda > 0$. Then

$$\|(T+\lambda)u\|^2 = \|Tu\|^2 + 2\lambda\langle u, Tu\rangle + \lambda^2 \|u\|^2 \geq \lambda^2 \|u\|^2.$$

It follows that $N_{T+\lambda} = \{0\}$ and that $\operatorname{ran}(T+\lambda)$ is closed. Since $\operatorname{ran}(T+\lambda)^\perp = N_{T+\lambda} = \{0\}$ it follows that $\operatorname{ran}(T+\lambda) = H$. Thus $T+\lambda$ is one-to-one, onto and bounded. By the inverse mapping theorem (27.5 in these notes) $(T+\lambda)^{-1}$ is bounded. Thus $-\lambda \notin \sigma(T)$. $\square$

LECTURE 35

Functional calculus and polar decomposition

The spectral theorem states that a compact Hermitian operator T is of the form:

$$T \;=\; \sum_{n=1}^{\infty} \lambda_n \langle \cdot, \phi_n \rangle \phi_n. \tag{$\star$}$$

Here λ_n are the eigenvalues and ϕ_n is an orthonormal basis. The notation indicates that

$$T\phi \;=\; \sum_{n=1}^{\infty} \lambda_n \langle \phi, \phi_n \rangle \phi_n.$$

In fact, since $\lambda_n \to 0$ the sum in $(\star)$ is absolutely convergent:

$$\left\| T - \sum_{n=1}^{N} \lambda_n \langle \cdot, \phi_n \rangle \right\| = \sup_{n>N} |\lambda_n| \;\longrightarrow\; 0.$$

If $f : \sigma(T) \to \mathbb{C}$ is a bounded function we define

$$f(T) \;=\; \sum_{n=1}^{\infty} f(\lambda_n) \langle \cdot, \phi_n \rangle \phi_n,$$

so

$$f(T)\phi \;=\; \sum_{n=1}^{\infty} f(\lambda_n) \langle \phi, \phi_n \rangle \phi_n.$$

It is an easy calculation to see that if $f(x) = p(x)$ is a polynomial in x then $f(T)$ defined this way agrees with plugging T into the polynomial. For instance,

$$\sum_{n=1}^{\infty} \lambda_n^2 \langle \phi, \phi_n \rangle \phi_n = \sum_{n=1}^{\infty} \lambda_n \langle \sum_{m=1}^{\infty} \lambda_m \langle \phi, \phi_m \rangle \phi_m, \phi_n \rangle = T^2 \phi,$$

by the pairwise orthogonality of ϕ_m. The map $f \mapsto f(T)$ is called the <u>functional calculus</u> for T and has the following properties:

Theorem 35.1. *Let T be a compact Hermitian operator. To every bounded function $f : \sigma(T) \to \mathbb{C}$ we assign a unique operator, $f(T)$, such that*

(1) $f(\sigma) \equiv 1 \implies f(T) = \mathbf{I}$.
(2) $f(\sigma) = \sigma \implies f(T) = T$.
(3) The map $f \mapsto f(T)$ is an injective homomorphism of the ring of bounded functions on $\sigma(T)$ into the $\mathcal{L}(H)$:

$$(f+g)(T) = f(t) + g(T), \quad (fg)(T) = f(T)g(T), \quad \textit{and } f(T) \equiv 0 \textit{ iff } f \equiv 0 \textit{ on } \sigma(T).$$

(4) $f(T)^\dagger = f^\star(T)$.

(5) The map is an isometry:

$$\|f\|_\infty = \sup_{\lambda\in\sigma(T)} |f(\lambda)| = \|f(T)\|.$$

This map has the properties

(6) If $f : \sigma(T) \to \mathbb{R}$ then $f(T)$ is Hermitian.
(7) If $f : \sigma(T) \to [0,\infty)$ then $f(T) \geq 0$.
(8) If $f : \sigma(T) \to \{|z| = 1\}$ then $f(T)$ is a unitary map, that is an isometry of H onto H.
(9) If $\lim_{\lambda\to 0} f(\lambda) = f(0) = 0$ then $f(T)$ is compact.

Remark. Properties (3-5) show that the map $f \mapsto f(T)$ is an isomorphism of $C^\star$ algebras, something we haven't defined yet but that we will see in the fall.

Proof. We already defined the map we will show it is unique in a moment. It is clear that (1) and (2) hold. It is easy to see that $(f+g)(T) = f(T) + g(T)$ and the argument given for T^2 above extends to a product fg. Thus (3) holds. To see that (4) holds note that

$$\langle f(T)^\dagger u, v\rangle = \langle u, f(T)v\rangle = \left\langle u, \sum_n f(\lambda_n)^\star \langle v, \phi_n\rangle \phi_n \right\rangle$$
$$= \sum_n f(\lambda_n)^\star \langle u, \phi_n\rangle\langle \phi_n, v\rangle = \left\langle \sum_n f(\lambda_n)^\star \langle u, \phi_n\rangle \phi_n, v\right\rangle = \langle f^\star(T)u, v\rangle.$$

Since

$$\|f(T)u\|^2 = \sum_n |f(\lambda_n)|^2 |\langle u, \phi_n\rangle|^2 \leq \|f\|_\infty \|u\|^2$$

and

$$f(T)\phi_n = f(\lambda_n)\phi_n,$$

(5) follows. (6) and (7) are easy calculations. To see (8) note that if $|f(x)|^2 = 1$ then

$$\|f(T)u\|^2 = \sum_n |\langle u, \phi_n\rangle|^2 = \|u\|^2.$$

(9) is an easy exercise.

Finally to see that the map is unique, note that (1), (2), and (3) specify the map for polynomials $p(x)$. Since T is compact, $\sigma(T)$ is discrete away from 0. Thus a bounded map on $\sigma(T)$ is continuous if and only if it is continuous at 0. Thus (5) and Stone-Weierstrass imply that the map for polynomials extends uniquely to bounded functions continuous at 0. In fact the map is uniquely defined on all bounded functions, but we will defer the proof of this to next term. □

Corollary 35.2. *If $T \geq 0$ and is compact Hermitian then there is a unique positive square root of T: $\sqrt{T} \geq 0$ and $\sqrt{T}^2 = T$.*

The square root is very useful as it allows us to define

Definition 35.1. Let T be a compact operator on a Hilbert space H. The absolute value of T, denoted $|T|$, is the operator $\sqrt{T^\dagger T}$.

The absolute value $|T|$ is compact (why?) so it has eigenvalues, which are all non-negative since $|T| \geq 0$. The eigenvalues of $|T|$ are called the singular values of T. Note that

$$\|T\phi\|^2 = \langle \phi, T^\dagger T\phi \rangle = \langle \phi, |T|^2 \phi \rangle = \||T|\phi\|^2 .$$

It follows that the map $|T|\phi \mapsto T\phi$ defined on the range of $|T|$ is a linear isometry. This extends to the closure of $\operatorname{ran}|T|$. Let this map be denoted U and define U to be zero on $\operatorname{ran}|T|^\perp = N_{|T|} = N_T$. Note that

$$T = U|T|.$$

THEOREM 35.3 (Polar decomposition). *Every compact operator T on a Hilbert space H may be factored, as*

$$T = UA$$

with $A \geq 0$ and $U^\dagger U \mid_{\operatorname{ran} A} = \mathbf{I}$. The map $A = |T|$ and U is uniquely determined if we specify $U \equiv 0$ on $\operatorname{ran} A^\perp$.

COROLLARY 35.4. *Any compact operator T on a Hilbert space has a singular value decomposition*

$$T = \sum_{n=1}^{\infty} \mu_n \langle \cdot, \phi_n \rangle \psi_n,$$

with $\mu_n \geq 0$, $\mu_n \to 0$ and ϕ_n, ψ_n (possibly distinct) orthonormal sequences.

PROOF. Let μ_n, ϕ_n be the eigenvalues/vectors of $|T|$. Let $\psi_n = U\phi_n$ with $T = U|T|$. $\square$

In fact, the functional calculus, square root and polar decomposition extend to non-compact operators. These are some theorems we will prove next term:

THEOREM 35.5. *Every positive operator $A \geq 0$ on a Hilbert space has a unique positive square root.*

THEOREM 35.6 (Polar Decomposition). *Every operator $T \in \mathcal{L}(H)$ may be factored as*

$$T = UA$$

with $A \geq 0$ and $U^\dagger U \mid_{\operatorname{ran} A} = \mathbf{I}$. The map $A = |T| = \sqrt{T^\dagger T}$ and U is uniquely determined if we specify $U \equiv 0$ on $\operatorname{ran} A^\perp = N_T$.

THEOREM 35.7 (Spectral Theorem: functional calculus version). *To every self adjoint operator $T \in \mathcal{L}(H)$ there is associated a unique injective $C^\star$ homomorphism from $C(\sigma(T)) \to \mathcal{L}(H)$ with the properties (1)-(8) above. Conversely, given any compact set $\sigma \subset \mathbb{R}$ and an injective $C^\star$ homomorphism with these properties there is a self adjoint operator T so that the homomorphism is realized as $f \mapsto f(T)$.*