Functional Analysis

Lecture Notes MTH 920 Spring 2022

Jeff Schenker

MICHIGAN STATE UNIVERSITY
Email address: schenke6@msu.edu

Contents

Comments and course information

These are lecture notes for Functional Analysis (Math 920), Spring 2022. The text for this course is:

- P. D. Lax. *Functional Analysis*. New York: Wiley, 2002

In some places I will 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 I: Functional Analysis*. San Diego, CA: Academic Press, 1980.
- W. Rudin. *Functional Analysis*. 2nd ed. International Series in Pure and Applied Mathematics. New York: McGraw-Hill, 1991.
- J. B. Conway. *A Course in Functional Analysis*. Vol. 96. Graduate Texts in Mathematics. New York, NY: Springer New York, 2007.

For background on undergraduate analysis, see:

- W. Rudin. *Principles of Mathematical Analysis*. 3rd ed. International Series in Pure and Applied Mathematics. McGraw-Hill, 1976.

For measure theory and complex analysis background, see:

- T. Tao. *An Introduction to Measure Theory*. Graduate Studies in Mathematics 126. Providence, RI: American Mathematical Soc., 2011.
- B. Simon. *Real Analysis*. Providence, Rhode Island: American Mathematical Society, Nov. 2015
- D. Sarason. *Complex Function Theory*. Second. Providence, RI: American Mathematical Society, 2007.
- G. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1999.
- W. Rudin. *Real and Complex Analysis*. 3rd ed. New York: McGraw-Hill, 1987.

Part 1

Linear Spaces and the Hahn Banach Theorem

LECTURE 1

Linear spaces, linear maps and the index

Reading: Chapters 1 and 2 of Lax

Linear Spaces

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

(1) $F[X]$ = set of polynomials over a field F.
(2) $C(M)$ = space of continuous functions ($\mathbb{R}$ or $\mathbb{C}$ valued) on a manifold M.
(3) $A(U)$ = space of analytic functions in a domain $U \subset \mathbb{C}$.
(4) $L^1(\mu) = \{$ equivalence classes of integrable functions on a measure space $\mathcal{M}, \mu$ modulo equality μ-almost everywhere $\}$.

The key features here are the axioms of linear algebra,

Definition 1.1. A *linear space* X over a field F (in this course we always take $F = \mathbb{R}$ 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*. That is, the operation $+$ is commutative and associative, and identity and inverses exist.
- The identity is called 0 ("zero").
- The 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 1.2. It follows from the axioms that $0x = 0$ and $-x = (-1)x$.

Let X be a linear space. Let us recall various facts and definitions from linear algebra:

- A set of vectors $S \subset X$ is *linearly independent* if

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

- The *dimension* of X, $\dim X$, is the *cardinality*[1] of a maximal linearly independent set in X (it is a theorem of linear algebra that all such sets have the same cardinality).

[1] Recall that cardinal numbers can be *finite* (0,1,2,3,...), *countable* (the cardinality of $\mathbb{N}$), or *uncountable* (any cardinal that is neither finite or countable, for example $\mathfrak{c}$ the cardinality of $\mathbb{R}$).

- 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 $\operatorname{span} S = X$.
- The dimension $\dim X$ is also the the cardinality of a *minimal spanning set*.

Definition 1.3. Given linear spaces X, Y and a mapping $T : X \to Y$, we say that T is *linear* if

$$T(x + ay) = T(x) + aT(y)$$

for all $x, y \in X$ and $a \in F$. A *linear isomorphism* is a linear mapping that is one-to-one and onto, and two linear spaces are *isomorphic* if there is a linear isomorphism between them.

It is a basic fact of linear algebra that two linear spaces are isomorphic if and only if they have the same dimension (try proving this without using finiteness of the dimension!). For example, regarding the example linear spaces above, $F[X]$ is the only space with countable dimension. Both $C(M)$ and $A(U)$ have uncountable dimension. The space $L^1(\mu)$ could be finite dimensional (if μ is supported on a finite set) or could have uncountable dimension.

What is functional analysis?

If you are only familiar with finite dimensional linear algebra, 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.[2] 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.

All that explains the "analysis" in "functional analysis." "Functional" is a somewhat 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" is the analysis of functions defined on function spaces.

Linear Maps

As we will see, not all linear maps are continuous (in infinite dimensions). By demanding that our linear maps be continuous we will obtain a much richer and more useful structure than what is provided by simple linear algebra. Nonetheless there a few concepts that can be introduced and studied in great generality, without topology.

There is a natural *algebra* on linear maps, inherited from the algebra on general maps:

- Given linear maps $T, S : X \to Y$, their sum $T + S$, defined by $(T + S)(x) = T(x) + S(x)$ is a linear map. Similarly, the scalar multiple aT is defined by $(aT)(x) = aT(x)$. These two operations make the space $\mathcal{L}(X, Y)$ of linear maps from X to Y into a linear space in its own right.

[2] The usual topology on the finite dimensional space F^n is a metric topology given by the Euclidean metric $d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$. However, this topology can also be characterized, in a completely algebraic fashion, as the minimal topology such that every linear map from F^n to F is continuous (when F is given its usual metric topology).

- Given linear maps $T : X \to Y$ and $S : Y \to Z$, the *product* ST is the composite map: $(ST)(x) = S(T(x))$. As in linear algebra we write this operation as a product, rather than using the composition symbol $\circ$.

Definition 1.4. A linear map $T : X \to Y$ is *invertible* if it is one-to-one and onto.

An invertible map has an inverse map $T^{-1} : Y \to X$ defined by $T^{-1}(y) = x$ if and only if $T(x) = y$.

Exercise 1.1. Verify that the inverse $T^{-1} : Y \to X$ of an invertible linear map is linear.

Definition 1.5. The *nullspace* of T (or kernel of T), denoted N_T or $\ker T$, is the set

$$\ker T \ := \ \{x \in X \mid Tx = 0\} \ .$$

The *range* of T (denoted $\operatorname{ran} T$) is the image of X under T:

$$\operatorname{ran} T \ := \ \{y \in Y\}\, y = Tx \quad \text{for some } x \in X \ .$$

Recall the following notions from linear algebra:

- A set $U \subset X$ is a *subspace* if it is a linear space under the operations inherited from X, that is if $x + ay \in U$ whenever $x, y \in U$ and $a \in F$.
- Given a subspace $U \subset X$, the *quotient space* X/U is the set of cosets $\{x + U \ : \ x \in X\}$, which is a linear space under the natural operations $(x + U) + (y + U) = (x + y) + U$ and $a(x + U) = ax + U$.

The following theorem lists several standard results. The proof is left as an exercise:

Theorem 1.6. *Let $T : X \to Y$ be a linear map.*

(1) $\ker T$ and $\operatorname{ran} T$ are linear subspaces of X and Y, respectively.
(2) T is invertible if and only if $\ker T = \{0\}$ and $\operatorname{ran} T = Y$.
(3) T maps the quotient $X/\ker T$ one-to-one onto $\operatorname{ran} T$.
(4) If T is invertible and $S : Y \to Z$ is an invertible linear map, then ST is invertible and $(ST)^{-1} = T^{-1}S^{-1}$.
(5) If ST is invertible, then $\ker T = \{0\}$ and $\operatorname{ran} S = Z$.

Exercise 1.2. Prove Theorem 1.6.

Note that it can happen (even in finite dimensions) that ST is invertible although T and S are not separately invertible. For example, consider the linear maps $T : \mathbb{R}^2 \to \mathbb{R}^3$ and $S : \mathbb{R}^3 \to \mathbb{R}^2$ given by the following matrices:

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \ .$$

However, if ST is invertible and $X = Y = Z$ *AND* X is finite dimensional, then S and T are separately invertible.

Similarly, *if X is finite dimensional* and $T : X \to X$ is linear, then

- if $\ker T = \{0\}$, then $\operatorname{ran} T = X$ and T is invertible,
- if $\operatorname{ran} T = X$, then $\ker T = \{0\}$ and T is invertible.

These facts are no longer true if X is infinite dimensional. For example, consider the space ℓ^∞ of all bounded, complex sequences. That is,

$$\ell^\infty := \left\{ (a_1, a_2, \ldots) \;\middle|\; a_j \in \mathbb{C} \quad \text{and} \quad \sup_j |a_j| < \infty \right\} .$$

Let R be the *right shift*,

$$R(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots) ,$$

and L be the *left shift*,

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

Formally R and L correspond to the infinite matrices

$$R = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ 0 & 1 & 0 & & \\ & 0 & 1 & 0 & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \quad \text{and} \quad L = \begin{bmatrix} 0 & 1 & 0 & & & \\ & 0 & 1 & 0 & & \\ & & 0 & 1 & 0 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix} .$$

Note that

- $\operatorname{ran} L = \ell^\infty$ but $\ker L = \operatorname{span}\{(1, 0, \ldots)\}$, so L is not invertible;
- $\ker R = \{0\}$ but $\operatorname{ran} R = \{(a_1, a_2, \ldots) \,:\, a_1 = 0\} \neq \ell^\infty$, so R is not invertible;
- $LR = I$ (the identity);
- but $RL(a_1, a_2, \ldots) = (0, a_2, \ldots)$.

So, *new and interesting things can happen in infinite dimensions!*

Pseudo-invertible maps

There is a a very general notion, *the index*, that quantifies the phenomenon we just saw for the right and left shifts. It applies to a special class of maps called *pseudoinvertible* maps.

Definition 1.7. A linear map K is *degenerate* if $\operatorname{ran} K$ is finite dimensional.

Note that *any linear* map on a finite dimensional space is degenerate. The degenerate maps have two imporant properties:

Theorem 1.8. *Let X, Y, U, V be linear spaces and let $\mathcal{K}(X, Y)$ denote the set of degenerate maps from X into Y. Then*

(1) $\mathcal{K}(X, Y)$ is a linear space, and
(2) if $K \in \mathcal{K}(X, Y)$, $T \in \mathcal{L}(U, X)$, and $S \in \mathcal{L}(Y, V)$, then SK and KT are degenerate.

Exercise 1.3. Prove Theorem 1.8

Definition 1.9. Let X, Y be linear spaces. Linear maps $T : X \to Y$ and $S : Y \to X$ are *pseudoinverses* of each other if there are degenerate maps K, K' such that

$$ST = I + K \quad \text{and} \quad TS = I + K' ,$$

where I denotes the identity map on the appropriate space (X or Y). We say that a map is *pseudoinvertible* if it has a pseudoinverse.

Exercise 1.4. Show that the right and left shifts defined above are pseudoiverses to each other.

Theorem 1.10. *(1) If T and S are pseudoiverses of each other, so are $T + K$ and $S + K'$ for arbitrary degenerate maps K, K'.*
(2) If $T : X \to Y$ and $S : Y \to Z$ have pseudoinverses $T' : Y \to X$ and $S' : Z \to Y$, respectively, then $T'S'$ is a pseudoiverse for ST.

Exercise 1.5. Prove theorem 1.10 .

Definition 1.11. Given a subspace $U \subset X$ of a linear space X, the *codimension* of U is the dimension of X/U :

$$\operatorname{codim} U := \dim(X/U) .$$

The following theorem characterizes pseudoinvertible maps:

Theorem 1.12. *A linear map $T : X \to Y$ is pseudoinvertible if and only if*

$$\dim \ker T < \infty \quad \textit{and} \quad \operatorname{codim} \operatorname{ran} T < \infty . \tag{1.1}$$

To prove Theorem 1.12, we will use two lemmas. The first is

Lemma 1.13. *If $K : X \to X$ is a degenerate linear map, then*

$$\dim \ker(I + K) < \infty \quad \textit{and} \quad \operatorname{codim} \operatorname{ran}(I + K) < \infty .$$

PROOF. If $x \in \ker(I + K)$ then $x = -Kx = K(-x) \in \operatorname{ran} K$. So $\ker(I + K) \subset \operatorname{ran} K$ is finite dimensional. By Theorem 1.6, K maps $X/\ker K$ one-to-one onto $\operatorname{ran} K$. Thus $\operatorname{codim} \ker K = \dim \operatorname{ran} K < \infty$. If $x \in \ker K$ then $(I + K)x = x$. Thus $\operatorname{ran}(I + K) \supset \ker K$, so $\operatorname{codim} \operatorname{ran}(I + K) \leq \operatorname{codim} \ker K < \infty$. □

We can use this lemma prove half of Theorem 1.12:

PROOF OF THE FORWARD DIRECTION OF THEOREM 1.12. Suppose that T has a pseudoinverse S. Then

$$\dim \ker T \leq \dim \ker ST \quad \text{and} \quad \operatorname{codim} \operatorname{ran} T \leq \operatorname{codim} \operatorname{ran} TS ,$$

since $\ker T \subset \ker ST$ and $\operatorname{ran} T \supset \operatorname{ran} TS$. Thus (1.1) follows from the following Lemma: □

For the reverse implication, we need a definition and a lemma to proceed.

Definition 1.14. Let $U, V \subset X$ be subspaces. We say that V is a *complementary* subspace to U (or that U and V are *complementary*) if $U \cap V = \{0\}$ and $U + V = \{u + v \ : \ u \in U \quad \text{and} \quad v \in V\} = X$.

Lemma 1.15. *Let $U \subset X$ be a linear subspace. Then there is a subspace V complementary to U.*

This proof is going to be unsatisfying. In the generality of the statement written here, there is no constructive proof. Instead we will use the *Kuratowski-Zorn lemma*, more commonly known as *Zorn's lemma*. This is the following result equivalent to the axiom of choice:

Theorem 1.16 (Kuratowski 1922 and Zorn 1935)**.** *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 1.17. A *partially ordered set* (poset) S is a set on which an order relation $a \leq b$ is defined on pairs $a, b \in S$, with the following properties

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

The relation $\leq$ is called a *partial order*. A subset T of S is *totally ordered* if any two elements of T are comparable, i.e.,

$$x, y \in T \implies x \leq y \text{ or } y \leq x.$$

An element $u \in S$ is an *upper bound* for $T \subset S$ if $x \leq u$ for all $x \in T$. A *maximal element* $m \in S$ satisfies $m \leq b \implies m = b$.

Note that it is *not* required that any two elements of a poset S are comparable: we may have, for given $a, b \in S$, $a \not\leq b$ and $b \not\leq a$. Similarly, a maximal element m need not be comparable to all elements of S – it just has the property that *if it is comparable* with a then $a \leq m$.

A proof of Zorn's lemma, based on certain facts about ordinal numbers, is included in an appendix. In these lectures, we will take it as given.

PROOF OF LEMMA 1.15. Consider the collection $\mathcal{Z}$ of all subspaces V such that $V \cap U = \{0\}$. Clearly $\mathcal{Z}$ is non-empty, since $\{0\} \in \mathcal{Z}$. We can partially order $\mathcal{Z}$ by taking $V \leq V'$ if $V \subset V'$. Given any totally ordered collection of elements of $\mathcal{Z}$, we find that the union of the elements is itself an element of $\mathcal{Z}$ (this is an elementary exercise) and thus an upper bound. By Zorn's lemma $\mathcal{Z}$ has a maximal element, i.e., there is a $V \in \mathcal{Z}$ such that if $Y \in \mathcal{Z}$ and $Y \supset V$ then $Y = V$. Suppose that $U + V \neq X$ and let $y \in X \setminus (U + V)$. It follows that $V + \{ay : a \in F\}$ is an element of $\mathcal{Z}$ strictly larger than V, contradicting the maximality of V. Thus we must have $U + V = X$. □

Remark 1.18. If U and V are complementary, we will write $U \oplus V = X$. It follows that each coset in X/U contains a unique element of V. Thus there is a linear isomorphism from X/U to V.

PROOF OF THE REVERSE DIRECTION OF THEOREM 1.12. Suppose that (1.1) holds. Let $U \subset X$ and $V \subset Y$ be subspaces such that $X = \ker T \oplus U$ and $Y = \operatorname{ran} T \oplus V$. Note that $\dim V = \operatorname{codim} \operatorname{ran} T < \infty$. By Theorem 1.6, T maps $X/\ker T$ one-to-one and onto $\operatorname{ran} T$. Since $X/\ker T$ is isomorphic to U we conclude that the restriction of T to U is an invertible map from U to $\operatorname{ran} T$. Thus every element of Y can be written uniquely as $Tu + v$ with $u \in U$ and $v \in V$. Define $S : Y \to X$ as follows, if $y = Tu + v$ as above, then

$$Sy := u .$$

Then $\ker S = V$, $\operatorname{ran} S = U$, and

$$ST = I - P \quad \text{and} \quad TS = I - Q ,$$

where $P(x + u) = x$ for $x \in \ker T$, $u \in U$, and $Q(y + v) = v$ for $y \in \operatorname{ran} T$, $v \in V$. Since Q and P are degenerate, it follows that S is a pseudoiverse for T. □

LECTURE 2

The Index; The Hahn Banach Theorem

Reading: Chapter 3 of Lax

The Index

Definition 2.1. Let $T : X \to Y$ be a pseudoinvertible map. The *index* of T is the integer

$$\operatorname{ind} T \;=\; \dim\ker T - \operatorname{codim}\operatorname{ran} T\,. \tag{2.1}$$

Note that if X and Y are finite dimensional, then $\operatorname{ind} T = \dim X - \dim Y$. Indeed, we have $\dim X = \dim\operatorname{ran} T + \dim\ker T$ by the "rank-nullity theorem." Thus

$$\operatorname{ind} T \;=\; \dim\ker T - \dim Y + \dim\operatorname{ran} T \;=\; \dim X - \dim Y\,.$$

On the other hand, in infinite dimensions the index is not just a function of the domain and target spaces, but actually depends on the map. Indeed, for the right and left shift defined above, we have

$$\operatorname{ind} R \;=\; -1 \quad\text{and}\quad \operatorname{ind} L \;=\; 1\,.$$

It is no accident that $\operatorname{ind} R + \operatorname{ind} L = 0$. Indeed, this follows from $LR = I$ and the following

Theorem 2.2. *If $T : X \to Y$ and $S : Y \to Z$ are pseudoinvertible linear maps, then ST has a pseudoinverse and*

$$\operatorname{ind}(ST) \;=\; \operatorname{ind} S + \operatorname{ind} T\,.$$

To prove the theorem we will use the following algebraic notion:

Definition 2.3. Consider a finite sequence $V_0, \dots, V_n$ of linear spaces with maps $T_j : V_j \to V_{j+1}$ for $j = 0, \dots, n-1$,

$$V_0 \xrightarrow{T_0} V_1 \xrightarrow{T_1} \cdots \xrightarrow{T_{n-1}} V_n\,.$$

This sequence is called *exact* if $\operatorname{ran} T_j = \ker T_{j+1}$ for $j = 0, \dots, n-2$.

Lemma 2.4. *Suppose that we have an exact sequence of finite dimensional vector spaces with* $\dim V_0 = \dim V_n = 0$. *Then*

$$\sum_{j=0}^{n} (-1)^j \dim V_j = 0\,.$$

PROOF. Decompose each V_j as $V_j = N_j \oplus Y_j$ with $N_j = \ker T_j$ (for $j = 0, \dots, n-1$). We take $N_n = \{0\}$. It follows that T_j is an isomorphism of Y_j with N_{j+1} for each $j = 0, \dots, n-1$. Thus

$$\dim V_j = \dim N_j + \dim Y_j = \dim N_j + \dim N_{j+1}\,, \quad j = 0, \dots, n-1$$

Hence, summing the relation from $j = 0$ to $n - 1$, with alternating signs, leads to

$$\sum_{j=0}^{n-1}(-1)^j \dim V_j = \sum_{j=0}^{n-1}(-1)^j \dim N_j - \sum_{j=1}^{n}(-1)^j \dim N_j = \dim N_0 - \dim N_n = 0,$$

since $N_0 = N_n = \{0\}$. □

PROOF OF THEOREM 2.2. To prove the theorem we use the exact sequence

$$0 \to \ker T \xrightarrow{J} \ker ST \xrightarrow{T} \ker S \xrightarrow{P} Y/\operatorname{ran} T \xrightarrow{S'} Z/\operatorname{ran} ST \xrightarrow{Q} Z/\operatorname{ran} S \to 0\,,$$

where

(1) J is the identification map $x \overset{J}{\mapsto} x$ (note that $\ker T \subset \ker ST$);
(2) P is the map $y \overset{P}{\mapsto} y + \operatorname{ran} T$;
(3) S' is the map $y + \operatorname{ran} T \overset{S'}{\mapsto} Sy + \operatorname{ran} ST$; and
(4) Q is the map $z + \operatorname{ran} ST \overset{Q}{\mapsto} z + \operatorname{ran} S$.

Exercise 2.1. Check that these maps are well defined and that this is indeed an exact sequence.

As a result, we have from Lemma 2.4

$$\begin{aligned} 0 &= \dim\ker T - \dim\ker ST + \dim\ker S - \operatorname{codim}\operatorname{ran} T + \operatorname{codim}\operatorname{ran} ST - \operatorname{codim}\operatorname{ran} S \\ &= \operatorname{ind} T + \operatorname{ind} S - \operatorname{ind} ST\,. \end{aligned}$$ □

One further important result is stability of the index under degenerate perturbations:

Theorem 2.5. *Let $T : X \to Y$ be a pseudoinvertible linear map and $K : X \to Y$ a degenerate linear map. Then $T + K$ is pseudoinvertible and*

$$\operatorname{ind}(T + K) = \operatorname{ind} T\,.$$

For a proof of this result, see Theorem 12 in Lax 2002.

The Hahn Banach Theorem

Definition 2.6. A *linear functional* is a linear map $\ell : X \to F$ from a linear space X over a field F to the field itself.

Given a linear space X it is not immediately obvious that any non-zero linear functionals actually exist. However, there is a very general logical principle, called the Hahn-Banach Theorem, that allows us to show that many linear functionals do exist. It is not a hard result, but it does rely on Zorn's lemma. The theorem gives a sufficient condition for a linear functional defined initially on a *linear subspace* $Y \subset X$ to be extended to all of X.

Theorem 2.7 (Hahn 1927 and 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 2.8. Let $X = B[0,1]$, the space of **all** bounded, real valued functions on $[0,1]$, and let $Y = C[0,1]$, the space of continuous, real valued functions on $[0,1]$. Let $\ell(f) = \int_0^1 f(t)dt$ for $f \in Y$, where the integral is the *Riemann integral*:

$$\ell(f) \ = \ \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(k/n).$$

Let $p : B \to \mathbb{R}$ be $p(f) = \sup\{|f(x)| \ : \ x \in [0,1]\}$. Then p satisfies (1) and (2) and $\ell(f) \leq p(f)$. Thus we can extend ℓ to *all of* $B[0,1]$.

We will return to this example later and see that we can extend ℓ so that $\ell(f) \geq 0$ whenever $f \geq 0$. With this extension we can define a set function $\mu(S) = \ell(\chi_S)$ on arbitrary subsets $S \subset [0,1]$. Since ℓ is linear, μ is *finitely additive*: $\mu(S \cup T) = \mu(S) + \mu(T)$ if $S \cap T = \emptyset$. As you might expect, for Lebesgue measurable S, $\mu(S)$ is nothing other than the Lebesgue measure of S. However, *that* does not follow from Hahn-Banach. This is typical of applications of Hahn-Banach. The Hahn-Banach theorem usually accomplishes the *soft* part of an argument, but in a way that doesn't give us much information without further analysis.

The proof of Hahn-Banach is *not* constructive, but relies on Zorn's lemma:

PROOF OF THEOREM 2.7. To apply Zorn's Lemma, we use the following 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 linear subspace $Y' \subset X$, such that

(1) $Y \subset Y' \subset X$,
(2) $\ell'(y) = \ell(y)$ for all $y \in Y$, and
(3) $\ell'(y) \leq p(y)$ for all $y \in Y'$.

We partially order S as follows

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

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, this 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 $(\ell', Y') \in S$ has an extension dominated by p whenever $Y' \neq X$. Fortunately, this reduces to a finite dimensional problem! Suppose $Y' \neq X$ and let $x_0 \in X \setminus Y'$. Let

$$Y'' \ = \ \{ax_0 + y \mid y \in Y, \ a \in \mathbb{R}\}.$$

To define an extension ℓ'' of ℓ' on Y'', we need only define $\ell''(x_0)$. If we didn't care about dominating ℓ'' by p, we could choose $\ell''(x_0)$ as we like. However, to guarantee that ℓ'' is dominated by p, we must have

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

for all $a \in \mathbb{R}$ and $y \in Y'$. Since Y' is a subspace and p is homogeneous, this is the same as

$$\pm\ell''(x_0) \leq p(y \pm x_0) - \ell'(y) \tag{2.2}$$

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

$$\ell'(y') - p(y' - x_0) \leq p(x_0 + y) - \ell'(y) \text{ for all } y, y' \in Y', \tag{2.3}$$

or equivalently

$$\ell'(y' + y) \leq p(x_0 + y) + p(y' - x_0) \text{ for all } y, y' \in Y'. \tag{2.4}$$

Since

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

eq. (2.4) holds. Thus a choice of $\ell'(x_0)$ satisfying eq. (2.2) exists. □

If $\operatorname{codim} Y < \infty$ has *finite co-dimension*, then one can follow the second paragraph of the proof to give a constructive proof of the Hahn-Banach theorem involving only finitely many choices. However, when Y has infinite co-dimension, the approach via Zorn's lemma is necessary and typically involves uncountably many "choices."

Geometric Hahn-Banach Theorems

One key use of the Hahn-Banach Theorem is to understand something of the geometry of linear spaces. In particular, we want to understand if the following picture holds in infinite dimension: Let X be a linear space over $\mathbb{R}$.

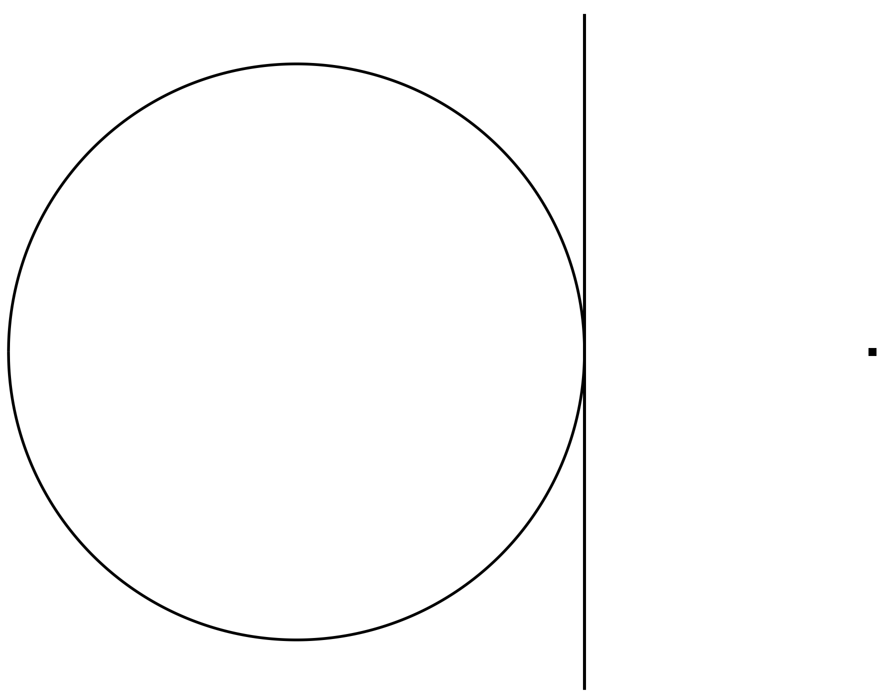

FIGURE 2.1. Separating a point from a convex set by a ~~line~~ hyperplane

Definition 2.9. 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 2.10. A point $x \in S \subset X$ is an *interior point* of S if for every $y \in X$ there is $\varepsilon > 0$ such that $|t| < \varepsilon \implies x + ty \in S$.

Remark 2.11. We could define a topology using this notion, letting $U \subset X$ be open whenever all $x \in U$ are interior. From the standpoint of abstract linear algebra this seems to be "the" 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 consideration. Much of functional analysis centers around the interplay of different topologies.

We are aiming at the following

Theorem 2.12. *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 there is a linear functional $\ell : X \to \mathbb{R}$ s.t.*

$$\ell(x) \leq \ell(y) \text{ for all } x \in K, \tag{2.5}$$

with strict inequality for all interior points x of K.

This is the "hyperplane separation theorem." It 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 2.13. 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

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

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

Lemma 2.14. *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 $p_K(x+y) \leq a + b$. Optimizing over a, b we obtain the result. □

PROOF OF HYPERPLANE SEPARATION THEOREM. It suffices to assume $0 \in K$ and is interior. 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 of K (why?), so

$$p_K(x) < 1 \iff x \text{ if an interior point of} K.$$

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

$$\ell(ay) \;\leq\; p_K(ay)$$

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 eq. (2.5). □

An extension of this is the following

Theorem 2.15. *Let H, M be disjoint convex subsets of X, at least one of which has an interior point. Then H and M can be separated 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 2.16 (Complex Hahn-Banach: Bohnenblust 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

$$|\ell(y)| \leq p(y) \tag{2.6}$$

for all $y \in Y$. Then ℓ can be extended to all of X so that (2.6) *holds for all $y \in X$.*

Remark 2.17. 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

$$\ell(y) = \ell_1(y) - \mathrm{i}\ell_1(\mathrm{i}y). \tag{2.7}$$

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 (2.7). 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,

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

Lax gives another beautiful extension of Hahn-Banach, due to Agnew and Morse, which involves a family of commuting linear maps.

Part 2

Banach Spaces, including Hilbert Spaces

LECTURE 3

Norms and Banach Spaces

Reading: Chapter 5 of Lax.

The Hahn-Banach theorem made use of a dominating function $p(x)$. When this function is positive for $x \neq 0$, it can be understood roughly as a kind of "distance" from a point x to the origin. Such a function is called a *norm*:

Definition 3.1. 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) $\|x\| = 0 \iff x = 0$ (non-degeneracy),
(2) $\|x + y\| \leq \|x\| + \|y\|$ (sub-additivity), and
(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 $\|\cdot\|$.

On any normed space we define an associated metric

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

Exercise 3.1. Show that d is a metric. Note that this proof uses only non-degeneracy and sub-additivity.

Thus any normed space is a metric space and the metric is easily seen to be

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

Associated to the metric, is the metric topology. In particular we have the following notions:

(1) a sequence $(x_n)_{n=1}^{\infty}$ *converges to* x, denoted $x_n \to x$, 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

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

Exercise 3.2. Show that equivalent norms define the same topology. That is, they generate the same family of open sets and the same family of convergent sequences.

Recall from real analysis that a metric space X is *complete* if every *Cauchy* sequence $(x_n)_{n=1}^{\infty}$ converges in X. In a normed space, a *Cauchy* sequence $(x_n)_{n=1}^{\infty}$ 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*. Banach spaces were introduced by Banach 1922. Fréchet 1926 coined the term *Espace de Banach* (Banach space).

Not every normed space is complete, and whether or not it is complete may depend on the norm. For example $C[0,1]$ with the norm

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

fails to be complete, although it is complete with respect to 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.

1. Examples of Normed and Banach spaces

Example 3.2. For each $p \in [1, \infty)$ let

$$\ell^p \;= \{p \text{ summable sequences}\} \;=\; \left\{ (a_1, a_2, \dots) \;\middle|\; \sum_{j=1}^{\infty} |a_j|^p < \infty \right\}.$$

Define a norm on ℓ^p via

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

Then ℓ^p is a Banach space. Note that $\boldsymbol{a} \in \ell^1$ is *summable*, i.e., $\sum_{j=1}^{\infty} a_j$ is convergent and

$$\left| \sum_{j=1}^{\infty} a_j \right| \;\leq\; \|\boldsymbol{a}\|_1.$$

Example 3.3. Let

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

with norm

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

Then ℓ^∞ is a Banach space.

Example 3.4. Let

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

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

Example 3.5. 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.

Example 3.6. Let $D \subset \mathbb{R}^d$ be a domain and let $p \in [1, \infty)$. 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.

Example 3.7. Let $D \subset \mathbb{R}^d$ be a domain and let $p \in [1, \infty)$. Let X denote the set of C^1 functions on D such that

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

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.

Separable Spaces

Definition 3.8. 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. For example,

(1) ℓ^p is separable for $1 \leq p < \infty$.
(2) ℓ^∞ is 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) $L^\infty(D)$ is not separable.
(4) $L^p(D)$ is separable for $1 \leq p < \infty$.
(5) The space M of all signed (or complex) measures μ on, say, D with norm

$$\|\mu\| \;=\; \int_D |\mu|(\mathrm{d}x)$$

is a non-separable Banach space. Here $|\mu|$ denotes the *total variation* of μ,

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

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.

Noncompactness of the Unit Ball

Theorem 3.9. *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.*

The fact that the unit ball is compact if X is finite dimensional is the Heine-Borel Theorem from Real Analysis. To prove the converse, we use the following

Lemma 3.10 (Riesz 1916)**.** *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. Since Y is proper, there is $x \in X \setminus Y$. As Y is closed and $x \notin Y$,

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

There may not be a minimizing y, but we can certainly find y_0 such that

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

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

$$\|z - y\| = \frac{\|x - y_0 - \|x - y_0\| y\|}{\|x - y_0\|} > \frac{d}{2d} = \frac{1}{2}$$

for $y \in Y$. □

PROOF THAT THE UNIT BALL IN AN INFINITE DIMENSIONAL SPACE IS NOT COMPACT. It suffices to show that if X is infinite dimensional then there is a sequence in $B_1(0)$ with no convergent subsequence.[1]

Let y_1 be any unit vector and recursively define a sequence of unit vectors 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. So the Lemma guarantees the existence of y_k. Since X is infinite dimensional the process never stops. No subsequence of y_j can be Cauchy, much less convergent. □

[1]Since $B_1(0)$ is a metric space, it is compact if and only if it is sequentially compact, namely if and only if every sequence in $B_1(0)$ has a convergent subsequence.

LECTURE 4

Uniform Convexity and Bounded Linear Maps

Reading: Chapter 5 of Lax.

Uniform convexity

The following theorem may be easily shown using compactness:

Theorem 4.1. *Let $X = \mathbb{R}^n$ with the usual Euclidean norm. 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*

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

Exercise 4.1. Prove this theorem. (Hint: existence of a minimizer follows from compactness; 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 4.2. 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) $\left\| \frac{1}{2}(x+y) \right\| \le 1 - \epsilon(\|x - y\|)$ for all $x, y \in B_1(0)$, the unit ball of X.

Theorem 4.3 (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

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

for suitable $t \in (0,1)$. By convexity, $t y_n + (1-t) y_m \in K$ so

$$\|t y_n + (1-t) y_m\| \ge s.$$

Thus

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

as $n, m \to \infty$. 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 $\|y_0\| = s$. □

Exercise 4.2. Show that the minimizer y_0 found above is unique.

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. In fact we can have

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

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 $K = \{\text{functions that vanish on } [-1,0]\}$. 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 $0 \leq g(x) \leq 1$.

On the other hand, many familiar Banach spaces are uniformly convex. For example, ℓ^p with $1 < p < \infty$ and $L^p(D)$ with D any measure space and $1 < p < \infty$ are uniformly convex.

Bounded Linear maps

Definition 4.4. 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 4.5. *A linear map $T : X \to Y$ between normed spaces X and Y is continuous if and only if it is bounded.*

PROOF. First suppose that $T : X \to Y$ is bounded. If T is the zero map, then it is constant, hence continuous. If T is not the zero map, then $\|T\| > 0$. Let $x \in X$ and let $\epsilon > 0$. If $\|x' - x\| < \frac{\epsilon}{\|T\|}$, then

$$\big\|T(x') - T(x)\big\| \;=\; \big\|T(x' - x)\big\| \;\leq\; \|T\|\big\|x' - x\big\| < \epsilon\,.$$

Thus T is continuous at x.

On the other hand, suppose that $T : X \to Y$ and that $\|T\| = \infty$. Then for each n there is $x_n \neq 0$ such that

$$\frac{\|T(x_n)\|}{\|x_n\|} \geq n\,.$$

Let

$$x'_n = \frac{1}{\sqrt{n}} \frac{1}{\|x_n\|} x_n .$$

So $\|x'_n\| = \frac{1}{\sqrt{n}}$ and $\lim_{n\to\infty} x'_n = 0$. But

$$\|T(x'_n)\| = \frac{1}{\sqrt{n}} \frac{\|T(x_n)\|}{\|x_n\|} \geq \sqrt{n} .$$

Thus T is not continuous at 0. By linearity, we conclude that T is discontinuous everywhere. □

Exercise 4.3. 1) Verify that the norm defined above is a *norm* on the space $\mathcal{B}(X,Y)$ of bounded linear maps from X to Y. 2) Let $T_1 \in \mathcal{B}(X,Y)$ and $T_2 \in \mathcal{B}(Y,Z)$ prove that $\|T_2 \circ T_1\| \leq \|T_2\|\|T_1\|$. Find an example to show that the inequality can be strict (hint use matrices).

Isometries

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 4.6 (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 4.7. 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 do this we will use the following:

Exercise 4.4. Let $M : X \to Y$ be a continuous map between normed spaces 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.

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

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

We need to show

$$2z' \;=\; x' + y'. \tag{$\star$}$$

Since M is an isometry, it follows that

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

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 *strictly sub-additive*, i.e., if

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

In general, however, there may be a number of points "half-way between x' and y'." This happens, for example in $C(X)$ or $L^1(X)$.

Let

$$A_1 = \left\{ u \in X \;\middle|\; \|x - u\| = \|y - u\| = \frac{1}{2}\|x - y\| \right\},$$

and

$$A_1' = \left\{ u' \in Y \;\middle|\; \|x' - u'\| = \|y' - u'\| = \frac{1}{2}\|x' - y'\| \right\}.$$

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

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

which is the set of "centers of A_1." If $v \in A_1$ then $2z - v \in A_1$:

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

Thus

$$d \;\geq\; \|2z - v - v\| \;=\; 2\|z - v\|$$

and so $z \in A_2$. Similarly, let

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

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

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

and

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

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

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

Since $z' \in A'_j$ for all j, $(\star)$ follows. □

LECTURE 5

Scalar Products and Hilbert Spaces

Reading: Chapter 6 of Lax.

Definition 5.1. 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: $y \mapsto \langle x, y \rangle$ is linear and $x \mapsto \langle x, y \rangle$ is *skewlinear*,

$$\langle x + ax', y \rangle = \bar{a} \langle x, y \rangle + \langle x', y \rangle .$$

(2) Skew symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle}$.
(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 5.2. 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, suppose $(\cdot, \cdot)$ is a real inner product on a complex linear space; if $(\cdot, \cdot)$ satisfies $(\star)$, then $(\star\star)$ defines 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 sub- additivity we need the following important

Theorem 5.3 (Cauchy-Schwarz Inequality)**.** *A real or complex scalar product satisfies*

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

with equality only if $ax = by$.

Remark 5.4. A corollary is that

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

from which follows sub-additivity

$$\|x+y\| \leq \|x\| + \|y\|.$$

PROOF. It suffices to consider the real case, since given x, y we can always find θ so that $\langle x, \mathrm{e}^{\mathrm{i}\theta}y\rangle = \mathrm{e}^{\mathrm{i}\theta}\langle x, y\rangle$ is real. Also, we may assume $y \neq 0$, as otherwise both sides of the inequality are zero.

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

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

The Cauchy-Schwarz inequality follows. □

Another important, related result, is the *parallelogram identity*

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

A result of Jordan and von Neumann 1935 states that any norm which satisfies the parallelogram law comes from an inner product.

Definition 5.5. A linear space with a scalar product that is complete in the induced norm is a *Hilbert space*.

That is a Hilbert space is a Banach space with a norm that comes from an inner product.

Any scalar product space can be completed in norm. It follows from the Schwarz inequality that the scalar product is continuous 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

$$\langle a, b\rangle = \sum_j \overline{a_j} b_j,$$

which is finite by Hölder's inequality.

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

$$\langle f, y\rangle = \int_0^1 \overline{f(t)} g(t)\mathrm{d}t.$$

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

Remarks. 1) 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 sometimes the second. 2) Hilbert 1906 used inner products in his study of linear integral equations. 3) The abstract definition of a Hilbert space, and the name, are due to von Neumann 1927 in his work on the foundations of quantum mechanics.

Orthogonal Projection

Definition 5.6. Two vectors in an inner product space are *orthogonal* if

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

The *orthogonal complement* of a set s is

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

Lemma 5.7. *Let H be a Hilbert space. If $S \subset H$, then $S^{\perp}$ 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

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

Theorem 5.8. *Let H be a Hilbert space. If Y is a closed subspace of H, then*

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

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

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

To prove this theorem, we need

Lemma 5.9. *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 from Clarkson's Theorem 4.3 if we show that H is uniformly convex. 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

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

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

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

for any t. In other words the function

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

which can occur only if 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 $z = 0$. The proof that $(Y^{\perp})^{\perp} = Y$ is left as a simple exercise. □

Riesz-Fréchet Theorem

We have already seen that for fixed $y \in H$, a Hilbert space, the map $\ell_y(x) = \langle y,\, x\rangle$ 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. Since

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

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*:

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

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

Theorem 5.10 (Riesz Representation Theorem). *Let $\ell(x)$ be a bounded linear functional on a Hilbert space H. Then there is a unique $y \in H$ such that*

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

Remark. The theorem in the abstract context is due to Riesz 1934. The result is sometimes called the "Riesz-Fréchet" theorem because of earlier work by Riesz 1907b and Fréchet 1907 in the context of function spaces.

Before turning to the proof, let us state several basic facts whose proofs are left as an exercise.

Lemma 5.11.

(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 OF THE RIESZ-FRÉCHET THM. 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 $Y^{\perp}$. The point $\hat{y}$ is unique up to a scalar multiple. Then $m(x) = \langle \hat{y},\, x\rangle$ is a linear functional, with null-space Y. Thus $\ell = \alpha m$ and we may take $y = \overline{\alpha}\hat{y}$.

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

LECTURE 6

Lax-Milgram Theorem and the Geometry of a Hilbert space

Reading: §6.3 of Lax.

Lax-Milgram Theorem

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

Definition 6.1. Let H be a Hilbert space over $\mathbb{R}$. A *bilinear 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

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

A bilinear or skew-linear form B on H is *bounded* if there is a constant $c > 0$ such that

$$|B(x, y)| \leq c\|x\|\|y\| ,$$

and is *bounded below* if

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

Theorem 6.2 (Lax and Milgram 1954)**.** *Let H be a Hilbert space over $\mathbb{R}$, or $\mathbb{C}$, and let B be a bounded bilinear, or skew-linear, form on H that is bounded from below. Then every bounded linear functional $\ell \in H^\star$ may be written*

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

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

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

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,

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

so

$$b\|y\|^2 \leq \|y\|\|z(y)\|,$$

and thus

$$b\|y\| \leq \|z(y)\|.$$

If y_n is any sequence then

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

Hence, if $z(y_n) \to z_0$, then y_n is Cauchy so $y_n \to y_0$ 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

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

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-Fréchet that any linear functional ℓ may be written $\ell(x) = \langle z(y),\, x\rangle = B(y,x)$ for some y. Uniqueness of y follows as above, since if $B(y,x) = B(y',x)$ for all x we conclude that $\|y - y'\| = 0$ since B is bounded from below. □

An Application of Riesz-Fréchet and Lax Milgram

The Reisz-Fréchet 5.10 and Lax Milgram 6.2 theorems can be used to give a simple proof of the existence of weak solutions to parabollic equations. The following is a sketch of the ideas. For more details, see Chapter 7 of Lax.

Let $D \subset \mathbb{R}^n$ be a bounded, connected, open set. To begin consider the Dirichlet problem for the Laplacian

$$-\Delta u = f \quad \text{ on } D, \quad u(x) = 0 \quad \text{ on } \partial D.$$

Let $H_0^1(D)$ be the homogeneous Sobolev space of first order over D. One can introduce this as the Hilbert space obtained by completing the smooth, compactly supported functions $C_c^\infty(D)$ in the norm

$$\|u\|_1 \;=\; \left(\sum_{j=1}^n \int_D \left|\partial_j u(x)\right|^2 ds \right) .$$

The following Lemma is often called the Poincare Lemma; Poincare 1890 proved a version of it for convex domains. The result as stated here is due to Zaremba 1909:

Lemma 6.3 (Zaremba 1909)**.** *For $u \in H_0^1(D)$ one has*

$$\|u\|_0 \;\le\; d\|u\|_1$$

where $\|u\|_0 = \left(\int_D |u(x)|^2 dx\right)$ is the L^2 norm and $d = \operatorname{diam}(D)$.

PROOF. By taking limits, it suffices to prove the inequality for $u \in C_c^\infty(D)$. Let $x \in D$ and let $\boldsymbol{e}$ be unit vector in $\mathbb{R}^n$. Let d_x be the smallest number such that $x + d_x e \in \partial D$. Then

$$u(x) \;=\; -\int_0^{d_x} \boldsymbol{e} \cdot \nabla u(x + te)dt \,.$$

Applying Cauchy-Schwarz yields

$$|u(x)|^2 \;\le\; d_x \int_0^{d_x} |\nabla u(x+te)|^2 dx.$$

Integrating of D yields and noting that $d_x \le d$ yields

$$\|u\|_0^2 \;\le\; d \int_0^d \int_D |\nabla u(x+te)|^2 dx \;\le\; d^2 \|u\|_1^2 \,. \quad \square$$

Given $f \in L^2(D)$, let $\ell_f(v)$ denote the linear functional

$$\ell_f(v) \;=\; \int_D f(x)v(x)dx \,.$$

By Cauchy-Schwarz and Zaremba's lemma, ℓ_f *bounded* on H_0^1:

$$\left|\ell_f(v)\right| \;\le\; \|f\|_0\|v\|_0 \;\le\; d\|f\|_0\|v\|_0 \,.$$

Thus by the Riesz-Fréchet Theorem, there is $u \in H_0^1$ such that

$$\ell_f(v) \;=\; \langle u,\, v\rangle \;=\; \int_D \nabla u(x)\cdot\nabla v(x)dx$$

for all $v \in H_0^1$. This u is the weak solution to the Dirichlet problem that we wished to find.

By using the Lax-Milgram Theorem, we can extend the above result to divergence form elliptic problems such as

$$-\nabla\cdot \boldsymbol{A}(x)\nabla u(x) + \phi(x)u(x) = f(x)\,, \quad x \in D\,, \tag{6.1}$$

with $u(x) = 0$ for $x \in \partial D$, provided

(1) $\boldsymbol{A}(x)$ is a positive definite $n \times n$ matrix that is bounded and uniformly elliptic

$$0 \;<\; \mu \;:=\; \inf_{x\in D}\lambda_1(x) \;\le\; \sup_{x\in D}\lambda_n(x)\,,$$

where $\lambda_1(x) \le \lambda_n(x)$ are the smallest and largest eigenvalues of $\boldsymbol{A}(x)$, and

(2) $-\frac{\mu}{d} \;<\; \inf_{x\in D}\phi(x) \;\le\; \sup_{x\in D}\phi(x) \;<\; \infty$.

The key is to define the bilinear form

$$B(u,v) \;=\; \int_D \left\{\nabla u(x)\cdot \boldsymbol{A}(x)\cdot\nabla v(x) + \phi(x)u(x)v(x)\right\}dx\,.$$

Exercise 6.1. Show that $B(u,v)$ is bounded on H_0^1, i.e., $|B(u,v)| \le c\|u\|_1\|v\|_1$, and positive, i.e., $B(u,u) \;\ge\; c\|u\|_1^2$. Hint: as in Zaremba's lemma it suffices to prove the bounds for $u, v \in C_c^\infty(D)$ and take limits. One needs the uniform ellipticity of $\boldsymbol{A}$ to get positivity.

Thus by Lax-Milgram, we can write the linear funtional $\ell_f(v)$ as

$$\ell_f(v) \;=\; B(u,v)$$

for some unique $u \in H_0^1$. This u is the weak solution to (6.1) that we seek.

Geometry of Hilbert Space

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:

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

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:

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

Proposition 6.4. *Let X be a Banach space. Then* $\text{c-span}\, S = \overline{\text{span}\, S}$.

The proof is left as an exercise.

In a Hilbert space we have a geometric characterization of $\text{c-span}\, S$:

Theorem 6.5. *Let $S \subset H$ be any subset of a Hilbert space H. Then*

$$\text{c-span}\, S = \left(S^{\perp}\right)^{\perp}.$$

That is, $y \in \text{c-span}\, S$ if and only if y is perpendicular to everything that is perpendicular to S:

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

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

Definition 6.6. A collection of vectors S in an inner product space H is called *orthonormal* 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 *orthonormal basis* if $\text{c-span}\, S = H$.

Lemma 6.7. *Let S be an orthonormal set of vectors in a Hilbert space H. Then the* $\text{c-span}\, S$ *consists of all vectors of the form*

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

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

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

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

Remark 6.8. Most orthonormal sets encountered in practice are countable, so we would tend to write $S = \{x_1, \ldots, \}$ and

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

However, the lemma holds even for uncountable orthonormal sets.

PROOF. It is clear that all vectors of the form ($\star$) are in $\overline{\operatorname{span} S} = \text{c-span}\, S$. 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 c-span S is contained in this subspace. Thus the two subspaces are equal.

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

Theorem 6.9. *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 c-span $S_{\max} \subsetneq X$. Then, c-span $S_{\max}^{\perp}$ is a non-trivial closed subspace. Let $y \in$ c-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 6.10 (Bessel's inequality)**.** *Let S be any orthonormal set in a Hilbert space H (not necessarily a basis), then*

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

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

LECTURE 7

Gram-Schmidt Process, Isometries of Hilbert Spaces, and Duality of Banach Spaces

Reading: Ch. 6 and 8 in Lax.

Gram-Schmidt Process

Recall that a metric space is *separable* if it contains a countable dense set. If a Hilbert space is separable, then any orthonormal basis is finite or countable. For such spaces we can avoid Zorn's Lemma in the construction of an orthonormal basis by using the *Gram-Schmidt process*.

Theorem 7.1 (Gram-Schmidt Process)**.** *Let $(y_j)_{j=1}^{\infty}$ be a sequence of vectors in a Hilbert space. Then there is a sequence $(x_j)_{j=1}^{\infty}$, with $N \in \mathbb{N} \cup \{\infty\}$, such that*

$$\langle x_j,\ x_k \rangle \ = \ 0 \quad \text{if } j \neq k,$$

$\|x_j\| = 1$ or 0 for each j and

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

The set $S = \{x_j \mid \|x_j\| = 1\}$ is an ortho-normal basis for c-span $\{y_j \mid j = 1, \ldots, \infty\}$.

PROOF. The proof is recursive. If $y_1 = 0$ let $x_1 = 0$. Otherwise 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

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

If $y_n \in \operatorname{span}\{y_1, \ldots, y_{n-1}\}$, let $x_n = 0$. Otherwise let

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

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

$$\langle x_n,\ x_k \rangle \ = \ 0 \quad \text{for } 1 \leq k < n$$

and

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

By induction, the result follows. □

Corollary 7.2. *Let H be a separable Hilbert space. Then H has a finite or countable orthonormal basis.*

Isometries of Hilbert spaces

Finally, let us discuss the isometries of Hilbert spaces.

Theorem 7.3. *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 c-span $S' \subset H'$. *Furthermore, any linear isometry of H with a subspace of H' is of this form.*

Corollary 7.4. *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 and any finite dimensional Hilbert space is isomorphic to $\ell^2(\{1, \dots, n\}) \cong \mathbb{C}^n$ for some n.*

Remark 7.5. 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.

Dual of a Banach Space

The dual of a Banach space X is the linear space $X^\star$ of bounded linear functionals on X. The operator norm

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

makes $X^\star$ a normed space.

Proposition 7.6. *The dual space X^* is a Banach space.*

PROOF. By exercise 4.3, $\|\cdot\|$ is a norm on $X^\star$. It remains to show that $X^\star$ is complete. Suppose that ℓ_n is a Cauchy sequence in $X^\star$. Then for each $x \in X$, $\ell_n(x)$ is a Cauch sequence in F, the base field of X. Let $\ell(x) = \lim_n \ell_n(x)$. The result map is linear by elementary properties of limits in F and furthermore ℓ is bounded since

$$|\ell(x)| \;=\; \lim_{n \to \infty} |\ell_n(x)| \;\leq\; \lim_{n \to \infty} \|\ell_n\| \|x\|.$$

Here, $\lim_n \|\ell_n\|$ exists since, by the triangle inequality,

$$|\|\ell\|_n - \|\ell\|_m| \;\leq\; \|\ell_n - \ell_m\| \,. \quad \square$$

We have the following dual characterization of the norm on X:

Theorem 7.7. *For every $x \in X$ we have*

$$\|x\| \;=\; \max_{\ell \neq 0\ \ell \in X^\star} \frac{|\ell(x)|}{\|\ell\|}.$$

PROOF. Since $\|\ell\|\|y\| \geq |\ell(y)|$ the left hand side is no smaller than the right hand side. 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. □

Since $X^\star$ is a Banach space, we can take form its dual $(X^\star)^\star$. There is a natural isometry $T : X \to (X^\star)^\star$ given by

$$T_x(\ell) \;=\; \ell(x)\;.$$

It is natural to ask whether this isometry is surjective or not, i.e., whether $X \cong (X^\star)^\star$?

As an example, recall the definitions of the following spaces, with associated norms

$$c_0 \;:=\; \Big\{(a_n)_{n=1}^\infty \;\Big|\; \lim_{n\to\infty} a_n \;=\; 0\Big\}\;, \qquad \|(a_n)_{n=1}^\infty\|_0 \;=\; \max_n |a_n|$$

$$\ell_1 := \Big\{(a_n)_{n=1}^\infty \;\Big|\; \sum_n |a_n| < \infty\Big\}\;, \qquad \|(a_n)_{n=1}^\infty\|_1 \;=\; \sum_n |a_n|\;,$$

$$\ell_\infty := \Big\{(a_n)_{n=1}^\infty \;\Big|\; \sup_n |a_n| < \infty\Big\}\;, \qquad \|(a_n)_{n=1}^\infty\|_\infty \;=\; \sup_n |a_n|\;,$$

each of which is a Banach space. Note that

$$\ell_1 \subsetneq c_0 \subsetneq \ell_\infty$$

and that c_0 is a closed subspace of ℓ_∞. We have

Theorem 7.8. *$c_0^\star = \ell_1$, $\ell_1^\star = \ell_\infty$, and $\ell_\infty^\star \supsetneq \ell_1$.*

PROOF. Let ℓ be a linear functional on c_0. Evaluating ℓ on the sequences $\boldsymbol{e}^k$ with

$$e_n^k = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n \end{cases}$$

produces a sequence $b_n = \ell(\boldsymbol{e}^n)$. If $\boldsymbol{a}$ is eventually zero, i.e., $a_n = 0$ for $n > n_0$, then we have $\boldsymbol{a} = \sum_{n=1}^{n_0} a_n \boldsymbol{e}^n$. Thus, by linearity we have

$$\ell(\boldsymbol{a}) \;=\; \sum_{n=1}^{n_0} a_n b_n\;.$$

Taking $a_n = \mathrm{e}^{-i \arg b_n}$ for $n \leq n_0$ and 0 for $n > n_0$, we find that

$$\sum_{n=1}^{n_0} |b_n| \;=\; \ell(\boldsymbol{a}) \;\leq\; \|\ell\|\;.$$

Taking $n_0 \to \infty$ we see that $\boldsymbol{b} \in \ell_1$. For arbitrary $\boldsymbol{a} \in c_0$, we have $\boldsymbol{a} = \lim_{n\to\infty} \sum_{j=1}^n a_j \boldsymbol{e}^j$. Thus by linearity and continuity $\ell(\boldsymbol{a}) = \lim_{n\to\infty} \sum_{j=1}^n a_j b_j$, where the right hand side is absolutely summable since $\boldsymbol{b} \in \ell_1$.

The same idea works to prove $\ell_1^\star = \ell_\infty$. Finally, it is clear that $\ell_1 \subset \ell_\infty^\star$, however we can construct linear functionals on $\ell_\infty^\star$ that are not associated to any element of ℓ_1. For instance, consider the space

$$L \;=\; \{\boldsymbol{a} : \lim_{n\to\infty} a_n \text{ exists}\}\;.$$

This is a closed subspace of ℓ_∞ and the linear functional $\ell(\boldsymbol{a}) = \lim_n a_n$ satisfies

$$|\ell(\boldsymbol{a})| \leq \|\boldsymbol{a}\|_\infty$$

for all $\boldsymbol{a} \in L$. So by the Hahn-Banach theorem ther is an extension of ℓ to ℓ_∞ satisfying the same bound. This linear functional satisfies $\ell(\boldsymbol{e}^k) = 0$ for all k and thus cannot be represented as $\sum_n b_n a_n$ for some $\boldsymbol{b} \in \ell_1$. □

On the other hand there are spaces with $X \cong (X^\star)^\star$.

Definition 7.9. A Banach space is called *reflexive* if $(X^\star)^\star \cong X$. That is if every norm-bounded linear functional on $X^\star$ is of the form $\ell \mapsto \ell(x)$ for some $x \in X$.

Theorem 7.10. *Every Hilbert space is reflexive.*

PROOF. By the Riesz-Fréchet theorem 5.10, every linear functional $\ell \in X^*$ can be represented as $\ell(x) = \langle y_\ell, x \rangle$ for a unique $y_\ell \in X$. It follows that $\|\ell\| = \|y_\ell\|$, so the map $\ell \mapsto y_\ell$ is an isometry. If X is a *real* Hilbert space, the map $\ell \mapsto y_\ell$ is also linear and we see that $X \cong X^*$.

If X is a *complex* Hilbert space, the situation is slightly more complicated. The map $S\ell = y_\ell$ is not linear, but is rather conjugate linear. Thus we do not have a linear isometry between X and X^*. However, X^* is a complex Hilbert space, with

$$\langle \ell, \ell' \rangle_{X^*} = \langle S\ell', S\ell \rangle_X .$$

Note that we have reversed the order of $S\ell$ and $S\ell'$ in the inner product on the right to guarantee the correct skew-linearity of $\langle \cdot, \cdot \rangle_{X^*}$. By applying Riesz-Fréchet to X^* we obtain a conjugate linear isometry $S' : (X^*)^* \to X^*$. The compostion SS' is a linear isometry from $(X^*)^*$ to X. □

LECTURE 8

L^p spaces

Reading: Ch. 8 in Lax

Let (X, μ) be a measure space, where X and μ is a measure defined on a sigma algebra $\mathcal{S}_\mu$ of subsets of X. For $1 \leq p < \infty$, the L^p space with respect to μ is defined to be

$$L^p(\mu; F) \;=\; \left\{ f : X \to F \;:\; \int_X |f(x)|^p d\mu(x) < \infty \right\} , \tag{8.1}$$

where $F = \mathbf{C}$ or $\mathbf{R}$ and we identify functions that are equal μ-almost everywhere. In the discussion below we will use $L^p(\mu)$ to denote either $L^p(\mu; \mathbf{R})$ or $L^p(\mu; \mathbf{C})$ whenever the choice of specific base field is unimportant. The L^p norm is defined to be

$$\|f\|_p \;:=\; \left(\int_X |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} . \tag{8.2}$$

Similarly, we define the L^∞ norm

$$\|f\|_\infty \;:=\; \inf\{t \in \mathbb{R} \;:\; \mu(\{x \in X \;:\; |f(x)| > t\} = 0)\} , \tag{8.3}$$

and the space

$$L^\infty(\mu; F) \;=\; \{f : X \to F \;:\; \|f\|_\infty < \infty\} . \tag{8.4}$$

It is clear from the definitions that we have

Proposition 8.1. *For $1 \leq p \leq \infty$, we have*

(1) $\|f\|_p \geq 0$ and $\|f\|_p = 0$ if and only if $f = 0$ almost everywhere.
(2) $\|\alpha f\|_p = |\alpha| \|f\|_p$ for all $\alpha \in F$.

For $p = 1$ and $p = \infty$ the triangle inequalities,

$$\|f + g\|_1 \;\leq\; \|f\|_1 + \|g\|_1 \quad \text{and} \quad \|fg\|_\infty \;\leq\; \|f\|_\infty + \|g\|_\infty ,$$

also follow directly from the definitions. To prove the triangle inequalty for $1 < p < \infty$, we require

Theorem 8.2 (Hölder's Inequality)**.** *Let (X, μ) be a measure space, let $1 \leq p < \infty$, and let q be the conjugate exponent with $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\mu)$ and $g \in L^q(\mu)$ then $fg \in L^1(\mu)$ and*

$$\int_X |fg| \mathrm{d}\mu \;\leq \|f\|_p \|g\|_q .$$

PROOF. If $p = 1$ and $q = \infty$, this is immediate from the pointwise almost everywhere inequality $|fg| \leq \|g\|_\infty |f|$.

For $1 < p < \infty$, it suffices to consider $f, g \geq 0$. One has

$$fg \;=\; \exp(\log f + \log g) \;=\; \exp\!\left(\tfrac{1}{p} \log f^p + \tfrac{1}{q} \log g^q \right) \;\leq\; \tfrac{1}{p} f^p + \tfrac{1}{q} g^q ,$$

by convexity of the exponential. Thus

$$\int_X fg d\mu \;\leq\; \tfrac{1}{p}\int_X f^p d\mu + \tfrac{1}{q}\int_X g^q d\mu \;<\; \infty\,, \tag{8.5}$$

so $fg \in L^1$. Applying (8.5) with $\frac{1}{\|f\|_p} f$ and $\frac{1}{\|g\|_q} g$ in place of f and g yields

$$\frac{\int_X fg d\mu}{\|f\|_p \|g\|_q} \;\leq\; \frac{1}{p} + \frac{1}{q} \;=\; 1\,. \quad \square$$

Recall that a measure μ is *semi-finite* if whenever $\mu(E) = \infty$ there is $E' \subset E$ such that $0 < \mu(E') < \infty$. That is a measure is semi-finite if it has no infinite atoms.

Corollary 8.3. *Let (X, μ) be a measure space, let $1 \leq p < \infty$, and let q be the conjugate exponent: $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\mu)$, then*

$$\|f\|_p \;=\; \sup\left\{\left|\int_X fg d\mu\right| \;:\; g \in L^q(\mu) \text{ and } \|g\|_q = 1\right\}\,. \tag{8.6}$$

If μ is semi-finite then (8.6) *holds also for $p = \infty$.*

Remark. Note that (8.6) definitely fails for $p = \infty$ if μ is not semi-finite. Indeed, if $f = 1_E$ with where $\mu(E) = \infty$ and E has no subset of finite measure then $\|f\|_\infty = 1$ but $\int fg = 0$ for any $g \in L^1$..

PROOF. Hölder's equality guarantees that $\|f\|_p$ is an upper bound for the right hand side of (8.6). To show equality, we will show that $\int_X fg d\mu = \|f\|_p$ for a suitable choice of g. If $\|f\|_p = 0$, then $f = 0$ almost everywhere, and any g will do.

If $\|f\|_p > 0$, let

$$g \;=\; \frac{1}{\|f\|_p^{p-1}} \begin{cases} \frac{|f|^{p-1}}{f}\,, & \text{if } |f| > 0 \\ 0 & \text{if } f = 0\,. \end{cases}$$

Then $fg = \frac{1}{\|f\|_p^{p-1}} |f|^p$ so $\int_X fg d\mu = \|f\|_p$. It remains to see that $\|g\|_q = 1$. If $p = 1$ then $g = \frac{|f|}{f} I[|f| > 0]$, so $|g| = 1$ almost everywhere on the support of f and $\|g\|_\infty = 1$. If $p > 1$ then $q < \infty$ and

$$|g|^q \;=\; \frac{|f|^{q(p-1)}}{\|f\|_p^{q(p-1)}} \;=\; \frac{|f|^p}{\|f\|^p}$$

so $\|g\|_q = 1$.

If $p = \infty$ and μ is semi-finite, then for any $t < \|f\|_\infty$ we can find a measurable set E_t with $0 < \mu(E_t) < \infty$ such that $|f| > t$ on E_t. Taking $g = \frac{1}{\mu(E_t)} 1_{E_t}$ we have

$$\int_X |fg| d\mu \;=\; \frac{1}{\mu(E_t)} \int_{E_t} |f| d\mu \;>\; t\,.$$

As this was possible for any $t < \|f\|_\infty$ we see that (8.6) holds. $\square$

Now we can show that $\|\cdot\|$ is a norm, and furthermore that L^p is a Banach space.

Theorem 8.4. *Let (X, μ) be a measure space and $1 \leq p \leq \infty$. Then $\|\cdot\|_p$ is a norm and $L^p(\mu)$ is a Banach space.*

Remark. The completeness of L^p is often refered to as the *Riesz-Fischer Theorem*. Riesz 1907a and Fischer 1907 studied the completeness of $L^2([a,b])$ for an interval on the real line.

SKETCH OF PROOF. The positivity and scaling properties of the norm were already stated in Prop. 8.1. It remains to show the triangle inequality. For $1 \le p < \infty$ this follows from Cor. 8.3 since

$$\|f+g\|_p = \sup_{\|h\|_q=1} \left|\int_X (f+g)hd\mu\right| \le \sup_{\|h\|_q=1} \left|\int_X fhd\mu\right| + \sup_{\|h\|_q=1} \left|\int_X ghd\mu\right| = \|f\|_p + \|g\|_p .$$

For $p = \infty$, one has $|f+g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty$ almost everywhere, so the inequality holds in that case too.

To show that L^p is a Banach space, we will use the following

Lemma 8.5. *Let $(f_n)_{n=1}^\infty$ be a sequence of measurable functions. Then*

$$\left\|\sum_{n=1}^\infty |f_n|\right\|_p \le \sum_{n=1}^\infty \|f_n\|_p , \tag{8.7}$$

and if the sum on the right hand side is finite, then the series $g = \sum_{n=1}^\infty f_n$ converges absolutely μ-almost everywhere, $\|g\|_p \le \sum_n \|f_n\|_p$, and $\lim_N \left\|g - \sum_{n=1}^N f_n\right\|_p = 0$.

PROOF OF LEMMA. Eq. (8.7) generalizes the triangle inequality to infinite sums. For $p = \infty$, this result and the following conclusions are immediate from the definition of the norm. For $1 \le p < \infty$, we may prove (8.7) by noting that

$$\int_X \left|\sum_{n=1}^N |f_n|\right|^p d\mu \le \left(\sum_{n=1}^N \|f_n\|_p\right)^p \le \left(\sum_{n=1}^\infty \|f_n\|_p\right)^p$$

by the usual triangle inequality. Eq. (8.7) follows by the monotone convergence theorem.

From (8.7), we conclude that the series defining g converges absolutely μ-almost everywhere provided $\sum_n \|f_n\|_p < \infty$. It follows easily that $\|g\|_p \le \sum_n \|f_n\|_p$. Finally, $g - \sum_{n=1}^N f_n = \sum_{n=N+1}^\infty$, so $\|g\|_p \le \sum_{n=N+1}^\infty \|f_n\|_p \to 0$. ☐

Returning to the proof that L^p is a Banach space, suppose that $(f_n)_{n=1}^\infty$ is a Cauchy sequence in L^p. It suffices to show that a subsequence has a limit. By passing to a suitable subsequence, we may assume that $\|f_n - f_m\| \le 2^{-n}$ for all $n \le m < \infty$. Let $g_n = f_{n+1} - f_n$ for each n. Then we have $\sum_n \|g_n\|_p \le \sum_n 2^{-n} = 1$. Thus $\sum_n g_n$ converges almost everywhere to a function in L^p by the Lemma. However,

$$f_{N+1} = f_1 + \sum_{n=1}^N g_n .$$

We conclude that $f_N \to f_1 + \sum_{n=1}^\infty g_n$ in L^p. Therefore L^p is complete, and thus a Banach space. ☐

Theorem 8.6. *Let (X,μ) be a measure space. For $1 < p < \infty$, we have $L^p(\mu)^\star \cong L^q(\mu)$ where $\frac{1}{p} + \frac{1}{q} = 1$. If μ is σ-finite, then $L^1(\mu)^\star \cong L^\infty(\mu)$.*

Remark. Recall that μ is σ-finite if $X = \cup_{j=1}^\infty E_j$ with $\mu(E_j) < \infty$.

PROOF. The Hölder inequality shows that $L_q(X) \hookrightarrow L^p(X)^\star$, via the map $g \mapsto \ell_g$ defined by

$$\ell_g(f) \;=\; \int_X f(x)g(x)\mathrm{d}\mu(x), \quad f \in L^p \text{ and } g \in: L^q.$$

Furthermore, by Cor. 8.3 we have $\|\ell_g\| = \|g\|_q$, so this mapping is an isometry.

Let us first prove the result under the assumption that μ is finite, using the *Radon-Nikodym* theorem:

Theorem 8.7 (Radon 1913 and Nikodym 1930; see Thm. 6.10 of Rudin 1987)**.** *Let (X, μ) be a σ-finite measure space. If ν is a complex measure on $\mathcal{S}_\mu$ such that $\nu(E) = 0$ whenever $\mu(E) = 0$, then there is $f \in L^1(\mu; \mathbb{C})$ such that $\nu(E) = \int_X f d\mu$ for any measurable set E.*

Suppose now that μ is finite and ℓ is a linear functional on $L^p(X)$. For each measurable set A, we have $1_A \in L^p(X)$, with $\|1_A\|_p = \mu(A)^{1/p}$. So

$$\nu(A) \;:=\; \ell(1_A)$$

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

$$\left\| 1_{\bigcup_j A_j} - \sum_{j=1}^{n} 1_{A_j} \right\|_p \;=\; \left(\mu\left(\textstyle\bigcup_{j=n+1}^{\infty} A_j\right)\right)^{\frac{1}{p}} \;\to\; 0\,.$$

Furthermore $\nu(A) = 0$ if $\mu(A) = 0$. Thus by the Radon-Nikodym Theorem, there is $g \in L^1(\mu)$ such that

$$\ell(1_A) \;=\; \nu(A) \;=\; \int_A g(x)\mathrm{d}\mu(x) \;=\; \int_X g(x)1_A(x)\mathrm{d}\mu(x).$$

By taking limits of simple functions we have

$$\ell(f) \;=\; \int_X g(x)f(x)\mathrm{d}\mu(x) \tag{8.8}$$

whenever $f \in L^\infty(\mu)$, using dominated convergence.

It remains to show that $g \in L^q$, for then (8.8) extends to all of L^p by the Hölder inequality and density of L^∞ in L^p. First consider the case $p > 1$ and 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 $|f_t| \leq t^{q-1}$ so $f_t \in L^\infty \subset L^p$. Thus

$$\int_{\{|g|<t\}} |g|^q \mathrm{d}\mu = \int f_t g \mathrm{d}\mu \;\leq\; \|\ell\| \left[\int |f_t|^p \mathrm{d}\mu\right]^{\frac{1}{p}}.$$

But $|f_t|^p = |g|^{pq-p} = |g|^q$ on $\{|g(x)| < t\}$. Thus

$$\left[\int_{\{|g|<t\}} |g|^q \mathrm{d}\mu(x)\right]^{\frac{1}{q}} \leq \|\ell\|$$

and it follows by monotone convergence that $\|g\|_q \leq \|\ell\|$.

When $p = 1$ we modify the above by setting

$$f_t(x) = \frac{|g(x)|}{G(x)} 1_{A_t} \quad \text{where} \quad A_t = \{x \mid |g(x)| > t\}.$$

Note that $f_t \in L^\infty$ and that

$$t\mu(A_t) < \int_{A_t} |g| \mathrm{d}\mu = \int f_t g \mathrm{d}\mu = \ell(f_t) \leq \|\ell\| \|f_t\|_{L^1} = \|\ell\| \mu(A_t).$$

It follows that $\mu(A_t) = 0$ when $t > \|\ell\|$, so $\|g\|_\infty \leq \|\ell\|$.

If μ is σ-finite, we can write X as a countable disjoint union $X = \bigcup_j E_j$ with $\mu(E_j) < \infty$. Let $\mu_j(A) = \mu(A \cap E_j)$. Note that $L^p(\mu_j)$ embeds in $L^p(\mu)$ in a natural way (extending a function on E_j to be zero on the rest of X). Thus ℓ defines a bounded linear functional on $L^p(\mu_j)$ and by the above argument there is a function $g_j \in L^q(\mu_j)$ such that $\ell(1_{E_j} f) = \int_{E_j} g_j f d\mu$ for all $f \in L^p(\mu)$. Furthermore, by finite additivity, we have

$$\ell(1_{\bigcup_{j=1}^n E_j} f) = \sum_{j=1}^n \int_{E_j} g_j f d\mu$$

for all $f \in L^p(\mu)$. It follows that $\left\| \sum_{j=1}^n 1_{E_j} g_j \right\|_q \leq \|\ell\|$. Let $g = \sum_{j=1}^n 1_{E_j} g_j$, noting that at most one term in the sum is nonzero. It follows by monotone convergence that $\|g\|_q \leq \|\ell\|$, and by dominated convergence that

$$\int_X g f d\mu = \sum_j \int_{E_j} g_j f d\mu = \sum_j \ell(1_{E_j} f) = \ell(f),$$

for every $f \in L^p(\mu)$.

For a proof that the inequality extends to general (non σ-finite) μ in case $1 < p < \infty$, see Theorem 6.15 in Folland 1999. In the proof, one shows that any linear functional ℓ on $L^p(\mu)$ vanishes except on functions supported on a σ-finite piece of X. □

Corollary 8.8. *$L^p(X)$ is reflexive for $1 < p < \infty$.*

This result also follows from

Theorem 8.9 (Milman 1938). *Any uniformly convex Banach space is reflexive*

In general L^1 is not reflexive: $(L^1)^\star = L^\infty$ but L^∞ contains linear functionals that are not in L^1. The proof breaks down even if $\mu(X) < \infty$. Given a linear functional ℓ on L^∞ we can define a set function

$$\nu(A) = \ell(1_A)$$

as above. It is certainly *finitely* additive, and clearly $\nu(A) = 0$ if $\mu(A) = 0$. However ν is *not* in general countably additive since

$$\left\| \sum_{j=1}^n 1_{A_j} - 1_{\bigcup_j A_j} \right\|_{L^\infty} = 1$$

as long as $\cup_{j=n+1}^\infty A_j$ has positive measure.

The inequality $(L^\infty)^\star \neq L^1$ also follows from:

Theorem 8.10. *Let X be a Banach space. If $X^\star$ is separable so is X.*

PROOF. Let $\{\ell_n\}$ be a countable dense subset of $X^\star$. 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 $\{x_n \mid n = 1, 2, \ldots\}$ is dense in X.

Suppose contrarily that c-span $\{x_n \mid n = 1, 2, \ldots\} \neq X$. Then there is a non-zero linear functional $\ell \in X^\star$ 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 c-span $\{x_n \mid n = 1, 2, \ldots\} = X$. □

LECTURE 9

Point set topology in a nutshell

We will make a brief interlude now to discuss point-set topology. The immediate reason is to be able to formulate the Riesz-Markov-Kakutani theorem on the dual of $C(Q)$ with Q a compact Hausdorff space. Later on we will introduce locally convex spaces; for that we will need a bit of topology. This will be a very brief presentation. For more details, see Chapter 2 of Simon 2015b.

1. Definitions

Definition 9.1. A *topological space* is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T}$ is a family of subsets of X satisfying

(1) $\emptyset, X \in \mathcal{T}$,
(2) $\mathcal{T}$ *is closed under arbitrary unions*: if $\mathcal{V} \subset \mathcal{T}$ then $\bigcup_{U \in \mathcal{V}} U \in \mathcal{T}$,
(3) $\mathcal{T}$ *is closed under finite intersections*: if $U_1, \dots, U_n$ are in $\mathcal{T}$ then $\bigcap_j U_j \in \mathcal{T}$.

A family $\mathcal{T}$ satisfying (1-3) is called a *topology*. The sets in $\mathcal{T}$ are called *open sets*; if $U \in \mathcal{T}$ and $x \in U$ we say that U *is an open neighborhood of* x. A set $F \subset X$ is *closed* if $F^c = X \setminus F$ is open.

Proposition 9.2. *Let $(X, \mathcal{T})$ be a topological space and $E \subset X$. Let $\mathcal{T}_E = \{U \cap E \ : \ U \in \mathcal{T}\}$. Then $\mathcal{T}_E$ is a topology on E.*

The topology $\mathcal{T}_E$ is called the *relative topology*; its elements are called *relatively open subsets of E*.

Definition 9.3. Let X be a topological space and $E \subset X$. The *closure of E*, denoted $\overline{E}$ is the smallest closed set containing E, that is

$$\overline{E} = \bigcap \{F \supset E \ : \ F \text{ is closed}\} .$$

The *interior of E*, denoted E°, is the largest open set contained in E, that is

$$E^\circ = \bigcup \{U \subset E \ : \ U \text{ is open}\} .$$

Topological spaces provide a general framework for discussing convergence and continuity:

(1) A sequence $(x_n)_{n=1}^\infty$ in X *converges* to $x \in X$ if for every open neighborhood U of x there is N such that $x_n \in U$ for $n \geq N$.
(2) If $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are topological spaces, we say that $f : X \to Y$ is *continuous* if $V \in \mathcal{T}_Y \implies f^{-1}(V) \in \mathcal{T}_X$.

A continuous map $f : X \to Y$ that is invertible and has a continuous inverse is called a *homeomorphism*. The set of continuous maps from X to Y is denoted $C(X, Y)$. The following proposition shows that topological spaces form a category, with continuous functions the morphisms. The proof is left as an exercise.

Proposition 9.4. *Let X, Y and Z be topological spaces.*
(1) *The identity map $I(x) = x$ is continuous on X.*
(2) *If $f \in C(X, Y)$ and $g \in C(Y, Z)$ then $g \circ f \in C(X, Z)$.*

The following two exercises relate convergence of sequences to continuity.

Exercise 9.1. Let $Y = \mathbb{N} \cup \{\infty\}$ and let $\mathcal{T}$ consist of 1) all subsets of $\mathbb{N}$, and 2) sets of the form $S \cup \{\infty\}$ where $\mathbb{N} \setminus S$ is a finite set. Show that $\mathcal{T}$ is a topology on Y and that a map $f : Y \to X$, with X another topological space, is continuous if an only if the sequence $(f(n))_{n=1}^{\infty}$ converges to $f(\infty)$.

Exercise 9.2. Let $f : X \to Y$ be a continuous function and $(x_n)_{n=1}^{\infty}$ a convergent sequence in X, with limit x. Show that $(f(x_n))_{n=1}^{\infty}$ converges to Y.

Exercise 9.3 (See Simon 2015b, Example 2.6.1)**.** Let X be an uncountable set and let $\mathcal{T} = \{S \subset X : X \setminus S \text{ is countable}\}$ and let $\mathcal{D} = \{S : S \subset X\}$.
(1) Prove that $\mathcal{T}$ is a topology and that a sequence $(x_n)_{n=1}^{\infty}$ converges in X if and only if it is eventually constant, i.e., if, for some N, one has $x_n = x$ for all $n \geq N$.
(2) Note that $\mathcal{D}$ is also a topology (it is called the discrete topology). Let $f(x) = x$, considered as a map from $(X, \mathcal{T})$ to $(X, \mathcal{D})$. Show that $\lim_n f(x_n) = f(\lim_n x_n)$ for any convergent sequence but f is not continuous (hint: $\{x_0\}$ is an open set in $(X, \mathcal{D})$).

The last example shows that sequencial convergence does not capture all the features of a general topological space. This can be remedied by introducing *nets*.

Definition 9.5. A *directed set* $(I, \leq)$ is a set I with a partial order $\leq$ such that any two elements of I have a common upper bound, i.e., $\alpha, \beta \in I \implies$ there is $\gamma \in I$ with $\alpha \leq \gamma$ and $\beta \leq \gamma$. Given a statement $S(\alpha)$ that depends on $\alpha \in I$, we say that $S(\alpha)$ is *eventually true* if there is α_0 such that $S(\alpha)$ is true for $\alpha \geq \alpha_0$.

Definition 9.6. A *net* in a topological space X is a map $I \to X$ where I is a directed set, denoted by $(x_\alpha)_{\alpha \in I}$. We say that a net $(x_\alpha)_{\alpha \in I}$ in X *converges to* $x \in X$ if for any open set $U \ni x$ it holds that x_α is eventually in U. We denote the limit of a convergent net by $x = \lim_\alpha x_\alpha$ or $x_\alpha \to x$, as for sequences.

Note that a sequence is a special case of a net, with $I = \mathbb{N}$. General nets may have uncountable and/or partially ordered index sets.

Exercise 9.4. Let X, Y be topological spaces. Show that
(1) A set $F \subset X$ is closed if and only if whenever a net $(x_\alpha)_{\alpha \in I}$ in F converges to $x \in X$, we have $x \in F$.
(2) A map $f : X \to Y$ is continuous if and only $\lim_\alpha f(x_\alpha) = f(\lim_\alpha x_\alpha)$ for any convergent net $(x_\alpha)_{\alpha \in I}$ in X.

2. Separation Axioms

The above axioms for topological spaces were introduced by Hausdorff 1914 in his text *Grundzüge der Mengenlehre* (*Principles of Set Theory*). Hausdorff included also the following axiom:

(T_2) *open sets separate points*: for every $x, y \in X$ with $x \neq y$ we can find open sets $U, V \in \mathcal{T}$ such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

In modern treatments, axioms (1-3) are taken as fundamental and a topological space that satisfies (4) is called *Hausdorff* (or T_2). The letter T in T_2 stands for *Trennungsaxiom* (separation axiom in German); there is a family of separation axioms T_j, $j = 0, \dots, 6$ as well as several interpolations $T_{2\frac{1}{2}}$ and $T_{3\frac{1}{2}}$. However the most important additional axiom is T_4, which is the following: *a topological space X is* normal *(or T_4) if it is Hausdorff and* open sets separate closed sets:

(T_4) X is Hausdorff and if $C, D \subset X$ are closed then there are open sets U, V such that $C \subset U$, $D \subset V$, and $U \cap V = \emptyset$.

Proposition 9.7. *Let X be a Hausdorff topological space. Then*
(1) one point sets in X are closed, and
(2) limits of nets are unique, i.e., if $x_\alpha \to y$ and $x_\alpha \to z$ then $y = z$.

Remark. The proof is left as an exercise. It can happen, in a general topological space, that one point sets are not closed and limits are not unique.

The following result is known as *Urysohn's lemma* — it is stated by Urysohn 1925 as a lemma for the proof of his metrization theorem. The result is, however, an important theorem in its own right.

Theorem 9.8 (Urysohn 1925)**.** *Let X be a normal topological space and let $C, D \subset X$ be closed and disjoint. Then there exists a continuous function $f : X \to [0, 1]$ such that*

$$f(x) = 1 \quad \text{on } C \quad \text{and} \quad f(x) = 0 \quad \text{on } D.$$

SKETCH OF PROOF. Note that we can reformulate the T_4 axiom as follows:

- let C be closed and V open with $C \subset V$. Then there are an open set U and a closed set K with $C \subset U \subset K \subset V$.

(To see this apply T_4 to C and V^c, letting K be the complement of the neighborhood of V^c that is obtained.)

With the above reformulation in mind, let $K_1 = C$ and $U_0 = D^c$. We can find $U_{1/2}$ and $K_{1/2}$ such that

$$K_1 \subset U_{1/2} \subset K_{1/2} \subset U_0 \ .$$

Now iterate the construction to find further sets betwee K_1 and $U_{1/2}$ and $K_{1/2}$ and U_0:

$$K_1 \subset U_{3/4} \subset K_{3/4} \subset U_{1/2} \subset K_{1/2} \subset U_{1/4} \subset K_{1/4} \subset U_0 \ .$$

Proceeding in this way we define, recursively, U_α and K_α for every dyadic rational $\alpha = m/2^n$ in $[0, 1]$ such that $U_\alpha \subset K_\alpha$ and $K_\beta \subset U_\alpha$ for $\alpha < \beta$.

Now we simply define

$$f(x) \ := \ \sup \{\alpha \ : x \in U_\alpha\} \ . \tag{9.1}$$

Similarly, one may define

$$g(x) \ := \ \inf \{\alpha \ : \ x \notin K_\alpha\} \ . \tag{9.2}$$

In fact, $g(x) = f(x)$. To see this note first that $f(x) \leq g(x)$ since $U_\alpha \subset K_\alpha$ for every α. On the other hand, given $\epsilon > 0$ we can find a dyadic rational $\alpha < f(x) + \epsilon$ such that $x \notin U_\alpha$. Taking another dyadic rational $\beta < \alpha + \epsilon$, we have $x \notin K_\beta$ and thus $g(x) \leq \beta \leq \alpha + \epsilon \leq f(x) + 2\epsilon$. As ϵ was arbitrary, we have $g(x) \leq f(x)$.

By (9.1), we have for any t that $\{x \ : \ f(x) > t\} = \bigcup_{\alpha > t} U_t$ is open , while by (9.2) $\{x \ : \ f(x) < t\} = \bigcup_{\alpha < t} K_\alpha^c$ is open. It follows that f is continuous. Since $f(x) = 0$ on $U_0^c = D$ and $f(x) = 1$ on $K_1 = C$, we are done. □

An important corollary of Urysohn's lemma is the following

Theorem 9.9 (Tietze Extension Theorem; Tietze 1915)**.** *Let X be a normal topological space. If $E \subset X$ is closed and $f : E \to \mathbb{R}$ is a bounded function that is continuous in the relative topology, then there is a bounded continuous function $g : X \to \mathbb{R}$ so that $g|_E = f$.*

Remark. For a direct exercise leading to the proof of the Tietze Extension Theorem, see Problem 4, §2.3 of Simon 2015b.

3. Metric Spaces

Metric spaces are an improtant example of topological spaces. Recall that a metric space (X, d) is a set X together with a *metric* d, which is map $X \times X \overset{d}{\mapsto} \mathbf{R}$ such that

(1) $d(x, y) \geq 0$ for all x, y and equals zero if and only if $x = y$,
(2) $d(x, y) = d(y, x)$ for all x, y, and
(3) $d(x, z) \leq d(x, y) + d(y, z)$ for all x, y, z.

The *open ball $B_r(x)$ of radius $r > 0$ and centered at x* is the set $B_r(x) = \{y \in x \mid d(y, x) < r\}$. The *metric topology $\mathcal{T}_d$* is the collection of sets U such that $x \in U \implies B_r(x) \subset U$ for some $r > 0$.

Exercise 9.5. Let (X, d) be a metric space. Verify that $\mathcal{T}_d$ is a topology and that X is Hausdorff.

Theorem 9.10. *Every metric space is normal.*

To prove this, we will use the following

Lemma 9.11. *Let $C \subset X$ be a closed subset of a metric space and let*

$$d_C(x) = \inf_{y \in C} d(x, y) \ .$$

Then d_C is a continuous function and $d_C(x) = 0$ if and only if $x \in C$.

PROOF. By the triangle inequality $|d_C(x) - d_C(x')| \leq d(x, x')$. It follows that d_C is continuous. Clearly $d_C(x) \geq 0$ and $d_C(x) \leq d(x, x) = 0$ if $x \in C$. On the other hand if $d_C(x) = 0$ then there are $x_n \in C$ such that $d(x, x_n) \leq \frac{1}{n}$, i.e., $x_n \to x$. Since C is closed, we have $x \in C$. □

PROOF OF THM. 9.10. Let C, D be disjoint closed sets. Let $f : X \to [-, 1]$ be defined as follows

$$f(x) \ := \ \frac{d_C(x) - d_D(x)}{d_C(x) + d_D(x)} \ ,$$

with d_C and d_D as in the lemma. Since $C \cap D = \emptyset$ we have $d_C(x) + d_D(x) > 0$ for all x, so the expression is defined and f is continuous. Also $f(x) = 1$ on D and $f(x) = 0$ on C. Taking $U = f^{-1}(-\infty, 0)$ and $V = f^{-1}(0, \infty)$ we obtain open sets that separate C and D. □

The topology on a metric space does not determine the metric — two metrics d_1 and d_2 are called equivalent if $\mathcal{T}_{d_1} = \mathcal{T}_{d_2}$.

Exercise 9.6. Show that d_1 and d_2 are equivalent if and only if there is $c > 1$ such that $\frac{1}{c} d_1(x, y) \leq d_2(x, y) \leq c d_1(x, y)$ for all $x, y \in X$.

A topological space $(X, \mathcal{T})$ is called *metrizable* if there is a metric d on X such that $\mathcal{T} = \mathcal{T}_d$. Clearly a necessary condition for metrizability is that X be normal. But not every normal space is metrizable. There are necessary and sufficient conditions for metrizability — see Bing 1951; Nagata 1950; Smirnov 1951 — however, they are technical to state. Urysohn gave a famous sufficient condition, based on the following notion. A *base* for a topology is a collection $\mathcal{B} \subset \mathcal{T}$ such that if $x \in U \in \mathcal{T}$ then there is $V \in \mathcal{B}$ such that $x \in V \subset V$. A space is called *second countable* if it has a countable base.

Theorem 9.12 (Urysohn 1925). *If $(X, \mathcal{T})$ is a second countable, normal space, then X is metrizable.*

For a proof, see Simon 2015b.

4. Compact Spaces

Definition 9.13. An *open cover* of a topological space X is a collection $\mathcal{C}$ of open subsets of X such that $X = \bigcup_{U \in \mathcal{C}} U$. A topological space X is *compact* if every open cover of X has a finite subcover.

Lemma 9.14. *Let X be a compact space. 1) If $F \subset X$ is closed, then F is compact (with respect to the relative topology). 2) If X is Hausdorff and $F \subset X$ is compact, then F is closed.*

PROOF. For 1, let $\{S_\alpha : \alpha \in I\}$ be an open cover of F by relatively open sets. For each $S_\alpha = U_\alpha \cap F$ with U_α open in X. Since F^c is open in X, we have $X = F^c \cup \bigcup_\alpha U_\alpha$. Thus there is a finite subcover. This in turn gives a finite subcover of F.

For 2, we will show that F^c is open. Let $x \in F^c$. Then for each $y \in F$ there is $U_y \ni y$ and $V_y \ni x$ open with $U_y \cap V_y = \emptyset$. Since F is compact, we have $F \subset U_{y_1} \cup \ldots \cup U_{y_n}$ for some finite number of points. Let $S_x = V_{y_1} \cap \ldots \cap V_{y_n}$. Then $x \in S_x$ is open and $S_x \subset F^c$. Clearly $F_c = \bigcup_{x \in F^c} S_x$, so F^c is open. □

Theorem 9.15. *If X is a compact Hausdorff space, then X is normal.*

PROOF. Let us first show that a closed set C can be separated from a one point set $\{x\}$ with $x \notin C$ (this is a property called *regularity* or T_3). For each $y \in C$ we can find open sets $U_y \ni x_0$ and $V_y \ni y$ such that $U_y \cap V_y = \emptyset$. Clearly $\{V_y : y \in C\}$ is an open cover of C. Since C is compact, we may find a finite subcover V_{y_j}, $j = 1, \ldots, n$. Let $V = \bigcup_{j=1}^n V_{y_j}$ and $U = \bigcap_{j=1}^n U_{y_j}$. Then U, V are open, $U \cap V = \emptyset$, $x \in U$ and $C \subset V$.

Now repeat the argument with two closed sets C, D. For each $x \in D$ find open $U_x \ni x$ and $V_x \supset C$ such that $U_x \cap V_x = \emptyset$. Since D is compact, there are $x_1, \ldots, x_n$ such that $U = U_{x_1} \cup \ldots \cup U_{x_n} \supset D$. Let $V = V_{x_1} \cap \ldots \cap V_{x_n}$. Then $U \supset D$, $V \supset C$ are open and $U \cap V = \emptyset$. □

Definition 9.16. Let X be a Hausdorff space. If $x \in X$, a *compact neighborhood of x* is a compact set K such that $x \in K^\circ$. A Hausdorff space X is *locally compact* if every $x \in X$ has a compact neighborhood.

If X is locally compact, the *one point compactification of* X is $X_\infty = X \cup \{\infty\}$, where ∞, the point at infinity, is a point not in X and the topology $\mathcal{T}_\infty = \mathcal{T}_X \cup \{X_\infty \setminus C : C \text{ is a compact subset of } X\}$.

Exercise 9.7. Check that $\mathcal{T}_\infty$ is a topology.

Theorem 9.17. *Let X be a locally compact Hausdorff space. Then the one-point compactification X_∞ is a compact Hausdorff space.*

PROOF. Let $\{U_\alpha : \alpha \in I\}$ be an open cover of X_∞. We must have $\infty \in U_{\alpha_0}$ for some α_0. The $U_{\alpha_0} = X_\infty\ C$ with C compact. The remaining sets must cover C. So there is a finite subcover $C \subset U_{\alpha_1} \cup \ldots U_{\alpha_n}$. Clearly then $X_\infty \subset U_{\alpha_0} \cup U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$. Thus X_∞ is compact.

If $x, y \in X_\infty$ and both are in X then we can separate them with open sets because X is Hausdorff. Suppose that one of them is ∞, say $y = \infty$. Let K be a compact neighborhood of $x \in X$. Then $x \in K^\circ$ and $y \in X_\infty \setminus K$. Thus X_∞ is Hausdorff. □

It is an amusing fact that we only used the local compactness of X to prove that X_∞ is Hausdorff. The one-point compactification exists for any space and is always compact. For it to be Hausdorff we need to start with a locally compact Hausdorff space.

LECTURE 10

The Riesz-Markov-Kakutani theorem

Reading: Appendix A of Lax 2002. See also Ch. 4 of Simon 2015b, §4.4 of Reed and Simon 1980, Ch. 2 of Rudin 1987, and Ch. 7 of Folland 1999.

1. Banach spaces of continuous functions

Given a Hausdorff topological space X, let $C(X)$ denote the set of all continuous, complex-valued functions on X. If X is locally compact, we say that $f \in C(X)$ *vanishes at infinity* if every $\epsilon > 0$ there is a compact $K \subset X$ such that $|f(x)| < \epsilon$ if $x \in K^c$, and let $C_0(X) \subset C(X)$ denote those functions that vanish at infinity. Let $C_{\mathbb{R}}(X)$, respectively $C_{0;\mathbb{R}}(X)$, denote the real valued elements of $C(X)$, respectively $C_0(X)$. If X is compact, then $C(X) = C_0(X)$.

Proposition 10.1. *If $f \in C_0(X)$, with X a locally compact Hausdorff space, then there is a point $x_0 \in X$ such that $|f(x_0)| = \sup_{x \in X} |f(x)| < \infty$.*

PROOF. Since f vanishes at infinity, $K = \{|f(x)| \geq 1\}$ is compact. Since f is continuous, $\mathcal{C} = \left\{f^{-1}(\{|z| < t\}) \, : \, t > 1\right\}$ is an open cover for K. As K is compact, there is a finite sub-cover. Because the sets in $\mathcal{C}$ are increasing with t, we see that there is $t_0 > 1$ such that $K \subset \{|f(x)| < t_0\}$. Thus $\kappa = \sup_x |f(x)| < t_0$.

If $\kappa = 0$, then $f(x) = 0 = \kappa$ for all x. If $\kappa > 0$, let $F = \{x \, : \, |f(x)| \geq \kappa/2\}$. Then F is compact. Let $(x_n)_n$ be a sequence in F such that $|f(x_n)| \to \kappa$. Every sequence in a compact space has a convergent sub-sequence (see Ex. 10.1). Thus there is a subsequence $x_{n_k} \to x_0$. Then we have $|f(x_0)| = \kappa$ by continuity. □

Exercise 10.1. Let $(x_n)_{n=1}^{\infty}$ be a sequence in a compact space X. Show that $(x_n)_{n=1}^{\infty}$ has a convergent subsequence. (Hint: Suppose there is a sequence with no convergent subsequence and use this to create an open cover with no finite sub-cover.)

Definition 10.2. Let X be a locally compact Hausdorff space. The *uniform norm* on $C_0(X)$ and $C_{0;\mathbb{R}}(X)$ is

$$\|f\|_u \; = \; \max_{x \in X} |f(x)| \, . \tag{10.1}$$

Lemma 10.3. *Let X be a locally compact Hausdorff space. Then the uniform norm defined in* (10.1) *is a norm. Furthermore $C_0(X)$ and $C_{0;\mathbb{R}}(X)$ are Banach spaces over $\mathbb{C}$ and $\mathbb{R}$, respectively.*

Exercise 10.2. Prove Lemma 10.3.

2. Measure theory on compact Hausdorff spaces

Definition 10.4. Let X be topological space. A *G_δ set* in X is a countable intersection of open sets. An *F_σ set* in X is a countable union of closed sets.

Proposition 10.5. *1) In a metric space X, every closed set is G_δ. 2) If X is second countable and normal, then every closed set is G_δ.*

PROOF. (1) $\implies$ (2) by the Urysohn metrization theorem 9.12. To prove (1), note that if $F \subset X$ is closed then $F = \bigcap_{n=1}^{\infty} \{x : d_F(x) < 2^{-n}\}$ where $d_F(x) = \inf_{y \in F} d(x,y)$. □

There are important compact Hausdorff spaces (e.g., uncountable products of compact sets) in which not every closed set is G_δ. However, the level sets of continuous real valued functions are always closed G_δ's sets.

Proposition 10.6. *Let X be a topological space and let $f : X \to \mathbb{R}$ be continuous. Then $f^{-1}([a,\infty))$ is a closed G_δ for each $a \in \mathbb{R}$.*

PROOF. The set $[a,\infty)$ is a closed G_δ in $\mathbb{R}$ (since $\mathbb{R}$ is a metric space). Thus $f^{-1}([a,\infty))$ is a closed G_δ, since f^{-1} preserves intersections, closed sets and open sets. □

Definition 10.7. Let X be a Hausdorff space. The *Baire σ-algebra*, $\mathsf{Baire}(X)$, is the smallest σ-algebra containing all compact G_δ's. The *Borel σ-algebra*, $\mathsf{Borel}(X)$, is the smallest σ-algebra containing all open sets.

Clearly $\mathsf{Baire}(X) \subset \mathsf{Borel}(X)$. Although these σ-algebras can be distinct in a general space X, they agree on $\mathbb{R}$ and $\mathbb{C}$, and more generally on any space which is a countable union of compact G_δ sets. We will always consider $\mathbb{C}$ and $\mathbb{R}$ with the Baire/Borel σ-algebra.

Proposition 10.8. *Let X be a compact Hausdorff space. 1) If $f \in C(X)$, then f is measurable with respect to* $\mathsf{Baire}(X)$*. 2)* $\mathsf{Baire}(X)$ *is the smallest σ-algebra such that every $f \in C(X)$ is measurable.*

For the proof see Simon 2015b.

Definition 10.9. Let X be a locally compact Hausdorff space. A *positive Baire measure* on X is a countably additive map $\mu : \mathsf{Baire}(X) \to [0,\infty]$ such that $\mu(K) < \infty$ for any compact K. A *complex Baire measure* is an countably additive map $\mu : \mathsf{Baire}(X) \to \mathbb{C}$. A *signed Baire measure* is a complex Baire measure with $\mu(E) \in \mathbb{R}$ for all E.

Remark. A signed or complex measure μ has $\sum_j |\mu(E_j)| < \infty$ whenever $(E_j)_{j=1}^{\infty}$ are disjoint. This follows from the fact that a sum is absolutely convergent if and only if it is convergent and invariant under all rearrangements — see Rudin 1976, Thms. 3.54 and 3.55. The *total variation measure* $|\mu|$, defined by

$$|\mu|(E) \;=\; \sup\left\{ \sum_{j=1}^{\infty} |\mu(E_j)| \;:\; (E_j)_{j=1}^{\infty} \text{ are disjoint and } E = \bigcup_j E_j \right\},$$

is a positive, finite measure on X — see Rudin 1987, Thm. 6.2. By the Radon-Nikodym Theorem, one has $d\mu = f d|\mu|$ where $f : X \to \mathbb{C}$ is a measurable function of modulus one — Rudin 1987, Thm. 6.12.

Theorem 10.10. *Let μ be a positive, finite Baire measure on a compact Hausdorff space X. Then μ is* regular*, meaning that for any $E \in$* $\mathsf{Baire}(X)$ *we have:*

(1) Inner regularity*: $\mu(E) = \sup\{\mu(K) : K \subset E$ is Baire and compact.$\}$, and*
(2) Outer regularity*: $\mu(E) = \inf\{\mu(U) : U \supset E$ is Baire and open.$\}$.*

SKETCH OF PROOF. One can show that the collection of sets for which inner and outer regularity holds is a σ-algebra — see Simon 2015b, Lemma 4.5.5. Let K be a compact

G_δ. Then inner regularity holds trivially. To prove outer regularity, note that we have $K = \bigcap_n V_n$ with V_n open. It could happen that V_n is not a Baire set for some n. However, by Urysohn's Lemma, Thm. 9.8, there are continuous functions $f_n : X \to [0,1]$ such that $f_n = 1$ on K and $f_n = 0$ on V_n^c. The sets $U_n = \{f_n > 0\}$ are open and Baire, by Prop. 10.6, since $U_n = \bigcup_{m=1}^{\infty}\{f_n \geq 2^{-m}\}$. Since $K \subset U_n \subset V_n$, we have $K = \bigcap_v U_n$ with U_n open and Baire. Thus $\mu(K) = \lim_n \mu(U_n)$ by dominated convergence and inner regularity holds. □

Remark 10.11. A *Borel measure* on X is a measure defined on the Borel σ-algebra. A Borel measure μ is regular if it satsfies

(1) *Inner regularity*: $\mu(E) = \sup\{\mu(K) \ : \ K \subset E \text{ compact.}\}$, and
(2) *Outer regularity*: $\mu(E) = \inf\{\mu(U) \ : \ U \supset E \text{ open.}\}$.

Regular Borel measures are sometimes called *Radon measures*. Unfortunately, the analogue of Thm. 10.10 does not hold for finite Borel measures — for an example of a finite, non-regular Borel measure on a compact Hausdorff space, see Folland 1999, §7.2 Problem 15. However, every Baire measure on a compact Hausdorff space can be extended uniquely to a regular Borel measure — see Dudley 2002, Thm. 7.31.

Definition 10.12. Let X be a locally compact Hausdorff space. Let $\mathcal{M}_+(X)$ denote the set of all positive Baire measures. Let $\mathcal{M}(X)$, respectively $\mathcal{M}_\mathbb{R}(X)$, denote the set of all complex, respectively signed, Baire measures. These last two spaces are normed linear spaces, over $\mathbb{C}$ and $\mathbb{R}$ respectively, with the total variation norm $\|\mu\| \ = \ |\mu|(X)$.

Proposition 10.13. *$\mathcal{M}(X)$ and $\mathcal{M}_\mathbb{R}(X)$ are Banach spaces in the total variation norm.*

The proof is left as an exercise.

3. The Riesz-Markov-Kakutani Theorem

The following theorem is called the Riesz-Markov-Kakutani Theorem. Riesz 1911 showed that the dual of $C([a,b])$ could be described using Stieltjes integrals, while Markov 1938 and Kakutani 1941 independently derived the general result stated here.

Theorem 10.14 (Riesz-Markov-Kakutani Theorem)**.** *Let Q be a compact Hausdorff space. Then to every bounded linear functional $\ell \in C_\mathbb{R}(Q)^\star$ (respectively, $C(Q)^\star$) is associated a unique signed (respectively, complex) Baire measure μ such that*

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

Furthermore the norm of ℓ is the total variation $\|\ell\| \ = \ |\mu|(Q)$. Thus $C_\mathbb{R}(Q)^* \ \cong \ \mathcal{M}_\mathbb{R}(Q)$ and $C(Q)^\star \cong \mathcal{M}(Q)$.

We will prove the real valued case; the complex case follows easily. Let

$$C_+(Q) = \{f \in C_\mathbb{R}(Q) \ : \ f(x) \geq 0 \text{ for all } x \in Q\} \ .$$

A linear functional $\ell \in C_\mathbb{R}(Q)$ is called *positive* if $\ell(f) \geq 0$ whenever $f \in C_+(Q)$. Note that a positive linear functional satisfies $\|\ell\| = \ell(1)$, since clearly $\ell(1) \leq \|\ell\|$ but also

$$0 \leq \ell(\|f\|_u 1 \pm f) \ = \ \|f\|_u \ell(1) \pm \ell(f) \ ,$$

from which we conclude that $|\ell(f)| \leq \|f\|_u \ell(1)$ and thus $\|\ell\| \leq \ell(1)$.

Theorem 10.15. *Given $\ell \in C_{\mathbb{R}}(Q)^\star$ there is a unique deomposition $\ell = \ell_+ - \ell_-$ with $\ell_\pm$ positive linear functionals and $\|\ell\| = \ell_+(1) + \ell_-(1)$.*

PROOF. For $f \in C_+(Q)$ define $\ell_+(f) = \sup\{\ell(h) : 0 \le h \le f\}$. It is clear that $\ell_+(tf) = t\ell_+(f)$ for $t \ge 0$ and that $\ell_+(f) \ge \ell(0) = 0$. Given $f, g \in C_+(Q)$, $f + g \in C_+(Q)$ and

$$\ell_+(f+g) \ge \sup\{\ell(h_1) + \ell(h_2) \ : \ 0 \le h_1 \le f \text{ and } 0 \le h_2 \le g\} \ = \ \ell_+(f) + \ell_+(g).$$

The opposite inequality follows from the following:

Claim 10.16. *Given $f, g \in C_+(Q)$ and $0 \le h \le f + g$ we can write $h = h_1 + h_2$ with $h_{1,2} \in C_+(Q)$, $0 \le h_1 \le f$ and $0 \le h_2 \le g$.*

PROOF OF CLAIM. Let $h_1 = \min\{f, h\}$. Then h_1 is continuous and $0 \le h_1 \le f$. Let $h_2 = h - h_1$. Since $h_1 \le h$, we have $h_2 \ge 0$. When $h_1 = f$ we have $h_2 \le f + g - f = g$, while when $h_1 = h$ we have $h_2 = 0$ so $h_2 \le g$. □

So, $\ell_+(tf + sg) = t\ell_+(f) + s\ell_+(g)$ if $t, s \ge 0$ and $f, g \in C_+(Q)$. Extend ℓ_+ to $C(Q)$ 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

$$\ell_-(f) = \sup\{\ell(h) - \ell(f) \ : \ 0 \le h \le f\} = \sup\{\ell(k) \ : \ -f \le k \le 0\} \ge 0,$$

for $f \in C_+(Q)$.

To prove $\|\ell\| = \ell_+(1) + \ell_-(1)$ note first that $\|\ell\| \le \|\ell_+\| + \|\ell_-\| = \ell_+(1) + \ell_-(1)$. On the other hand, by the definition of ℓ_+ and the corresponding equality for ℓ_- we have

$$\begin{aligned}\ell_+(1) + \ell_-(1) &= \sup\{\ell(h) \ : \ 0 \le h \le 1\} + \sup\{\ell(k) \ : \ -1 \le k \le 0\} \\ &= \sup\{\ell(g) \ : \ -1 \le g \le 1\} \ \le \ \|\ell\|. \end{aligned}$$ □

We are now ready to prove the Riesz-Kakutani Theorem in the real valued case. By the splitting of a linear functional into positive and negative parts, it suffices to show

Theorem 10.17 (Riesz, Markov, Kakutani)**.** *Let ℓ be a positive linear functional on $C_{\mathbb{R}}(Q)$ with Q a compact Hausdorff space. Then there is a unique positive Baire measure m on Q such that $\ell(f) = \int_Q f \mathrm{d}m$ and $\|\ell\| = m(Q)$. Conversely, any positive Baire measure m gives a positive linear functional on $C(Q)$ via $f \mapsto \int_Q f \mathrm{d}m$.*

Before the proof let us introduce some notation and a lemma. Given an open set $U \subset Q$, we write $f \prec U$ to indicate that $f \in C(Q)$, $\operatorname{supp} f \subset U$, and $0 \le f \le 1$. Similarly, if K is compact, $f \succ K$ indicates $1 - f \prec K^c$, i.e., $1_K \le f$ and $0 \le f \le 1$.

Lemma 10.18 (Finite partitions of unity)**.** *Let $K \subset Q$ be compact and let $U_1, \ldots, U_n$ be a finite open cover of K. Then there are $f_j \prec U_j$, $j = 1, \ldots, n$, such that $\sum_j f_j \succ K$.*

Remark 10.19. $(f_j)_{j=1}^n$ is called a *partition of unity on K subordinate to $(U_j)_{j=1}^n$.*.

PROOF. The case $n = 1$ is Urysohn's lemma, Thm. 9.8. Suppose the lemma is verified for covers of size $n - 1$ and let $K \subset U_1 \cup \cdots \cup U_n$. Then $K \setminus U_n$ is compact and covered by $(U_j)_{j=1}^{n-1}$. Let $(g_j)_{j=1}^{n-1}$ be a partition of unity on $K \setminus U_n$ subordinate to $(U_j)_{j=1}^{n-1}$ and let

$g_n \prec U_n$ be such that $g_n \equiv 1$ on $K \setminus \bigcup_{j=1}^{n-1} U_j$. We have $q = \sum_{j=1}^n g_j \geq 1$ on K. Let $V = \{q(x) > \frac{1}{2}\}$. Then V is open and $K \subset V$. Let $K \prec h \prec V$ and take

$$f_j(x) = \frac{h(x)}{q(x)} g_j(x) . \quad \square$$

PROOF OF RIESZ-MARKOV-KAKUTANI THEOREM. First, it is clear that any finite positive Baire measure m gives rise to a positive linear functional by $\ell(f) = \int_Q f \mathrm{d}m$. Since the functional is positive $\|\ell\| = \ell(1) = m(Q)$.

To see uniqueness, suppose that $m \neq m'$ are distinct Baire measures. By regularity, Thm. 10.10, we must have $m(U) \neq m'(U)$ for some open Baire U. By inner regularity, we can find sequences $(K_n)_{n=1}^\infty$ and $(K_n')_{n=1}^\infty$ of compact Baire subsets of U such that $m(U) = \lim_n m(K_n)$ and $m'(U) = \lim_n m'(K_n')$. Let $F_n = \bigcup_{j=1}^n (K_n \cup K_n')$ and let $F_n \prec f_n \prec U$. Then

$$m(K_n) \leq m(F_n) \leq \int_Q f_n dm \leq m(U) \quad \text{and} \quad m'(K_n') \leq m'(F_n) \leq \int_Q f_n dm' \leq m'(U) .$$

Thus $\lim_n \int_Q f_n dm = m(U) \neq m'(U) = \int_Q f_n dm'$. It follows that $\int_Q f_n dm \neq \int_Q f_n dm'$ for some n, so m and m' generate distinct linear functionals.

It remains to show any positive linear functional $\ell \in C_{\mathbb{R}}(Q)^\star$ is of the form $\ell(f) = \int_Q f dm$ for a positive Baire measure. The proof we give will actually construct the completion of the measure, its restriction to the Baire sets will be the Baire measure we seek.

Given an open set U we take

$$m^*(U) := \sup\{\ell(f) : f \prec U\}, \tag{10.2}$$

Since $1 \prec Q$, it follows that $m^*(Q) = \ell(1)$. Since $m^*(U) \leq m^*(V)$ if $U \subset V$, we have $m^*(U) \leq \ell(Q)$ for all U. We extend m^* to arbitrary subsets of Q by taking

$$m^*(S) := \inf\{m^*(U) : S \subset U \text{ and } U \text{ is open}\}. \tag{10.3}$$

By the monotonicity already noted for open sets, this identity is consistent for open U.

Claim 10.20. *m^* is an outer measure.*

PROOF OF CLAIM. Recall that an *outer measure* satisfies (1) $m * (\emptyset) = 0$ and (2) monotonicity with countable subadditivity: if $B \subset \bigcup_{j=1}^\infty A_j$, then $m^*(B) \leq \sum_{j=1}^\infty m^*(A_j)$. Because $0 \prec \emptyset$, we have $m^*(\emptyset) = \ell(0) = 0$.

Suppose that $B \subset \bigcup_j A_j$ and let $U_j \supset A_j$ be open with $m^*(U_j) \leq m^*(A_j) + \frac{\epsilon}{2^j}$. Then $B \subset V = \bigcup_{j=1}^\infty U_j$. Let $f \prec V$ with $\ell(f) \geq m^*(V) - \epsilon$. Since $\operatorname{supp} f$ is compact, we have $\operatorname{supp} f \subset \bigcup_{j=1}^n U_j$ for some n. Let $(g_j)_{j=1}^n$ be a partition of unity on $\operatorname{supp} f$ subordinate to $(U_j)_{j=1}^n$. Since $g_j f \prec U_j$, we have

$$m^*(B) \leq m^*(V) \leq \ell(f) + \epsilon \leq \epsilon + \sum_{j=1}^n \ell(f g_j) \leq \epsilon + \sum_{j=1}^n m^*(U_j) \leq 2\epsilon + \sum_{j=1}^\infty m^*(A_j) .$$

As ϵ was arbitrary, we see that m^* is monotonic and countably subadditive. $\square$

Let $\mathcal{F}_m$ denote the set of *Carathéodory measurable sets*: $E \in \mathcal{F}_m$ if and only if $m^*(A) = m^*(A \cap E) + m^*(A \setminus E)$ for every $A \subset X$. By the Carathéodory extension theorem, $\mathcal{F}_m$ is a σ-algebra and $m^*|_{\mathcal{F}_m}$ is a measure — see Tao 2011, Theorem 1.7.3.

Claim 10.21. *Every open set is Caratheodry measurable.*

PROOF. Let U be open. Since m^* is subadditive, it suffices to show

$$m^*(A) \geq m^*(A \cap U) + m^*(A \setminus U) \quad \text{for all } A \subset Q. \tag{10.4}$$

We first prove (10.4) for A open. Then $A \cap U$ is open and there is $f \prec A \cap U$ with $\ell(f) \geq m(A \cap U) - \epsilon$. Also $A \setminus \operatorname{supp} f$ is open, so we can find $g \prec A \setminus \operatorname{supp} f$ with $\ell(g) \geq m^*(A \setminus \operatorname{supp} f) - \epsilon$. Then $f + g \prec A$, so

$$\begin{aligned} m^*(A) \ \geq\ \ell(f+g) \ =\ \ell(f) + \ell(g) \ \geq\ & m^*(A \cap U) + m^*(A \setminus \operatorname{supp} f) - 2\epsilon \\ \geq\ & m^*(A \cap U) + m^*(A \setminus U) - 2\epsilon\,. \end{aligned}$$

Taking $\epsilon \to 0$, we obtain (10.4) for A open.

Now suppose A is an arbitrary set. Then there is $V \supset A$ with $m^*(V) \leq m^*(A) + \epsilon$. By (10.4) for V we have

$$m^*(A) + \epsilon \ \geq\ m^*(V) \ \geq\ m^*(V \cap U) + m^*(V \setminus U) \ \geq\ m^*(A \cap U) + m^*(A \setminus U)\,.$$

Taking $\epsilon \to 0$, we obtain (10.4). ☐

Since open sets generate $\mathsf{Borel}(Q)$, we have $\mathcal{F}_m \supset \mathsf{Borel}(Q) \supset \mathsf{Baire}(Q)$. Let m denote the restriction of m^* to $\mathsf{Baire}(Q)$. To complete the proof we will need the following

Claim 10.22. *If $K \subset Q$ is a compact Baire set, then $m(K) = \inf\{\ell(f) \,:\, f \succ K\}$.*

PROOF OF CLAIM. Let $f \succ K$. Then $V_\epsilon = \{f > 1 - \epsilon\} \supset K$ is open. Let $g \prec V_\epsilon$ with $\ell(g) \geq m(V_\epsilon) - \epsilon$. Since $g \leq \frac{1}{1-\epsilon} f$, we have $\ell(f) \geq (1-\epsilon)\ell(g) \geq (1-\epsilon)m(V_\epsilon) - \epsilon \geq m(K) - O(\epsilon)$. Taking $\epsilon \to 0$, we find that $\ell(f) \geq m(K)$. Now let $U_\epsilon \supset K$ with $m(U_\epsilon) \leq m(K) + \epsilon$ and let $K \prec f \prec U$. Then $m(K) \leq \ell(f_\epsilon) \leq m(U) \leq m(K) + \epsilon$. Taking $\epsilon \to 0$, the result follows. ☐

It remains to show that $\int f \mathrm{d}m = \ell(f)$. By linearity and scaling, it suffices to prove this for $0 \leq f \leq 1$. By Fubini's Theorem,

$$\int_Q f \mathrm{d}m \ =\ \int_0^1 m(\{x : f(x) \geq t\})\mathrm{d}t = \int_0^1 m(\{x : f(x) > t\})\mathrm{d}t.$$

Thus

$$\int_Q f \mathrm{d}m \ \leq\ \sum_{j=1}^n \frac{1}{n} m\left(\left\{x : f(x) \geq \tfrac{j-1}{n}\right\}\right).$$

Let $g_{j;n}$ be such that $\left\{f(x) \geq \frac{j-1}{n}\right\} \prec g_{j;n} \prec \left\{f(x) > \frac{j-2}{n}\right\}$. Then

$$\int_Q f \mathrm{d}m \leq \sum_{j=1}^n \frac{1}{n} \ell\left(g_{j;n}\right) = \ell\left(\sum_{j=1}^n \frac{1}{n} g_{j;n}\right).$$

Now for every $x \in Q$

$$\sum_{j=1}^n \frac{1}{n} g_{j;n}(x) = \sum_{j=1}^{\lfloor nf(x) \rfloor} \frac{1}{n} + O(1/n) = \frac{1}{n}\lfloor nf(x) \rfloor + O(1/n) \ =\ f(x) + O(1/n).$$

We conclude that

$$\left\| \sum_{j=1}^{n} \frac{1}{n} g_{j;n}(x) - f(x) \right\|_n = O(1/n),$$

so $\ell(\sum_{j=1}^{n} \frac{1}{n} g_{j;n}) \to \ell(f)$, and $\int f \mathrm{d}m \leq \ell(f)$.

To show the reverse inequality, note that

$$\int_Q f \mathrm{d}m \;\geq\; \sum_{j=1}^{n} \frac{1}{n} m \left\{ f(x) > \frac{j}{n} \right\}.$$

Thus,

$$\int_Q f \mathrm{d}m \;\geq\; \ell \left(\sum_{j=1}^{n} \frac{1}{n} h_{j;n} \right),$$

with $\{f(x) \geq \frac{j+1}{n}\} \prec h_{j;n} \prec \{f(x) > \frac{j}{n}\}$. Again $\sum_j \frac{1}{n} h_{j,n}(x) = f(x) + O(1/n)$ and $\sum_j h_{j;n} \to f$ uniformly, so $\int f \mathrm{d}m \geq \ell(f)$, completing the proof. □

Part 3

Locally Convex Spaces

LECTURE 11

Locally Convex Spaces and the Schwartz Space

1. Topological Vector Spaces and Locally Convex Spaces

Reading: Lax 2002, §13.-13.2 and §B.1 and Reed and Simon 1980, Ch. 5 or Simon 2015b, Ch. 6.

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. Let X and Y be topological spaces. The *product topology* on $X \times Y$ is the smallest topology containing sets of the form $U \times V$ with $U \subset X$ and $V \subset Y$ open.

Remark. The product topology is also the smallest topology such that the coordinate maps $(x, y) \mapsto x$ and $(x, y) \mapsto y$ are continuous.

Definition 11.2. 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.

Theorem 11.3. *Let X be a TVS and let $U \subset X$ be open. Then*

(1) For any $x \in X$, $U + x = \{y + x \ : y \in U\}$ is open.
(2) For any scalar $k \neq 0$, $kU = \{ky \ : \ y \in U\}$ 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 $(y, x) \mapsto y + x$. Likewise (2) follows from continuity of $y \mapsto ky$.

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

The class of TVSs is very large. However, most of the TVSs important to analysis have the following property:

Definition 11.4. 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 base at the origin consisting of open convex sets.

Definition 11.5. Let X be a linear space. A subset $U \subset X$ is 1) *absorbing* if every 0 is an interior point of U, and 2) *balanced* if $x \in X \implies \omega x \in X$ whenever $\omega \in F$ has $|\omega| = 1$.

Remarks. 1) The term *absorbing* stems from the fact that U is absorbing if and only if $X = \bigcup_{t>0} tU$. 2) For $F = \mathbb{R}$, *balanced* is equivalent to symmetric, $U = -U$.

Proposition 11.6. *Let X be an LCS. Then X has a neighborhood base consisting of balanced, absorbing, convex sets.*

PROOF. Let $\mathcal{N}$ be a neighborhood base at 0 consisting of open, convex sets. Each element of $U \in \mathcal{N}$ is absorbing by Thm. 11.3, part 3. If $F = \mathbb{R}$, we may simply replace $\mathcal{N}$ by $\mathcal{N}' = \{U \cap (-U) : U \in \mathcal{N}\}$ to obtain the desired base.

If $F = \mathbb{C}$ we need to work a bit more. We claim that if $U \in \mathcal{N}$ there is a convex, open $V \ni 0$ and $\epsilon > 0$ such that if $x \in V$ then $\epsilon e^{i\theta} x \in U$ for all θ. To see this note that if $T(x,z) = zx$, then $T^{-1}(U)$ is open. Since $(0,0) \in T^{-1}(U)$ and X is an LCS, there is an open, convex set $V \ni 0$ and $\epsilon > 0$ such that $V \times (2\epsilon\mathbb{D}) \subset T^{-1}(U)$, where $\mathbb{D}$ denotes the unit disk. Take

$$U' = \bigcup_{\theta \in [0,2\pi]} \frac{1}{\epsilon} e^{-i\theta} V .$$

Then U' is open, balanced, and convex, and $U' \subset U$ by construction. Taking $\mathcal{N}' = \{U' : U \in \mathcal{N}\}$ gives the desired neighborhood base. □

Given a TVS X, we define the *dual* X' to be the set of all continuous linear functional on X. The dual X' is a linear space, but for a general TVS may not be very large, or may even consist of only the zero functional (examples will come later). However, the dual X' of an LCS has enough functionals to separate points of X:

Theorem 11.7. *Let X is a LCS. If $y \neq y'$ are points of X, then 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 2.12. Also, 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

$$\ell(x) = \ell_r(x) - i\ell_r(ix).$$

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 take U to be convex and balanced. Since all points of U are interior, Theorem 2.12 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

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

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 $\ell(x_0) = t$. As ℓ is linear,

$$\ell^{-1}(a-t, b-t) = \ell^{-1}(a,b) - x_0 \ni 0 .$$

Thus it suffices to suppose $a < 0 < b$ and show that $\ell^{-1}(a,b)$ contains an open neighborhood at 0. Let $s = \min\{-a, b\}$. Then, given $x \in U$,

$$\ell(sx) \leq p_U(sx) = sp_U(x) < s \quad \text{and} \quad \ell(-sx) \leq p_U(-sx) = sp_U(-sx) < s.$$

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

Converse to this construction is the following result

Theorem 11.8. *Let X be linear space and let L be any collection of linear functionals on X that separates points — i.e., for any $y, y' \in X$ there is $\ell \in L$ such that $\ell(x) \neq \ell(x')$. Endow X with the weakest TVS 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 11.9. 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 base of convex sets. But not every LCS has a norm which is compatible with the topology. By far the most important examples, are spaces of "test functions" and their duals, spaces of distributions.

2. Generation of an LCS by semi-norms

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 "generalized functions," or "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. It is common to use inner product and function notation for a distribution, writing

$$\ell(\psi) \;=\; \langle T,\, \psi\rangle \;=\; \int T(x)\psi(x)\mathrm{d}x,$$

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

A common, and useful, space of test functions is the *Schwartz space*

$$\begin{aligned}&\mathcal{S}(\mathbb{R}^d)\\ &\quad= \left\{ f \in C^\infty(\mathbb{R}^d) \;:\; f \text{ and its derivatives vanish faster than any polynomial} \right\} . \end{aligned} \tag{11.1}$$

Thus a function $f \in \mathcal{S}(\mathbb{R}^d)$ if

$$p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(f) \;:=\; \sup_{x\in\mathbb{R}^d} |\boldsymbol{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}} f(x)| \;<\; \infty \tag{11.2}$$

for every $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^d$, where

$$x^{\boldsymbol{\alpha}} \;=\; x_1^{\alpha_1}\cdots x_d^{\alpha_d} \quad \text{and} \quad D^{\boldsymbol{\beta}} \;=\; \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_d}}{\partial x_d^{\beta_d}} \, .$$

We want to topologize $\mathcal{S}(\mathbb{R}^d)$ in a way so that a sequence $f_n \to f$ if $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(f_n - f) \to 0$ for every $\boldsymbol{\alpha}, \boldsymbol{\beta}$.

Recall that a *semi-norm* 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, $p(x + y) \le p(x) + p(y)$. It is allowed that $p(x) = 0$ for $x \neq 0$. The maps $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ on the Schwartz space are semi-norms (in fact, they are norms). Note that this collection of semi-norms *separates points*, where

Definition 11.10. A collection $\mathcal{N}$ of semi-norms *separates points* if whenever $p(u) = 0$ for all $p \in \mathcal{N}$ it follows that $u = 0$.

Now endow $\mathcal{S}(\mathbb{R}^d)$ with the weakest topology such that $\mathcal{S}(\mathbb{R}^d)$ is a TVS and each of the semi-norms $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is continuous.

Claim 11.11. *$\mathcal{S}(\mathbb{R}^d)$ with this topology is a LCS*

The claim follows from the following general result.

Definition 11.12. Let X be a linear space and $\mathcal{N}$ a family of functions $f : X \to M$, M a topological space (usually $M = \mathbb{R}$ or $\mathbb{C}$). The *natural topology generated* by $\mathcal{N}$ is the weakest topology on X such that addition and all the functions in $\mathcal{N}$ are continuous.

Theorem 11.13. *Given a linear space X and a collection of semi-norms $\mathcal{N}$ that separates points, the natural topology generated by $\mathcal{N}$ makes X a locally convex space. Conversely, given a LCS X and $\mathcal{C}_0$ a neighborhood base at the origin consisting of balanced, convex, open sets, the LCS topology on X is the natural topology generated by $\mathcal{N} = \{p_U : U \in \mathcal{C}_0\}$.*

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 in $[0,\infty)$ for each $b > 0$. Next consider a set $p_U^{-1}(b,\infty)$. Let x be in this set and let $\alpha = p_U(x)$. Consider the open neighborhood $V = x + (\alpha - b)U$ then for $y \in V$ we have $y = x + (\alpha - b)y'$ with $y' \in U$ and thus

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

So $V \subset p_U^{-1}(b,\infty)$ and thus $p_U^{-1}(b,\infty)$ 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 suppose we are given a collection $\mathcal{N}$ of semi-norms that separates points. Any topology such that addition and every $p \in \mathcal{N}$ is continuous necessarily contains the collection

$$\mathcal{C} = \{x + U : x \in X \text{ and } U \in \mathcal{C}_0\},$$

where

$$\mathcal{C}_0 = \{p^{-1}[0,b) : p \in \mathcal{N} \text{ and } b \in (0,\infty)\}.$$

Consider the coarsest topology $\mathcal{T}$ containing $\mathcal{C}$. Since $p = p_U$ for $U = p^{-1}[0,1)$, we see from the above argument that all $p \in \mathcal{N}$ are continuous in $\mathcal{T}$. Thus, $\mathcal{T}$ is the natural topology generated by $\mathcal{N}$.

We must show that $(X,\mathcal{T})$ is a LCS. Because $\mathcal{N}$ separates points, $(X,\mathcal{T})$ is Hausdorff. Since $\mathcal{C}_0$ is a convex neighborhood base at the origin, the topology is locally convex. Because the collection $\mathcal{C}$ is invariant under translation ($U \in \mathcal{C} \Leftrightarrow x + U \in \mathcal{C}$), it follows that $\mathcal{T}$ is invariant under translation. Because, $zp^{-1}[0,b) = p^{-1}[0,|z|b)$ for any $z \in F$ and $b > 0$, we see that $\mathcal{C}$ is invariant under scalar multiplication. Thus $\mathcal{T}$ is invariant under scalar multiplication.

Exercise 11.1. Let $(X, \mathcal{T})$ be a linear space with a topology $\mathcal{T}$ that is invariant under translation and scalar multiplication and contains a convex neighborhood base at the origin. Show that addition and scalar multiplication are continuous and, thus, that $(X, \mathcal{T})$ is a LCS.

Conversely, let $\mathcal{T}$ denote a given LCS topology on X. Thus $\mathcal{T}$ is certainly *a* topology under which addition and all elements of

$$\mathcal{N} = \{p_U \mid U \text{ a balanced, convex, open neighborhood of } 0\}$$

are continuous. To prove it is the weakest such, we must show that any such topology contains $\mathcal{T}$. Any $U \in \mathcal{C}_0$ satisfies $U = p_U^{-1}([0,1))$. Thus any topology such that addition and all p_U are continuous certainly contains $\mathcal{C}_0$ and all its translates, and thus $\mathcal{T}$. □

LECTURE 12

Metrizable Spaces, Dual Spaces, and Tempered Distributions

Reading: Reed and Simon 1980, Ch. 5

1. Metrizable LCSs

Above we claimed that the topology on $\mathcal{S}(\mathbb{R}^d)$ 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. We will see examples of this later. 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 12.1. *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 balanced, convex, open sets; and
(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 balanced, convex, open neighborhoods of the origin the corresponding gauge function and vice versa. The details are left as an exercise.

Suppose X is metrizable. Then X has a countable neighborhood base (this is a property of metric spaces), and in particular a countable neighborhood base at the origin. Since X is a LCS we may find a balanced, convex, open set contained in each set of the base, thus obtaining a countable, balanced, convex neighborhood base at the origin.

Now, suppose X has a countable, balanced, convex neighborhood base as indicated, and let $\mathcal{T}$ denote the corresponding 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

$$d(x,y) \ = \ \sum_{j=1}^{\infty} 2^{-j} \frac{p_j(x-y)}{1+p_j(x-y)}, \tag{12.1}$$

and the metric topology $\mathcal{T}_d$.

Exercise 12.1. Prove that d defined in (12.1) is a metric. Note that the collection $\{p_U \ : \ U \in \mathcal{C}\}$ separates points since $\mathcal{C}$ is a base.

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 $x \in tU_k$

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

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. □

For a general locally convex space, not necessarily metrizable, we can define the notions of Cauchy nets and completeness.

Definition 12.2. Let X be a LCS. A net $(x_\alpha)_{\alpha\in I}$ in X is *Cauchy* if for any continuous seminorm p and any $\epsilon > 0$ there is $\alpha_0 \in I$ such that $p(x_\alpha - x_\beta) < \epsilon$ for $\alpha, \beta \ge \alpha_0$. The LCS X is *complete* if every Cauchy net is convergent to a limit in x.

Exercise 12.2. Let X be a LCS with a countable, increasing family of semi-norms as in the proof of the last theorem. Prove that a net in X is Cauchy (in the LCS sense) if and only if it is Cauchy with respect to the metric d defined in (12.1). Conclude that X is complete as an LCS if and only if it is complete as a metric space in the metric d.

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

Remark. If X is a LCS with a topology generated by a countable family of seminorms, it suffices to check that every Cauchy *sequence* is convergent to prove completeness.

Corollary 12.4. *The space $\mathcal{S}(\mathbb{R}^d)$ with the topology generated by the seminorms $p_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is a Fréchet space.*

PROOF. It is clear that $\mathcal{S}(\mathbb{R}^d)$ is a LCS generated by a countable family of seminorms. Suppose that $(f_n)_{n=1}^{\infty}$ is a Cauchy sequence. Since $(f_n)_{n=1}^{\infty}$ is Cauchy with repsect to $p_{0,0}(u) = \sup_x |u(x)|$, by completeness of $C_0(\mathbb{R}^d)$ we conclude that there is a function $f \in C_0(\mathbb{R}^d)$ such that $f_n \to f$ uniformly. Since $(f_n)_{n=1}^{\infty}$ is Cauchy with respect to $p_{\boldsymbol{\alpha},0}(u) = \sup_x |x^{\boldsymbol{\beta}} u(x)|$, we conclude that $\sup_x |x^{\boldsymbol{\beta}} f(x)| < \infty$. Similarly, for each $\boldsymbol{\beta}$, we find a limit function $g_{\boldsymbol{\beta}}$ such that $D^{\boldsymbol{\beta}} f_j \to g_{\boldsymbol{\beta}}$ uniformly and $\sup_x |x^{\boldsymbol{\beta}} g_{\boldsymbol{\beta}}(x)|$. Arguing by induction on $\boldsymbol{\beta}$, we conclude $g_{\boldsymbol{\beta}} = D^{\boldsymbol{\beta}} f$ (see Rudin 1976 Thm. 7.17). Thus $f \in \mathcal{S}(\mathbb{R}^d)$ and $f_n \to f$ in the Schwartz space. □

2. The dual space of a LCS

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

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

(It is useful to use the inner product notation to denote the pairing between elements of X and linear functionals $\ell(x) = \langle \ell,\ x \rangle$. Note that this inner product is *linear* in both factors even if we are dealing with complex spaces.) The X-weak topology on $X^\star$, $\sigma(X^\star, X)$, is also called the *weak* toplogy*. It is generated by the family of seminorms

$$p_x(\ell) = |\langle \ell, x\rangle|, \quad x \in X$$

Theorem 12.5. *If X is a LCS then $(X^\star, \sigma(X^\star, X))^\star = X$.*

Remark 12.6. That is every LCS is "reflexive" if we give $X^\star$ the weak* topology. 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 *topology matters*.

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

Theorem 12.7. *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 $U = \ell^{-1}(\mathbb{D})$ is a balanced, convex, open neighborhood of the origin in X, where $\mathbb{D}$ is the unit disc in F (so $\mathbb{D} = (-1, 1)$ if $F = \mathbb{R}$). 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,

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

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 $\ell^{-1}(\varepsilon\mathbb{D})$ contains an open set containing the origin for each $\varepsilon > 0$. But clearly

$$\left\{ x : C \sum_{j=1}^n p_j(x) < \varepsilon \right\} \subset \ell^{-1}(\varepsilon\mathbb{D}). \quad \square$$

3. 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 12.7, which may be proved similarly:

Theorem 12.8. *Let X, Y be LCSs generated by a families of semi-norms $\mathcal{S}, \mathcal{T}$ respectively. A linear map $T : X \to Y$ is continuous if and only if for any semi-norm $q \in \mathcal{T}$ there is a constant $C > 0$ and a finite collection $p_1, \ldots, p_n \in \mathcal{S}$ such that*
$$q(Tx) \le C \sum_{j=1}^n p_j(x) \quad \forall x \in X.$$

Corollary 12.9. *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
$$\left\langle \partial_j\psi, F \right\rangle = -\left\langle \psi, \partial_j F \right\rangle,$$
whenever F is a C^1 function of polynomial growth. Thus define for arbitrary $\ell \in \mathcal{S}^\star$:
$$\left\langle \psi, \partial_j \ell \right\rangle = -\left\langle \partial_j, \ell \right\rangle.$$

Proposition 12.10. *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 $\boldsymbol{\alpha} \in \mathcal{N}^d$ be a multi-index and let F be a polynomial L^1 bounded function. Then $D^{\boldsymbol{\alpha}}F$ is a tempered distribution:
$$\langle \psi, D^{\boldsymbol{\alpha}}F \rangle = (-1)^{\boldsymbol{\alpha}} \langle D^{\boldsymbol{\alpha}}\psi, F \rangle == (-1)^{\boldsymbol{\alpha}} \int_{\mathbb{R}^d} D^{\boldsymbol{\alpha}}\psi(x)F(x)\mathrm{d}x.$$

Theorem 12.11 (Structure Theorem for Tempered Distributions)**.** *Let $\ell \in \mathcal{S}(\mathbb{R}^d)$ be a tempered distribution. Then there is a polynomially bounded* continuous *function g and a multi-index $\boldsymbol{\alpha} \in \mathbb{N}^d$ such that $\ell = D^{\boldsymbol{\alpha}} g$.*

For the proof see Reed and Simon 1980, Ch. V.

LECTURE 13

Inductive Limits

Reading: Reed and Simon 1980, Ch. 5 or Simon 2015b, Ch. 9.

Consider the scale of spaces

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

where

(1) $C_c^\infty(\mathbb{R}^d)$ is the set of compactly, suported smooth functions,
(2) $\mathcal{S}(\mathbb{R}^d)$ is the Schwartz space defined above,
(3) $C_0^\infty(\mathbb{R}^d)$ is the set of smooth functions f such that f and all its derivatives vanish at ∞, and
(4) $C^\infty(\mathbb{R}^d)$ is the set of all smooth functions.

Each of these spaces has a natural complete LCS topology, which we will discuss here.

1. C_0^∞ and C^∞

We have already seen the topology on $\mathcal{S}(\mathbb{R}^d)$. The spaces $C_0^\infty(\mathbb{R}^d)$ and $C^\infty(\mathbb{R}^d)$ are similar. In fact, we can replace $\mathbb{R}^d$ by an arbitrary open subset $\Omega \subset \mathbb{R}^d$.

Definition 13.1. Let Ω be an open subset of $\mathbb{R}^d$ and let $C^\infty(\Omega)$ denote the set of all infinitely differentiable maps $f : \Omega \to F$. The space $C_0^\infty(\Omega)$ denotes the set of all maps $f \in C^\infty(\Omega)$ such that for every $\alpha \in \mathbb{N}^d$, $D^{\boldsymbol{\alpha}} f(x) \to 0$ whenever x approaches Ω^c or ∞.

Theorem 13.2. *Let $\Omega \subset \mathbb{R}^d$ be open.*

(1) With the LCS topology generated by the family of seminorms

$$\mathcal{N}_0 = \left\{ p_{\boldsymbol{\alpha}}(f) \;=\; \sup_{x\in\Omega} |D^{\boldsymbol{\alpha}} f(x)| \;:\; \boldsymbol{\alpha} \in \mathbb{N}^d \right\}\,,$$

$C_0^\infty(\Omega)$ is a Fréchet space.

(2) With the LCS topology generated by the family of seminorms

$$\mathcal{N} \;=\; \left\{ q_{\boldsymbol{\alpha},K}(f) \;=\; \sup_{x\in K} |D^{\boldsymbol{\alpha}} f(x)| \;:\; \boldsymbol{\alpha} \in \mathbb{N}^d \text{ and } K \subset \Omega \text{ is compact} \right\}\,,$$

$C^\infty(\Omega)$ is a Fréchet space.

Remark. Note that we can embed $C_0^\infty(\Omega) \hookrightarrow C_0^\infty(\mathbb{R}^d)$ by extending f to be zero on Ω^c. This imbedding is continuous and makes $C_0^\infty(\Omega)$ a closed subspace of $C_0^\infty(\mathbb{R}^d)$.

SKETCH OF PROOF. The proof that C_0^∞ and C^∞ are locally convex spaces follows closely the result for the Schwartz space. In fact, the proof for C_0^∞ is actually a bit easier since we don't need to verify that the resulting limiting function is polynomially bounded. Since $\mathcal{N}_0$ is countable, it follows immediately that $C_0^\infty(\Omega)$ is a Fréchet space.

Because there are uncountably many compact sets $K \subset \Omega$, the family $\mathcal{N}$ is not countable. However, one can choose a countable family $K_1 \subset K_2 \subset \cdots \subset \Omega$ of compact sets with $\Omega = \bigcup_n K_n^\circ$. For example,

$$K_n = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) \geq 2^{-n} \text{ and } |x| \leq 2^n\} . \tag{13.1}$$

Any compact set $K \subset \Omega$ is contained in one of the K_n, so $q_{\boldsymbol{\alpha},K} \leq q_{\boldsymbol{\alpha},K_n}$. It follows that the topology generated by $\mathcal{N}$ is the same as that generated by the countable collection $\{q_{\boldsymbol{\alpha},K_n} : \boldsymbol{\alpha} \in \mathbb{N}^d , n \in \mathbb{N}\}$. Thus $C^\infty(\Omega)$ is a Fréchet space as claimed. □

2. Inductive Limits and C_c^∞

The space $C_c^\infty(\Omega)$ requires a new concept: *the inductive limit topology*. The key idea is to think of $C_c^\infty(\Omega)$ as the union $\bigcup_n C_0^\infty(K_n^\circ)$ with $(K_n)_{n=1}^\infty$ as in (13.1). We want to put a topology on the union $C_c^\infty(\Omega)$ that is consistent with the LCS topology of each subspace $C_0^\infty(K_n^\circ)$. The key observation that allows us to do this is the fact that the topology on $C_0^\infty(K_n^\circ)$ obtained by considering it as a subspace of $C_0^\infty(K_{n+1}^\circ)$ is the *same* as its given LCS topology.

Theorem 13.3. *Let X be a linear space with $(X_n)_{n=1}^\infty$ a family of subspaces of X such that $X_n \subset X_{n+1}$ and $X = \bigcup_n X_n$. Suppose that each X_n has a locally convex topology $\mathcal{T}_n$ such that the restriction of $\mathcal{T}_{n+1}$ to X_n is $\mathcal{T}_n$. Let $\mathcal{U}$ be the collection of balanced, convex, absorbing sets $U \subset X$ such that $U \cap X_n$ is open in X_n for each n. If $\mathcal{T}$ is the natural topology generated by $\mathcal{U}$, then*

(1) $\mathcal{T}$ is the strongest locally convex topology on X such that the injections $X_n \hookrightarrow X$ are continuous.
(2) The resriction of $\mathcal{T}$ to X_n is the given topology $\mathcal{T}_n$ on X_n.
(3) If each $(X_n, \mathcal{T}_n)$ is complete, then so is $(X, \mathcal{T})$.

Definition 13.4. The LCS X in Thm. 13.3 is *the strict inductive limit of* $(X_n)_{n=1}^\infty$.

Remarks. 1) For the proof of Thm. 13.3, see Simon 2015b, Thm. 9.1.1. 2) Recall that the restriction of a topology $\mathcal{T}$ on X to a subspace $Y \subset X$ is the relative topology consiting of the sets $U \cap Y$ for $U \in \mathcal{T}$. 3) The inductive limit is *strict* because the topology $\mathcal{T}_n$ is the restriction of $\mathcal{T}_{n+1}$ to X_n. There is a more general definition in which one has continuous embeddings $\phi_n : X_n \to X_{n+1}$, see Conway 2007, §IV.5.

Theorem 13.5. *Let X be the strict inductive limit of an increasing family $(X_n)_{n=1}^\infty$ of LCSs, let Y be an LCS, and let $T : X \to Y$ be a linear map. Then T is continuous if and only if the restriction $T_j = T|_{X_j}$ is a continuous map for each j.*

PROOF. If T is continuous, then the restrction of T to any subspace is continuous. Conversely, suppose that the restriction T_j is continuous for each j. Since Y is an LCS, to prove that T is continuous it suffices to prove that $T^{-1}(V)$ is open for every balanced, convex, open set in Y. Because T is linear, the inverse image of a balanced, convex set is balanced and convex. Since $T^{-1}(V) \cap X_j = T_j^{-1}(V)$, we see that $T^{-1}(V) \cap X_j$ is balanced, convex and open for each j. Thus $T^{-1}(V)$ is open, by definition. □

Theorem 13.6. *Let X be the strict inductive limit of an increasing family $(X_n)_{n=1}^{\infty}$ of LCSs. Suppose that each X_n is a proper, closed subspace of X_{n+1}. If $(x_j)_{j=1}^{n}$ is a sequence in X, then $(x_j)_{j=1}^{\infty}$ converges if and only if, for some X_m, all $x_j \in X_m$ and $(x_j)_{j=1}^{\infty}$ converges in X_m.*

Lemma 13.7. *Let X be the strict inductive limit of an increasing family $(X_n)_{n=1}^{\infty}$ of LCSs. If each X_n is a proper, closed subspace of X_{n+1}, then each X_n is closed in X.*

PROOF. Since X_n is closed in X_{n+1} for each n, we see that X_n is closed in X_m for any $m > n$. Let x be a limit point of X_n. Then $x \in X_m$ for some m. If $m \leq n$ then $x \in X_n$. If $m > n$, then $x \in X_n$ since X_n is a closed subspace of X_m. □

PROOF OF THM. 13.6 . By Lem. 13.7, we see that if all x_j are in X_m and $x_j \to x$ in X, then $x \in X_m$. Hence to prove the result we just need to show that any convergent sequence lies in a single X_m for some m. We will show the contrapositive: *if $(x_j)_{j=1}^{\infty}$ does not lie in any one of the X_m, then $(x_j)_{j=1}^{\infty}$ does not converge.* By passing to a subsequence, we may suppose that we have $x_j \in Y_{j+1} \setminus Y_j$ with $Y_j = X_{m_j}$ and $X_{m_1} \subsetneq X_{m_2} \subsetneq X_{m_3} \cdots$.

We will show that $(x_j)_{j=1}^{\infty}$ does not converge by constructing a continuous linear functional $L : X \to F$ such that $L(x_j) \to \infty$. We will define L as a sum $L = \sum_{j=1}^{\infty} L_j$ with L_j constructed recursively. Let $\ell_1 : \operatorname{span}(Y_1, x_1) \to F$ be a linear functional such that $\ell_1|_{Y_1} \equiv 0$ and $\ell_1(x_1) = 1$. Since Y_1 is closed, ℓ_1 is continuous and thus, by the Hahn-Banach theorem, there is an extension L_1 of ℓ_1 to X. Now, given the maps $L_1, \ldots, L_{n-1}$, let $\ell_n : \operatorname{span}(Y_n, x_n) \to F$ be a linear functional such that $\ell_n|_{Y_n} \equiv 0$ and $\ell_n(x_n) = n - \sum_{j=1}^{n-1} L_j(x_n)$. By Hahn-Banach, since Y_n is closed, there is an extension L_n of ℓ_n to X. Let $L = \sum_{n=1}^{\infty} L_n$. If $x \in Y_m$, then $L_n(x) = 0$ for $n \geq m$ so $L|_{Y_m} = \sum_{n=1}^{m-1} L_n$ is continuous. Thus, L is well-defined on $X = \bigcup_m Y_m$ and L is continuous on X by Thm. 13.5. Since $L(x_n) = n$, we conclude that $(x_n)_{n=1}^{\infty}$ cannot converge. □

Theorem 13.8. *Let X be the strict inductive limit of an increasing family $(X_n)_{n=1}^{\infty}$ of LCSs. If each X_n is a proper, closed subspace of X_{n+1}, then X is not metrizable.*

PROOF. Suppose that X had a countable neighborhood base at the origin. Without loss of generality, we may suppose that the base $(U_n)_{n=1}^{\infty}$ is decreasing, $U_n \supset U_{n+1}$ for each n. Since X is Hausdorff, $\bigcap_n U_n = \{0\}$. Let $x_n \in U_n \setminus X_n$ — note that U_n cannot be a subset of X_n since X_n is a proper subspace and U_n is absorbing. Then $x_n \to 0$ but we cannot have $(x_n)_{n=1}^{\infty}$ contained in any one X_m by construction, contradicting Thm. 13.6. Thus X has no countable neighborhood base at 0 and is hence not metrizable. □

3. A simple example

Consider the three spaces

$$\mathcal{F} \subset c_0 \subset \mathcal{E} \ ,$$

where c_0 and ℓ^{∞} are the Banach spaces defined above and

(1) $\mathcal{E}$ = all F valued sequences, and
(2) $\mathcal{F}$ = sequences that are eventually zero.

We have

(1) $\mathcal{E}$ is a Frechét space, generated by the seminorms

$$p_N(\boldsymbol{a}) \ = \ \sup_{n \leq N} |a_n| \ .$$

(2) c_0 is a Banach space (so a Frechét space) generated by the single norm $\|\boldsymbol{a}\| = \sup_n |a_n|$.
(3) $\mathcal{F}$ is the inductive limit of $(F^n)_{n=1}^{\infty}$, where we embed $F^n \hookrightarrow \mathcal{F}$ as $F^n \mapsto$ sequences that are zero after the n-th term.
(4) $\mathcal{E}^\star = \mathcal{F}$ and $\mathcal{F}^\star = \mathcal{E}$; the dual of c_0 is, of course, ℓ^1.

LECTURE 14

$C_c^\infty(\mathbb{R}^d)$ and Distributions

Reading: Reed and Simon 1980, Ch. 5 and Simon 2015b, Ch. 11. For more details about distributions, see Rudin 1991, Ch. 6.

1. $C_c^\infty(\Omega)$ as an inductive limit

Let $\Omega \subset \mathbb{R}^d$ be open and let $(K_n)_{n=1}^\infty$ be an increasing sequence of compact subsets of Ω such that $\Omega = \bigcup_n K_n^\circ$. We topologize $C_c^\infty(\Omega)$ with the inductive limit topology given by $C_0^\infty(K_n^\circ)$. Since $C_0^\infty(K_n^\circ)$ is a Fréchet space for each n, $C_c^\infty(\Omega)$ is complete. Since $C_0^\infty(K_n^\circ)$ is a closed subspace of $C_0^\infty(K_{n+1}^\circ)$, the resulting topology is not metrizable. At the moment, this construction appears to depend on the choice of the sequence $(K_n)_{n=1}^\infty$. That it does not follows from the following

Exercise 14.1. Reed and Simon 1980, Ch.5, problem 46:

(a) Suppose X is the strict inductive limit of $(X_n)_{n=1}^\infty$ and that $(Y_n)_{n=1}^\infty$ is an increasing family of subspaces of X so that for any n, there is N with $X_n \subset Y_N$. Prove that X is the strict inductive limit of $(Y_n)_{n=1}^\infty$.

(b) Let $K \subset \Omega \subset \mathbb{R}^d$ with K compact and Ω open. Prove that if $C_c^\infty(\Omega)$ has the inductive limit topology given by $(C_c^\infty(K_n))_{n=1}^\infty$ for some family as described above, then the restriction of this topology to $C_0^\infty(K^\circ)$ is the Fréchet topology for this space defined above.

(c) Prove that the topology on $C_c^\infty(\Omega)$ is independent of the choice of the increasing family $(K_n)_{n=1}^\infty$ of compact sets.

2. Distributions

Let $\Omega \subset \mathbb{R}^d$ be open. A *distribution on* Ω is an element of $\mathcal{D}^\star(\Omega) = (C_c^\infty(\Omega))^\star$.

Proposition 14.1. *Let $T \in \mathcal{S}^\star(\mathbb{R}^d)$ be a tempered distribution. Then the restriction of T to $C_c^\infty(\mathbb{R}^d)$ is a distribution and the restriction map is one-to-one, i.e., if $T(\phi) = S(\phi)$ for all $\phi \in C_c^\infty(\mathbb{R}^d)$ then $T = S$.*

PROOF. To prove that T is continuous, it suffices (by Thm. 13.5) to prove that $T|_{C_0^\infty(K^\circ)}$ is continuous for any compact set $K \subset \mathbb{R}^d$. This follows since $C_0^\infty(K^\circ)$ is a closed subspace of $\mathcal{S}(\mathbb{R}^d)$. That the restriction map is one-to-one follows since $C_c^\infty(\mathbb{R}^d)$ is dense in $\mathcal{S}(\mathbb{R}^d)$. □

Thus we have the embedding $\mathcal{S}^\star \subset \mathcal{D}^\star$. Distributions in $\mathcal{D}^\star$ need not be bounded at ∞ – any locally integrable function, like $\exp(\exp(|x|^2))$ is a distribution.

As for tempered distributions, the derivatives $D^\alpha T$ of a distribution on Ω can be defined using integratation by parts

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

for $u \in C_c^\infty(\Omega)$. Using Thm. 13.5 it is easy to see that $D^{\boldsymbol{\alpha}} : \mathcal{D}^\star(\Omega) \to \mathcal{D}^\star(\Omega)$ is continuous. A direct analogue of the structure theorem 12.11 does not hold in this context. However, there is a local version:

Theorem 14.2. *Let $\Omega \subset \mathbb{R}^d$ be open and let T be a distribution on Ω. If $K \subset \Omega$ is compact then there is $F_K \in C_c(\Omega)$ and $\boldsymbol{\alpha}_K \in \mathbb{N}^d$ such that if* $\operatorname{supp} u \subset K$, *then*

$$\langle T,\, u\rangle \;=\; \langle D^{\boldsymbol{\alpha}_K} F_K,\, u\rangle\,.$$

SKETCH OF PROOF. Let U be open with $\overline{U}$ compact and $K \subset U \subset \overline{U} \subset \Omega$. Since T is continuous on $C_c^\infty(\Omega)$ there are $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n \in \mathbb{N}^d$ and C, depending on U, such that for $u \in C_0^\infty(U)$ we have

$$|\langle T,\, u\rangle| \;\le\; \sum_{j=1}^{n} \sup_{x\in U} \left| D^{\boldsymbol{\beta}_j} u(x) \right|.$$

Now let Φ be a smooth function that is one on K and vanishes on U^c. It follows that $\Phi T \in \mathcal{S}^\star(\mathbb{R}^d)$, so $\Phi T = D^{\boldsymbol{\alpha}} F$ for some F, with $F \equiv 0$ on U^c since $\Phi \equiv 0$ there. □

Corollary 14.3. *Let $\Omega \subset \mathbb{R}^d$ be open and let T be a distribution on Ω. Then there are sequences $(F_n)_{n=1}^\infty$ in $C_c(\Omega)$ and $(\boldsymbol{\alpha}_n)_{n=1}^\infty$ in $\mathbb{N}^d$ such that*

1. *For any compact $K \subset \Omega$ at most finitely many F_n are non-zero on K, and*
2. *$T = \sum_{n=1}^\infty D^{\boldsymbol{\alpha}_n} F_n$.*

Remark. The point here is that, the distribution can become less and less regular as we approach the boundary of Ω or ∞. Note that for any $u \in C_c^\infty(\mathbb{R}^d)$ only finitely many terms contribute to the sum $\langle T,\, u\rangle = \sum_n \langle D^{\boldsymbol{\alpha}_n} F_n,\, u\rangle$.

SKETCH OF PROOF. Write $\Omega = \bigcup_n K_n$ with K_n an increasing sequence of compact sets with $K_n \subset K_{n+1}^\circ$. Let $(\phi_n)_{n=1}^\infty$ be a sequence of smooth functions such that $\phi_n \equiv 0$ on K_n and $\phi_n \equiv 0$ on K_{n+1}^c.

Let $L_1 = K_1$ and, for $n > 1$, let $L_n = K_n \setminus L_{n-1}^\circ$. Then $(L_n)_{n=1}^\infty$ is a sequence of compact subsets of Ω with $\bigcup L_n = \Omega$ and any $x \in \Omega$ is in at most two of the L_n. Let $\Psi_n = \phi_n - \phi_{n-1}$ (with $\phi_0 \equiv 0$). Then $(\Psi_n)_{n=1}^\infty$ satisfy

1. $\sum_{n=1}^\infty \Psi_n \equiv 1$, and
2. $\operatorname{supp} \Psi_n \subset L_n \cup L_{n+1}$, in particular for any x at most three of the $\Psi_n(x)$ are non-zero.

Decompose T as $\sum_n \Psi_n T$. Following the proof of Thm. 14.2, write each distribution $\Psi_n T = D^{\boldsymbol{\alpha}_n} F_n$ with a suitably chosen $F_n \in C_c(\Omega)$. Looking at the proof of Thm. 14.2, we see that we may choose F_n with $\operatorname{supp} F_n \subset L_{n-1} \cup L_n \cup L_{n+1} \cup L_{n+2}$. □

The space $\mathcal{D}^\star(\Omega)$ is a complete LCS in the weak-$\star$ topology. Recall that a measurable map $f : \Omega \to K$ is *locally integrable* if $\int_K |f(x)|dx < \infty$ for any compact subset. The set of locally integrable functions on Ω is denoted $L^1_{\mathrm{loc}}(\Omega)$.

Exercise 14.2. Give $L^1_{\mathrm{loc}}(\Omega)$ the LCS topology generated by the semi-norms $p_K(f) = \int_K |f(x)|dx$ for compact $K \subset \Omega$. Show that $L^1_{\mathrm{loc}}(\Omega)$ is a Fréchet space.

Let $f \in L^1_{\mathrm{loc}}(\Omega)$. We think of f as an element of $\mathcal{D}^\star(\Omega)$ by associating it to the linear functional

$$\langle f,\, u\rangle \;=\; \int_\Omega f(x)u(x)dx\,.$$

This provides an embedding $L^1_{\text{loc}}(\Omega) \hookrightarrow \mathcal{D}^\star(\Omega)$.

Exercise 14.3. Show that this embedding of $L^1_{\text{loc}}(\Omega)$ is a continuous map. Note that to do this, it suffices to show that the map $f \mapsto \langle f,\ u\rangle$ is continuous for each $u \in C_c^\infty(\Omega)$.

In particular, we have an embedding $C_c^\infty(\Omega) \hookrightarrow \mathcal{D}^\star(\Omega)$.

Exercise 14.4. Show that the embedding $C_c^\infty(\Omega) \hookrightarrow \mathcal{D}^\star(\Omega)$ is continuous.

Derivatives are defined on $\mathcal{D}^\star(\Omega)$ via integration by parts. We want to see that they are continuous. There is a more general class of operators for which we can define continuity.

Definition 14.4. Let $T, T' : C_c^\infty(\Omega) \to C_c^\infty(\Omega)$ be continuous linear maps. We say that T' is the transpose of T if

$$\left\langle T'v,\ u\right\rangle \ = \ \langle v,\ Tv\rangle$$

for every $u, v \in C_c^\infty(\Omega)$.

Note that the transpose T' is uniquely determined by T (if it exists) and $(T')' = T$. This transpose operation obeys the usual rules, namely if T and S have transposes, then

(1) $aT + bS$ has a transpose and $(aT + bS)' = aT' + bS'$ for any $a, b \in F$, and
(2) ST has a transpose and $(ST)' = T'S'$.

Theorem 14.5. *Let $T : C_c^\infty(\Omega) \to C_c^\infty(\Omega)$ be a continuous map with a transpose. Let ℓ be a distribution and define $T\ell$ by*

$$\langle T\ell,\ u\rangle \ = \ \left\langle \ell,\ T'u\right\rangle \ .$$

Then $T\ell \in \mathcal{D}^\star(\Omega)$ and $T : \mathcal{D}^\star(\Omega) \to \mathcal{D}^\star(\Omega)$ is a continuous linear map.

PROOF. To see that $T\ell$ is a distribution, we must show that $u \mapsto \langle \ell,\ T'u\rangle$ is continuous. But this map is the composition of the continuous maps $u \mapsto T'u$ and $v \mapsto \langle \ell,\ v\rangle$

Linearity of T on $\mathcal{D}^\star(\Omega)$ is clear from the definition. To see that the map is continuous, note that we must show that $\ell \mapsto \langle \ell,\ T'u\rangle$ is continuous for any $u \in C_c^\infty(\Omega)$. This follows from the definition of the weak-$\star$ topology, since $T'u \in C_c^\infty(\Omega)$. ☐

Exercise 14.5. Show that the following are continuous linear maps with transposes on $C_c^\infty(\mathbb{R}^d)$:

(1) Multiplication by a function $f \in C^\infty(\Omega)$, $u \mapsto fu$.
(2) Differentiation to any order $\alpha \in \mathbb{N}^d$, $u \mapsto D^\alpha u$.
(3) Translation by $y \in \mathbb{R}^d$, $u \mapsto u(\cdot - y)$.
(4) Convolution with compactly supported $F \in L^1(\mathbb{R}^d)$,

$$u \ \mapsto \ F * u(x) \ = \ \int_{\mathbb{R}^d} F(x-y)u(y)dy \ .$$

(5) Composition with a diffeomorphism $\Phi : \mathbb{R}^d \to \mathbb{R}^d$, $u \mapsto u \circ \Phi$.

All of the operations in the previous exercise lift to distributions. These are most of the operations of smooth analysis, with the exception of products. In general products of distrbutions cannot be defined. However, there is a way to do this for distributions acting on distinct variables:

Exercise 14.6. Suppose that $\Omega = A \times B$ with $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^m$ open sets.

(1) Suppose that $\ell_A \in \mathcal{D}^\star(A)$ and $u \in C_c^\infty(\Omega)$. Define, for each $y \in B$,

$$\ell_A(u)(y) \;=\; \langle \ell_A,\, u(\cdot, y)\rangle \;,$$

where $u(\cdot, y)$ is the element of $C_c^\infty(A)$ obtained by restricting u to the level set $A \times \{y\}$. Show that this map is a continuous map from $C_c^\infty(\Omega) \to C_c^\infty(B)$.

(2) Suppose that $\ell_A \in \mathcal{D}^\star(A)$ and $\ell_B \in \mathcal{D}^\star(B)$ and define

$$\ell_B \ell_A(u) \;=\; \langle \ell_B,\, \ell_A(u)\rangle$$

with $\ell_A(u)$ as in part (1). Prove that $\ell_B\ell_A \in \mathcal{D}^\star(\Omega)$ and that $\ell_B\ell_A = \ell_A\ell_A$.

3. Spaces of Distributions

Each of the spaces

$$C_c^\infty(\Omega) \subset C_0^\infty(\Omega) \subset C^\infty(\Omega)$$

is dense in the larger space. It follows by arguments similar to Prop. 14.1 that the corresponding spaces of distributions satsify a reverse inclusion

$$\mathcal{E}^\star(\Omega) \subset \mathcal{D}_0^\star(\Omega) \subset \mathcal{D}^\star(\Omega)$$

where

(1) $\mathcal{E}^\star(\Omega) = (C^\infty(\Omega))^\star$, and
(2) $\mathcal{D}_0^\star(\Omega) = (C_0^\infty(\Omega))^\star$.

Distributions in $\mathcal{E}^\star$ have compact support, where

Definition 14.6. 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 $\mathcal{D}_0^\star(\Omega)$ are "bounded," in the sense that they satisfy an analogue of the Structure Theorem:

Theorem 14.7. *Let $\Omega \subset \mathbb{R}^d$ be open and let $T \in \mathcal{D}_0^\star(\Omega)$, then there is a bounded continuous function F on Ω and $\boldsymbol{\alpha} \in \mathbb{N}^d$ such that $\langle T,\, u\rangle = \langle D^{\boldsymbol{\alpha}} F,\, u\rangle$.*

SKETCH OF PROOF. We will use some facts from harmonic analysis. The basic idea to the proof is the following. If $T \in \mathcal{D}_0^\star(\Omega)$ then there are finitely many $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ such that

$$|\langle T,\, u\rangle| \;\le\; C \sum_{j=1}^{n} \sup_{x\in\Omega} |D^{\boldsymbol{\alpha}_j} u(x)| \;.$$

Now it turns out that for any smooth function that vanishes at ∞ fast enough there are constants such that

$$\sup_x |D^{\boldsymbol{\alpha}} u(x)| \;\le\; C_{\boldsymbol{\alpha},n} \sup_x |(-\Delta + 1)^n u(x)|$$

where $-\Delta = -\nabla\cdot\nabla$ is the Laplacian and $2n > |\boldsymbol{\alpha}|$ — this is closely related to Sobolev inequalities, see Simon 2015a, Thm. 6.3.2. Thus we can replace the finite family of seminorms by the bound

$$|\langle T,\, u\rangle| \;\le\; C \sup_x |(-\Delta+1)^n u| \;.$$

Now it turns out that there is a continuous map $(-\Delta+1)^{-n}$ which is its own transpose — for $\Omega = \mathbb{R}^d$ it is given by convolution with an explicit function, a *Bessel kernel*. The distribution $(-\Delta+1)^{-n}T$ satisfies the bound

$$\left|\left\langle(-\Delta+1)^{-n}T,\, u\right\rangle\right| \;\leq\; C \sup_x |u(x)| \,.$$

Thus $(-\Delta+1)^{-n}T$ can be extended to give a continuous linear function on $C_0(\Omega)$. By the Riesz representation theorem there is a finite Borel measure μ on Ω such that $(-\Delta+1)^{-n}T = \mu$. Convolving with $(-\Delta+1)^{-1}$ one more time we obtain

$$(-\Delta+1)-n-1T = (-\Delta+1)^{-1} * \mu \,,$$

where $G = (-\Delta+1) * \mu$ is a continuous function (actually it is even C^1). Thus $T = (-\Delta+1)^{n+1}G$. This is close to what was claimed, and the actual result can be obtained by a function F such that $D^{\boldsymbol{\beta}}F = (-\Delta+1)^{n+1}G$ where $\boldsymbol{\beta} = (2(n+1), 2(n+1), \cdots, 2(n+1))$ — see Simon 2015b, §6.2, Problem 8.q12w $\square$

Part 4

More about Banach Spaces

LECTURE 15

Baire category theorem and its consequences

Reading: Lax 2002, Ch. 10, Simon 2015b, §5.4

1. Baire Category Theorem

In this lecture we will consider several results about Banach spaces that follow from the *Baire Category Theorem*:

Theorem 15.1 (Baire Category Theorem)**.** *Let X be a complete metric space. If $(U_n)_{n=1}^{\infty}$ is a countable family of open dense sets, then $\bigcap_n U_n$ is dense.*

Remark. The theorem was proved for $X = \mathbb{R}^n$ by Baire 1899. The general version is due to Kuratowski 1930 and Banach 1930.

Before proving the theorem, let us present some consequences and discuss the name. Let X be a metric space $K \subset X$ is *nowhere dense* if $\overline{K}$ has empty interior. An equivalent statement of the Baire category theorem is the following:

Corollary 15.2 (Nowhere dense version of the Baire Category Theorm)**.** *Let X be a complete metric space. If $(K_n)_{n=1}^{\infty}$ is a countable family of nowhere dense subsets in X, then $\bigcup_n K_n$ is nowhere dense.*

Exercise 15.1. Prove that Thm. 15.1 and Cor. 15.2 are equivalent.

The word "category" in all this comes from the following definitions. A set K in a metric space is said to be *first category* if it is the countable union of nowhere dense sets and is *second category* otherwise. The Baire Category Theorem implies the following statements that are responsible for the name:

(1) A first category subset of a complete metric space has empty interior.
(2) A complete metric space is second category.

The Baire Category Theorem follows from the following

Lemma 15.3. *Let X be a complete metric space. If $(U_n)_{n=1}^{\infty}$ is a countable family of open dense sets, then $\bigcap_n U_n$ is non-empty.*

Exercise 15.2. Show that Lemma 15.3 implies Theorem 15.1. (Hint: use the lemma to prove that $\bigcap_n U_n \cap \overline{B_r(x)} \neq 0$ for any closed ball $\overline{B_r(x)} \subset X$.)

PROOF OF LEMMA 15.3. Suppose that $(U_n)_{n=1}^{\infty}$ is a countable family of open dense sets. Since U_1 is open and dense, it is a non empty open set. Thus, there are $x_1 \in X$ and $r_1 > 0$ such that $\overline{B_r(x_1)} \subset U_1$. Proceeding recursively, for each $n = 2, 3, \ldots$, since U_n is open and dense, we can find a point x_n and $0 < r_n \leq {}^{r_{n-1}}/2$ such that $\overline{B_{r_n}(x_n)} \subset B_{r_{n-1}}(x_{n-1}) \cap U_n$. Because $r_n \leq 2^{n-2} r_1$ and $x_n \in \underline{B_{r_m}(x_m)}$ for $n \geq m$, we see that $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence. Let $x = \lim_n x_n$, then $x \in \overline{B_{r_n}(x_n)} \subset B_{r_{n-1}}(x_{n-1}) \cap U_n$ for each n, so $x \in \bigcap_n U_n$. $\square$

In our applications of Thm. 15.1 we will use the following simpler version which follows directly from Lem. 15.3:

Lemma 15.4. *Let X be a complete metric space and let $(C_n)_{n=1}^{\infty}$ be a sequence of closed sets. If $X = \bigcup_n C_n$, then there is m such that C_m has non-empty interior*

PROOF. We show the contrapositive: *if $(C_n)_{n=1}^{\infty}$ are nowhere dense, then $X \neq \bigcup_n C_n$*. This follows directly from Lemma 15.3 applied to the open dense sets $U_n = X \setminus C_n$. □

2. Principle of Uniform Boundedness

The *Principle of Uniform Boundedness* (PUB), or *Banach-Steinhaus Theorem*, is the following result:

Theorem 15.5 (Principle of Uniform Boundedness; Banach and Steinhaus 1927)**.** *Let X be a Banach space, let Y be a normed space and let $\mathcal{L}$ be a collection of bounded linear maps from X to Y. If $\sup_{T \in \mathcal{L}} \|T(x)\|_Y < \infty$ for each $x \in X$, then $\sup_{T \in \mathcal{L}} \|T\| < \infty$.*

Remark. That is, *if a family of bounded linear maps is pointwise bounded, then it is uniformly bounded.*

The Principle of Uniform Boundedness follows from the following more general result about functions on a metric space:

Theorem 15.6 (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 $x \in X$, i.e.,*

$$\sup_{f \in \mathcal{F}} |f(x)| < \infty \quad \text{for each } x \in X,$$

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. Note that by assumption $X = \bigcup_n \{x : |f(x)| \leq n \text{ for all } f \in \mathcal{F}\}$, where each of the sets $\{x : |f(x)| \leq n \; \forall f \in \mathcal{F}\}$ is closed. By Lemma 15.4, at least one of the sets has non-empty interior, which is to say it contains an open set U. This is the open set claimed in the theorem. □

Suppose now that X is a Banach space and each function $f \in \mathcal{F}$ is

(1) sub-additive: $f(x+y) \leq f(x) + f(y)$; and
(2) absolutely homogeneous: $f(ax) = |a| f(x)$.

For instance each f could be of the form $f(x) = \|T(x)\|$ for some linear map T. In this case, we can translate the local bound provided by Thm. 15.6 into a uniform bound.

Corollary 15.7 (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. If $\mathcal{F}$ is bounded at each point $x \in X$, then functions $f \in \mathcal{F}$ are uniformly bounded, i.e., 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 hold. 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 $f \in \mathcal{F}$,

$$f(y) = f(y + x_0 - x_0) \le f(y + x_0) + f(x_0) \le 2M.$$

Thus for arbitrary $x \in X$ and $f \in \mathcal{F}$,

$$f(x) = \frac{2\|x\|}{\epsilon} f\left(\frac{\epsilon}{2\|x\|} x\right) \le \frac{4M}{\epsilon}\|x\|. \quad \square$$

Exercise 15.3. Show that Cor. 15.7 implies Thm. 15.5. (Hint: apply Cor. 15.7 to the family of functions $f(x) = \|T(x)\|$ with $T \in \mathcal{L}$.)

3. Open Mapping Theorem and related results

The second big consequence of the Baire category theorem is the *open mapping theorem*. A map $f : X \to Y$ between topological spaces is said to be *open* if it maps open sets to open sets, i.e., $f(U)$ is open if $U \subset X$ is open. This is in contrast to continuity, which requires $f^{-1}(V)$ to be open if $V \subset Y$ is open. Continuous maps need not be open, and open maps need not be continuous:

Exercise 15.4. 1) Show that $\phi(x) = x^2$ is not an open map from $\mathbb{R} \to \mathbb{R}$. 2) Let $f : [0, 2\pi) \to T^1 = \{z : |z| = 1\}$ be the map $f(\theta) = e^{i\theta}$. Then f is continuous, one-to-one and surjective. Show that the inverse map f^{-1} open but not continuous.

The open mapping theorem shows that surjective, bounded, linear maps do not suffer from this problem:

Theorem 15.8 (Open Mapping Theorem)**.** *Let X and Y be Banach spaces. If $T : X \to Y$ is a bounded linear map with $T(X) = Y$, then T is an open map.*

PROOF. For each n, let $A_n = T(B_n^X(0))$ where $B_n(0) = \{x \in X : \|x\| < n\}$ is the open ball of radius n in X. Since T is surjective, we have $Y = \bigcup_n A_n$. By Lemma 15.4, there is m such that $\overline{A_m}$ has non-empty interior. Thus there is $y_0 \in \overline{A_m}$ and $r > 0$ such that $B_r^Y(y_0) \subset \overline{A_m}$, where $B_r^Y(y_0) = \{y \in Y : |y - y_0| \le r\}$.

If $\|y\| < 1$ then

$$ry = (ry + y_0) - y_0 \in \overline{A_m} + \overline{A_m} \in \overline{A_{2m}}\ .$$

Thus $B_1^Y(0) \subset \overline{A_g}$ where $g = {}^{2m}/_{r}$. Given $y \in Y$, we have $\frac{y}{\|y\|} \in \overline{B_1^Y(0)}$. It follows that for any $y \in Y$ and $\varepsilon > 0$ we can find $x \in X$ such that

$$\|x\| \le g\|y\| \quad \text{and} \quad \|y - Tx\| < \epsilon\ .$$

(To see this note that $\frac{y}{\|y\|} \in \overline{B_1^Y(0)} \subset \overline{A_g}$.)

We will now show that $B_1^Y(0) \subset A_{2g}$. Let $\|y\| < 1$. Proceeding recursively, we can find $x_1, x_2, \ldots$ such that

$$\|x_n\| < g 2^{1-n} \quad \text{and} \quad \|y - T(x_1 + \cdots + x_n)\| < 2^{-n}$$

for each n. The sequence $\xi_n = x_1 + \cdots + x_n$ is then a Cauchy sequence; let ξ denote its limit. We have

$$\|\xi\| < g \sum_{j=1}^{\infty} 2^{1-j} = 2g$$

and $y = T(\xi)$, by continuity of T. Thus $y \in A_{2g}$ and, as y was arbitrary, $B_1^Y(0) \subset A_{2g}$.

Now let $U \subset X$ be open and let $y \in T(U)$. There is $x \in X$ and $\varepsilon > 0$ such that $y = T(x)$ and $B_\varepsilon^X(x) \subset U$. Since $T(B_\varepsilon^X(x)) = y + \frac{\varepsilon}{2g}T(B_{2g}^X(0)) = y + \frac{\varepsilon}{2g}A_{2g}$, we see that $B^Y_{\frac{\varepsilon}{2g}}(y) \subset T(U)$ and thus that $T(U)$ is open. □

The open mapping theorem has a number of important consequences.

Corollary 15.9. *Let X and Y be Banach spaces and let $T : X \to Y$ a bounded linear map. If* ran T *is closed and* $\ker T = \{0\}$*, then there is $\varepsilon > 0$ such that*

$$\|Tx\|_Y \ \geq \ \varepsilon\|x\|_X \quad \textit{for all } x \in X.$$

The hypothesis that ran T is closed is crucial here:

Exercise 15.5. Consider the map $T : L^1(0,1) \to L^1(0,1)$ given by $Tf(x) = xf(x)$. Show that $\ker T = \{0\}$ but that ran T is not closed and that there is no $\varepsilon > 0$ such that $\|Tf\|_1 \geq \varepsilon\|f\|_1$.

PROOF. Since ran T is closed it is a Banach space in the norm of Y. Thus we may replace Y by ran T and assume without loss of generality that T is surjective. Then, by the Open Mapping Theorem 15.8, $T(B_1^X(0))$ is open in Y. Thus there is $\varepsilon > 0$ such that $B_\varepsilon^Y(0) \subset T(B_1^X(0))$. Hence if $\|x\| \geq 1$ then $\|Tx\| \geq \varepsilon$. The result follows by scaling. □

A particular case of the last corollary is that bijective, bounded linear maps have bounded inverses:

Corollary 15.10. *Let X and Y be Banach spaces and let $T : X \to Y$ be a bounded linear map. If T is a bijection, then T^{-1} is a bounded linear map.*

PROOF. That T^{-1} is linear follows from elementary linear algebra. The key is to prove that T^{-1} is bounded. By the previous corollary, we have $\|Tx\| \geq \varepsilon\|x\|$ for some $\varepsilon > 0$. It follows that $\left\|T^{-1}y\right\| \leq \frac{1}{\varepsilon}\|y\|$. □

The final application of the Open Mapping Theorem is the *Closed Graph Theorem*, which gives a useful criterion for a linear map to be bounded. Given Banach spaces X and Y, let $X \oplus Y$ denote the Banach space of ordered pairs (x, y) with $x \in X$ and $y \in Y$ with norm

$$\|(x,y)\|_{X\oplus Y} \ = \ \|x\|_X + \|y\|_Y \, .$$

Exercise 15.6.
- Show that $\|\cdot\|_{X\oplus Y}$ is a norm and that $X \oplus Y$ is a Banach space.
- Let $1 < p < \infty$ and define $\|(x,y)\|_{X\oplus Y,p} = (\|x\|_X^p + \|y\|_Y^p)^{1/p}$. Show that $\|\cdot\|_{X\oplus Y,p}$ is a norm and that it is equivalent to $\|\cdot\|_{X\oplus Y}$.

Given a linear map $T : X \to Y$, the *graph of T*, denoted $\Gamma(T)$, is the linear subspace of $X \oplus Y$ given by

$$\Gamma(T) \ := \ \{(x, Tx) \ : \ x \in X\} \ . \tag{15.1}$$

Exercise 15.7. Suppose that $T : X \to Y$ is a bounded linear map. Prove that $\Gamma(T)$ is a closed subspace of $X \oplus Y$.

It turns out that the graph being closed is also a sufficient condition for T to be bounded.

Theorem 15.11 (Closed Graph Theorem)**.** *Let X and Y be Banach spaces and let $T : X \to Y$ be a linear map. If the* graph of T, $X \oplus TX = \{(x, Tx) : x \in X\}$, *is a closed subspace of $X \oplus Y$ then T is bounded.*

PROOF. Consider the coordinate maps $\pi_1 : X \oplus Y \to X$ and $\pi_2 : X \oplus Y \to Y$ given by

$$\pi_1(x, y) = x \quad \text{and} \quad \pi_2(x, y) = y ,$$

which both have norm bounded by one. Let S denote the restriction of π_1 to $\Gamma(T)$. Then $S : \Gamma(T) \to X$ is a bounded linear bijection. Since $\Gamma(T)$ is closed, it is a Banach space and thus the inverse $S^{-1} : X \to \Gamma(T)$ is bounded. Thus $T = \pi_2 \circ S^{-1}$ is bounded. □

4. Algebraic dimension of Banach spaces

As a final application of the Baire category theorem, we prove that any infinite dimensional Banach space has uncountable algebraic dimension:

Theorem 15.12. *Let X be a Banach space. If X is not finite dimensional, then X has uncountable algebraic dimension.*

PROOF. Suppose that there is a countable set $\{x_j\}_{j=1}^{\infty}$ in X that spans X. We must show that X is finite dimensional. Let

$$X_n = \left\{ \sum_{j=1}^{N} \alpha_j x_j \ : \ \alpha_1, \ldots, \alpha_N \in F \right\} .$$

By assumption $X = \bigcup_N X_N$. Furthermore, each set X_N is a closed subspace:

Exercise 15.8. Let X be a topological vector space and let $Y \subset X$ be a finite dimensional subspace. Prove that Y is closed.

Thus there is m such that X_m has non-empty interior. In particular, there is $x_0 \in X_m$ such that $B_r(x_0) \subset X_m$ for some $r > 0$. But then $B_r(0) = B_r(x_0) - x_0 \subset X_m$. By scaling it follows that $X_m = X$. □

LECTURE 16

Weak and weak* topologies

1. Weak convergence in Banach spaces

Consider a Banach space X and its Banach space dual $X^\star$. The weakest topology on X so that every element of $X^\star$ is continuous is called the *weak topology* on X. This topology was already introduced above in the general context of LCSs, where it was denoted above by $\sigma(X, X^\star)$. A net $(x_n)_{n\in I}$ in X *converges weakly* to x if it converges in the weak topology, i.e., if

$$\ell(x_n) \;\to\; \ell(x) \quad \text{for every } \ell \in X^\star.$$

This is denoted

$$x_n \rightharpoonup x \quad \text{or} \quad \operatorname*{wk\text{-}lim}_{n\to\infty} x_n = x\,.$$

Strong convergence refers to convergence in the norm topology: $\|x_n - x\| \to 0$, denoted $x_n \to x$. Weak convergence is in fact *weaker* than strong convergence. That is strong convergence implies weak convergence, but not conversely.

Exercise 16.1. Let X be a Banach space and let $(x_n)_{n=1}^\infty$ be a sequence in X that converges strongly to $x \in X$. Prove that $\text{wk-lim}\, x_n = x$.

Proposition 16.1. *Let $(x_n)_{n=1}^\infty$ be an orthonormal sequence in a Hilbert space H. Then $x_n \rightharpoonup 0$.*

Remark 16.2. Since $\|x_n\| = 1$, x_n does not converge strongly to 0.

PROOF. Fix $y \in H$. By Bessel's inequality 6.10

$$\sum_n |\langle y, x_n\rangle|^2 \;\le\; \|y\|^2,$$

we see that $\langle y, x_n\rangle \to 0$. By the Riesz-Fréchet Theorem 5.10 it follows that $x_n \rightharpoonup 0$. ☐

The following theorem is very useful in practice for establishing weak convergence, by focusing on proving convergence for a dense subset of linear functionals.

Theorem 16.3. *Suppose $(x_n)_{n=1}^\infty \in X$, 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' norm-dense in X'.

Then $x_n \rightharpoonup x$.

PROOF. Let $M = \max(\|x\|, \sup_n \|x_n\|)$. Let $\ell \in X'$ and $\epsilon > 0$. Then there is $\ell' \in Y'$ such that $\|\ell - \ell'\| < \frac{\epsilon}{3M}$. We have

$$\begin{aligned}|\ell(x_n - x)| \;&\le\; |\ell(x_n) - \ell'(x_n)| + |\ell'(x_n) - \ell'(x)| \;+\; |\ell'(x) - \ell(x)| \\ &\le\; \|\ell - \ell'\|(\|x_n\| + \|x\|) + |\ell'(x_n - x)| \;<\; \frac{2}{3}\epsilon + |\ell'(x_n - x)|.\end{aligned}$$

Choosing N large enough (depending on ℓ') we may obtain $|\ell'(x_n - x)| < \frac{\epsilon}{3}$ and thus $|\ell(x_n - x)| < \epsilon$ for $n \geq N$. □

The Principle of Uniform Boundednes (Thm. 15.5) provides a converse to the previous result: *weakly convergent sequences are bounded*. We will obtain this from a more general result about *weakly precompact sets*. We say that a set S in a topological space is *pre-compact* if $\overline{S}$ is compact. In partiuclar, the image of any convergent sequence is precompact.

Theorem 16.4. *Let X be a Banach space and let $S \subset X$ be a weakly pre-compact subset of X. Then S is bounded, i.e., 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.

PROOF. Since S is weakly pre-compact, for a given linear functional ℓ the set $\{\ell(x) : x \in X\}$ must be a bounded subset of F (otherwise we could find a weakly divergent sequence). By the Principle of Uniform Boundedness applied to $\mathcal{L} = X$, considered as a set of linear functionals on $X^\star$, there is a constant c such that $\|x\| \leq c$ for all $x \in S$. Here we have used Thm. 7.7 which states that the norm x as a linear functional on $X^\star$ is the same as the Banach space norm of x. □

The Banach space norm is, in general, not a weakly continuous function. For example, an orthonormal sequence in a Hilbert space converges to zero weakly, although each element of the sequence has norm one. It is, however, "lower semi-continuous." 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

$$f(x) \leq \liminf_n f(x_n) \tag{16.1}$$

for any convergent sequence $x_n \to x$. That is a lower semi-continuous function can "jump down" in a limit, but cannot "jump up." (To see (16.1), 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 16.5 (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 16.6. 1) This should remind you of Fatou's lemma. 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. □

2. The weak* topology

We may also consider a weak topology on the dual $X^\star$ of a Banach space. That is the *weak** topology $\sigma(X^\star, X)$. 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

$$\operatorname*{wk^\star\text{-}lim}_{n\to\infty} u_n = u.$$

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 requires all linear functionals in $X^{\star\star}$ to be continuous.

Theorem 16.7. *A weak⋆ convergent sequence u_n is uniformly bounded and*

$$\|u\| \le \liminf \|u_n\|,$$

if $u = \text{wk}^\star\text{-lim}\, u_n$.

PROOF. Exercise. □

A key advantage of the weak⋆ topology is that closed balls in $X^\star$ are *compact* in this topology, a result known as *Alaoglu's Theorem* or somtimes the *Banach-Alaoglu Theorem*. For the proof of this result we will need:

Theorem 16.8 (Tychonoff's Theorem)**.** *Let (K_α), $\alpha \in I$, be a collection of compact spaces. Then the product $T = \prod_{\alpha\in I} K_\alpha$ is compact in the* product topology.

Remarks. 1) Recall that the *Cartesian product* $T = \prod_{\alpha\in I} K_\alpha$ is the set of all maps $x : I \to \bigcup_\alpha K_\alpha$ such that $x(\alpha) \in K_\alpha$ for every $\alpha \in I$.] 2) The *product topology* on T is the weakest topology such that the coordinate maps $\Pi_\alpha(x) = x(\alpha)$ are continuous. That is, a neighborhood base for T is given by the collection of sets of the form

$$\{x \,:\, x(\alpha_j) \in U_j,\ j = 1, \dots, N\}\ ,$$

with U_j open in K_{α_j} and $\{\alpha_1, \dots, \alpha_N\} \subset I$ an arbitrary finite collection. 3) The key point here is that the collection I can be arbitrarily large. 4) The theorem, with $K_\alpha = [0,1]$ for each α, is due to Tychonoff 1930. A proof of the full version was first given by Cech 1937. 5) The proof of this theorem is beyond the scope of these notes. It can be found in standard references on point set topology, e.g., Munkres 1974, Ch. 5, Theorem 1.1. For a short proof using nets see Chernoff 1992.

Theorem 16.9 (Alaoglu 1940)**.** *Let $X^\star$ be the dual of a Banach space X. The unit ball of $X^\star$ is weak⋆ compact.*

PROOF. Let B be the unit ball in $X^\star$. Let T be the (uncountable) product space:

$$T = \prod_{x\in X} I_x, \quad I_x = [-\|x\|, \|x\|].$$

By Tychonoff's theorem 16.8, 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 $\ell \in B$, $|\ell(x)| \le \|\ell\|\|x\| \le \|x\|$ 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 weak⋆ 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 $\Phi^{-1}_{x,y;t}(\{0\})$ is a closed set. Thus

$$\overline{B} \subset \bigcap_{x,y,t} \Phi^{-1}_{x,y;t}(\{0\}),$$

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 $\overline{B} = B$. $\square$

The weak* compactness of the unit ball in $X^\star$ has important consequences, for example in proving the existence of minimizers of functionals and thus for solving partial differential equations with a variational principle. Often one simply needs the existence of convergent subsequences for a given sequence, which follows from a weaker property.

Definition 16.10. A subset C of a dual Banach space $X^\star$ is *weak* sequentially compact* if every sequence of points in C has a weak* convergent subsequence with weak* limit in C.

Sequential compactness is, in general, a strictly weaker notion than compactness, although the notions are equivalent in metric spaces. In the present context, X is metrizable in the weak-topology if and only if $X^\star$ is separable (see Thm. 12.1).

Theorem 16.11 (Helly 1912)**.** *Let X be a separable Banach space. Then the closed unit ball in $X^\star$ is weak* sequentially compact.*

Clearly Helly's Theorem follows from Alaoglu's theorem 16.9. However, the proof we will now give of Helly's theorem is much more useful. The proof of Alaoglu's theorem is non-constructive, as it relies on Tychonoff's theorem. Often times what one really wants is to find a convergent sequence. The following proof Helly's theorem gives you an idea how to construct it.

PROOF. Given $u_n \in X^\star$ 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. $\square$

Every reflexive space is, of course, a dual space. So Alaoglu's Theorem 16.9 implies the forward direction of the following

Theorem 16.12. *A Banach space is reflexive if and only if its unit ball is weakly compact.*

Remark. The "only if" direction is due to Eberlein 1947 and Smulian 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 $g \equiv 1$,

$$\int_0^\infty f_n(x)g(x) = 1.$$

Reading: §11.6 in Lax

3. Positive harmonic functions

Let $\Omega \subset \mathbb{C}$ be an open set. Recall that a function $u : \Omega \to \mathbb{R}$ is called *harmonic* if it satisfies the *mean value property*

$$u(z) = \frac{1}{\pi\varepsilon^2}\int_{\mathbb{D}_\varepsilon(z)} u(w)\mathrm{d}m(w), \tag{16.2}$$

whenever the disc $\mathbb{D}_\varepsilon(z)$ of radius ε centered at z is contained in Ω, where m is Lebesgue measure. We may apply weak* compactness to prove the following:

Theorem 16.13. *Let u be a harmonic function on the open unit disk $\mathbb{D} = \{|z| < 1\}$. If $u(z) \geq 0$ for all $z \in \mathbb{D}$, then here is a unique finite, non-negative, Borel measure μ on $\partial\mathbb{D} = \{|z| = 1\}$ such that*

$$u(z) \;=\; \int_{\partial D} \frac{1-|z|^2}{|z-w|^2}\mathrm{d}\mu(w). \tag{16.3}$$

Conversely, any such function is a non-negative Harmonic function on the disk.

Remark. The theorem implies $|u(z)| \leq \text{const.}/(1-|z|)$. Thus the singularities of a non-negative harmonic function are highly constrained as z approaches the boundary. No similar estimate holds for a general real valued harmonic function. For example $u(z) = \operatorname{Re}\exp(1/1-z)$ blows up very fast as $z \to 1$ from within the disk.

PROOF. Note that u is continuous. To see this, observe that by the mean value property

$$u(z+h) - u(z) = \frac{1}{\pi\varepsilon^2}\left[\int_{\mathbb{D}_\varepsilon(z+h)} u(w)\mathrm{d}m(w) - \int_{\mathbb{D}_\varepsilon(z)} u(w)\mathrm{d}m(w)\right].$$

By dominated convergence the difference of integrals on the r.h.s. converges to zero as h converges to zero. Since u is continuous, we may differentiate

$$\pi r^2 u(0) = \int_{\mathbb{D}_r(0)} u(z)\mathrm{d}m(z) \;=\; \int_0^r \int_0^{2\pi} u(s\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\theta s\mathrm{d}s$$

with respect r and conclude that for every r

$$u(0) = \frac{1}{2\pi}\int_0^{2\pi} u(re^{i\theta})d\theta. \tag{16.4}$$

For each $r \in (0,1)$, define a measure on the circle $\partial\mathbb{D}$ by $d\mu_r(\theta) = (2\pi)^{-1}u(re^{i\theta})d\theta$. Since u is non-negative we have $\|\mu\| = \mu(\partial\mathbb{D}) = u(0)$ by (16.4). Thus the family $(\mu_r)_{r\in(0,1)}$ lies in the ball of radius $u(0)$ centered at 0 in $\mathcal{M}(\partial D) = C(\partial\mathbb{D})^\star$. By Helly's theorem we may find a weak* convergent subsequence μ_{r_n}. That is, there is a Borel measure μ on $\partial\mathbb{D}$ such that

$$\int_{\partial\mathbb{D}} f(\theta)d\mu(\theta) = \lim_{n\to\infty}\frac{1}{2\pi}\int_{\partial\mathbb{D}} f(\theta)u(r_n e^{i\theta})d\theta \tag{16.5}$$

for every $f \in C(\partial\mathbb{D})$. We see from (16.5) that μ is a positive measure and, by taking $f \equiv 1$, that $\mu(\partial\mathbb{D}) = u(0)$.

To complete the proof, we will use the identity

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi}\frac{r^2-|z|^2}{|re^{i\theta}-z|^2}u(re^{i\theta})d\theta, \tag{16.6}$$

valid for $z \in \mathbb{D}_r(0)$. Let us defer the proof for the moment and show how (16.6) implies the representation (16.3). Fix z and let

$$f_{r,z}(e^{i\theta}) = \frac{r^2-|z|^2}{|re^{i\theta}-z|^2}.$$

Note that $f_{r,z} \to \frac{1-|z|^2}{|e^{i\theta}-z|^2}$ uniformly as $r \to 1$. Thus the weak* convergence $\mu_{r_n} \to \int \cdot\mu$ implies

$$u(z) = \lim_n \int_{\partial\mathbb{D}} f_{r_n;z}(e^{i\theta})d\mu_{r_n}(\theta) = \int_0^{2\pi}\frac{1-|z|^2}{|e^{i\theta}-z|^2}d\mu(\theta).$$

Exercise 16.2. Let $(x_n)_{n=1}^\infty$ be a sequence in a Banach space X and let $(\ell_n)_{n=1}^\infty$ be a sequence in the dual $X^\star$. If $\ell_n \rightharpoonup \ell$ and $x_n \to x$, prove that $\ell_n(x_n) \to \ell_\infty(x)$.

The identity (16.6) 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 right hand side. It is straightforward to show that v is harmonic in $\mathbb{D}_r(0)$ — for this it suffices to show that $\frac{|w|^2-|z|^2}{|w-z|^2}$ is harmonic in $\mathbb{D}_{|w|}(0)$ for fixed w and use Fubini's theorem. Furthermore, it is not too hard to show that

$$\lim_{s\uparrow r} v(re^{i\theta}) = u(re^{i\theta}),$$

since for any continuous function f on the circle

$$\frac{1}{2\pi}\lim_{s\uparrow r}\int_0^{2\pi}\frac{r^2-s^2}{|re^{i\theta}-se^{i\phi}|^2}f(e^{i\phi})d\phi = f(e^{i\theta}). \tag{16.7}$$

Exercise 16.3. Prove (16.7).

Thus $u(z) - v(z)$ is a harmonic function on $\mathbb{D}_r(0)$, 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. Recall that the maximum principle is a straightforward consequence of the mean value property and continuity. $\square$

4. Herglotz-Riesz Theorem

An important application of the above is the following:

Theorem 16.14 (Herglotz-Riesz)**.** *Let F be an analytic function in the unit disk $\mathbb{D} = \{|z| < 1\}$ such that $\operatorname{Re} F \geq 0$ in $\mathbb{D}$. Then there is a unique non-negative, finite, Borel measure μ on $T^1 = \partial\mathbb{D}$ such that*

$$F(z) \;=\; \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z}\mu(d\theta) + i \operatorname{Im} F(0).$$

Conversely every analytic function in the disk with positive real part can be written in this form.

Remark. This theorem is due to Herglotz 1911 and Riesz 1911.

PROOF. First apply the Thm. 16.13 to $\operatorname{Re} F$. Let

$$G(z) = \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z}\mu(d\theta).$$

Then 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 $F(z) - G(z) = i \operatorname{Im} F(0)$. ☐

The theorem is often used in the following form

Theorem 16.15. *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*

$$\int_{\mathbb{R}} \frac{1}{1+x^2} d\mu(x) < \infty$$

and

$$F(z) = Az + \operatorname{Re} F(i) + \int_{\mathbb{R}} \frac{1+xz}{x-z}\frac{1}{1+x^2} d\mu(x). \qquad (\star\star\star)$$

Furthermore

$$A = \lim_{z\to\infty} \frac{F(z)}{z},$$

and

$$d\mu(x) = \text{wk}^\star\text{-}\lim_{y\downarrow 0} \frac{1}{\pi} \operatorname{Im} F(x+iy)dx.$$

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

$$F(z) = Az + B + \int_{\mathbb{R}} \frac{1}{x-z} d\mu(x).$$

Remark 16.16. Note that

$$\frac{1+xz}{x-z}\frac{1}{1+x^2} = \frac{1}{x-z} - \operatorname{Re}\frac{1}{x-i}.$$

PROOF. Consider the function

$$G(\zeta) = -iF\left(i\frac{1-\zeta}{\zeta+1}\right).$$

This is an analytic map from the disk into the right half plane. By the Herglotz-Riesz theorem

$$F\left(i\frac{1-\zeta}{\zeta+1}\right) \;=\; i\int_0^{2\pi} \frac{e^{i\theta} + \zeta}{e^{i\theta} - \zeta} d\nu(\theta) + \operatorname{Re} F(i)$$

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

$$\lim_{z\to\infty}(F(z) - Az) = \operatorname{Re} F(\mathrm{i}) - \operatorname{Re}\int_{\mathbb{R}} \frac{1}{x-\mathrm{i}}\mathrm{d}\mu(x). \quad \square$$

Part 5

Convexity

LECTURE 17

Convex sets in a Banach space

Reading: §8.4 and Ch. 12 of Lax

Definition 17.1. 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, $S_M(\ell + m) \leq S_M(\ell) + S_M(m)$.
(2) $S_M(0) = 0$.
(3) Positive homogeneity, $S_M(a\ell) = aS_M(\ell)$ for $a > 0$.
(4) Monotonicity: for $M \subset N$, $S_M(\ell) \leq S_N(\ell)$.
(5) Additivity: $S_{M+N} = S_M + S_N$, where $M + N = \{x + y | x \in M \text{ and } y \in N\}$.
(6) $S_{-M}(\ell) = S_M(-\ell)$, where $-M = \{-x | x \in M\}$.
(7) $S_{\overline{M}} = S_M$.

Exercise 17.1. Prove properties 1-7.

Definition 17.2. Let M be a subset of a Banach space. The *convex hull* $\hat{M}$ of M is the smallest convex set containing M. The *closed convex hull* $\check{M}$ of M is the smallest closed convex set containing M.

Exercise 17.2. Prove that $\check{M}$ is equal to the closure $\hat{M}$.

In addition, to properties 1-8 above, we have

(8) $S_{\check{M}} = S_M$.

Let us prove property (8), assuming the other properties. First by (7) and exercise **??** it suffices to show $S_{\hat{M}} = S_M$. Since $M \subset \hat{M}$, by (5) we have $S_M \leq S_{\hat{M}}$. However,

$$\hat{M} = \left\{ \sum_{j=1}^{n} a_j x_j \ : \ x_j \in M, \ a_j > 0, \ \text{and} \ \sum_j a_j = 1 \right\}.$$

So for any point in $\hat{M}$ we have

$$\ell\left(\sum_{j=1}^{n} a_j x_j\right) = \sum_{j=1}^{n} a_j \ell(x_j) \leq S_M(\ell).$$

Thus $S_{\hat{M}} \leq S_M$.

Here are some examples:

- If $M = \{x_0\}$, S_M is just evaluation at x_0.
- If $M = B_R(0)$ then $S_M(\ell) = R\|\ell\|$.
- If $M = B_R(x_0)$ then $M = \{x_0\} + B_R(0)$ so $S_M(\ell) = R\|\ell\| + \ell(x_0)$.

- 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, $S_M : X^\star \to \mathbb{R}$, however in general we define S_M as a map from $X^\star \to \mathbb{R} \cup \{\infty\}$. As in measure theory, 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

$$\left|S_M(\ell) - S_M(\ell')\right| \le \max\{S_M(\ell - \ell'), S_M(\ell' - \ell)\} \le \text{const.}\left\|\ell - \ell'\right\|.$$

This fails if M is unbounded, and also in the weak* topology. Nonetheless, we have the following additional property

(9) S_M is *weak* lower semi-continuous* .

Here we extend the definition of lower semi-continuity to functions taking values in $\mathbb{R} \cup \{\infty\}$: a function f from a topological space Ω to $\mathbb{R} \cup \{\infty\}$ is lower semi-continuous if $\{x \in \Omega : f(x) > t\}$ is open for every $t \in \mathbb{R}$. Since S_M is defined to be a surpremum of the weak* continuous functions $\ell \mapsto \ell(x)$, property (9) follows from the fact that lower semi-continuity is preserved by taking supremums:

Proposition 17.3. *Let Ω be a topological space and let $\{f_\alpha : \Omega \to \mathbb{R} \cup \{\infty\} : \alpha \in I\}$ be an arbitrary collection of lower semi-continuous functions. Then $F(x) := \sup_x f_\alpha(x)$ is lower semi-continuous.*

PROOF. Note that

$$\{x : F(x) > t\} = \bigcup_{\alpha \in I} \{x : f_\alpha(x) > t\} . \quad \square$$

Theorem 17.4. *The closed convex hull $\check{M}$ of a subset M of a Banach space X over $\mathbb{R}$ is equal to*

$$\check{M} = \{x : \ell(x) \le S_M(\ell) \text{ for all } \ell\}.$$

PROOF. Since $S_M = S_{\check{M}}$ it follows that $\ell(x) \le S_M(\ell)$ for all $x \in \check{M}$.

Now, suppose $x \notin \check{M}$. Since $\check{M}$ is closed there is an open ball $B_R(x)$ disjoint from $\check{M}$. By the geometric Hahn-Banach theorem 2.15 we can find a linear functional ℓ_0 and $c \in \mathbb{R}$ such that

$$\ell_0(u) \le c < \ell_0(v) \quad \text{for all } u \in \check{M} \text{ and } v \in B_R(x).$$

In particular, if $\|y\| \le R$ then

$$\ell_0(-y) = -\ell_0(y + x) + \ell_0(x) < \ell_0(x) - c .$$

Thus ℓ_0 is bouned and $\|\ell_0\| \le R^{-1}(\ell_0(x) - c)$. Furthermore, we have

$$\ell_0(x) \ge c - \ell_0(y) \ge c + R\|\ell_0\|.$$

It follows that

$$\ell_0(z) \ge S_M(\ell_0) + R\|\ell_0\|.$$

Thus ℓ_0 is a linear functional such that $\ell_0(x) > S_M(\ell_0)$. $\square$

The theorem shows that a closed, convex set K in a Banach space may be specified as the set

$$K = \{z : \Phi_K(x) \le 0\}$$

where

$$\Phi_K(z) = \sup_{\ell:\|\ell\|\leq 1} [\ell(x) - S_K(\ell)].$$

Since $S_K : X^\star \to \mathbb{R} \cup \{\infty\}$, the function Φ_K is initially defined as a map $X \to \mathbb{R} \cup \{-\infty\}$. However, note that $\Phi(x) = -\infty$ for some x if and only if $S_K(\ell) = \infty$ for all ℓ, in which case $\Phi(x) = -\infty$ for all x 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 x, $\ell(x) - S_K(\ell)$ is *weak* upper semi-continuous*, that is for each t

$$\{\ell \ : \ \ell(x) - S_K(\ell) < t\}$$

is weak* open. This observation is relevant, since $\{\|\ell\| \leq 1\}$ is compact, and

Proposition 17.5. *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.*

PROOF. The sets $(\{F(x) < t\})_{t\in\mathbb{R}}$ 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 $t_m = F(x_m)$. ☐

It follows that,

$$\Phi_K(x) = \max_{\|\ell\|\leq 1} [\ell(x) - S_K(\ell)].$$

The function $\Phi_K(x)$, being a max of weakly continuous functions, is in turn weakly lower semi-continuous. In particular,

$$K = \{x \ : \ \Phi_K(x) \leq 0\}$$

is *weakly closed*! Thus we have obtained the following theorem:

Theorem 17.6 (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 $\implies$ strongly closed for any set.)

Exercise 17.3. Find a Banach space such that the complement of an open ball $\{x \ : \ \|x\| \geq 1\}$ is not weakly closed.

As a corollary we have

Corollary 17.7. *If X is reflexive, then a bounded, norm closed, convex subset K in X is weakly compact.*

Exercise 17.4. 1) Show that the set K of non-negative functions $\rho \in L^1(\mathbb{R})$ such that $\int \rho dx = 1$ is convex, norm closed and bounded, but not weakly compact. Thus Cor. 17.7 fails if the requirement that X be reflexive is dropped. 2) Show that the space of Baire probability measures on $\mathbb{R}$ is a norm closed, bounded and convex set of the space $\mathcal{M}_0(\mathbb{R})$ of finite Baire measures on $\mathbb{R}$, but that it is not weak* closed ($\mathcal{M}_0(\mathbb{R})$ is the dual to the space $C_0(\mathbb{R})$).

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 17.8. *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 $\ell(x) - S_K(\ell) \leq 0$ for all ℓ so the maximum is attained at $\ell = 0$ and $\Phi_K(x) = 0$.

If $x \notin K$, $u \in K$ and $\|\ell\| \leq 1$, then

$$S_K(\ell) \geq \ell(u) = \ell(u - x) + \ell(x) \geq \ell(x) - \|u - x\|.$$

Thus

$$\|u - x\| \geq \sup_{\|\ell\| \leq 1} [\ell(x) - S_K(\ell)].$$

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 Thm. 17.4, there is $\ell_0 \in X'$ such that

$$S_K(\ell_0) + R\|\ell_0\| = S_{K+B_R(0)}(\ell_0) < \ell_0(x).$$

By scaling, we may take $\|\ell_0\| = 1$ to conclude

$$R < \ell_0(x) - S_K(\ell_0) \leq \sup_{\|\ell\| \leq 1} [\ell(x) - S_K(\ell)].$$

As R was any number less than $\inf_{u \in K} \|u - x\|$ the reverse inequality follows. □

LECTURE 18

Convex sets continued; Krein Millman Theorem

Reading: §13.3

1. Duality between convex sets and homogeneous subadditive functions

In the last lecture we saw that a closed convex set K in a Banach space satisfied the equality

$$K = \{x \in X : \ell(x) \le S_K(\ell) \text{ for all } \ell \in X^\star\},$$

where $S_K(\ell) = \sup_{x\in K} \ell(x)$ is the support function of K. In this lecture, we turn things aroung and start with a function $S : X^\star \to \mathbb{R} \cup \{\infty\}$. Consider the set

$$K = \{z \in X : \ell(z) \le S(\ell) \text{ for all } \ell \in X^\star\}.$$

Then K is clearly convex and weakly closed, and its support function satisfies

$$S_K(\ell) = \sup_{x\in K} \ell(x) \le S(\ell). \tag{18.1}$$

Can strict inequality hold in (18.1)? Of course it can, e.g., if the function S is not sub-additive or homogeneous. However, if we assume that S is 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 18.1. *Let X be a reflexive Banach space and let $S : X^\star \to \mathbb{R} \cup \{\infty\}$ 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 $K = \{x : \ell(x) \le S(\ell) \text{ for all } \ell\}$.

PROOF. To begin, let us prove the theorem under the additional restriction that S is bounded: $|S(\ell)| \le \beta\|\ell\|$. Then $x \in K \implies \|x\| \le \beta$, so K is bounded. Since $S(0) = 0$ the identity holds for $\ell = 0$.

Now fix a non-zero linear functional ℓ_0. Let us define a linear functional $L \in X^{\star\star}$, the double dual, via Hahn-Banach. Begin on the one-dimensional subspace spanned by ℓ_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 $S(-\ell_0) \ge -S(\ell_0)$ by sub-additivity.) Thus by Hahn-Banach there is a linear functional L on $X^\star$ which satisfies $L(\ell) \le S(\ell)$ for every $\ell \in X^\star$. Since $S(\ell) \le \beta\|\ell\|$ this functional is bounded.

As X is reflexive there is $x \in X$ such that $L(\ell) = \ell(x)$. Since $\ell(x) \le S(\ell)$ for all ℓ, we have $x \in K$, and since

$$S(\ell_0) = L(\ell_0) = \ell_0(x),$$

we have

$$S(\ell_0) = \max_{y \in K} \ell_0(y).$$

As ℓ_0 was arbitrary, this completes the proof for S bounded.

To extend this to unbounded S, note that

$$-\infty < \inf_{\|\ell\| \leq 1} S(\ell) \leq 0,$$

by wk* lower semi-continuity. Let $-\beta = \inf_{\|\ell\| \leq 1} S(\ell)$. Then by positive homogeneity,

$$S(\ell) \geq -\beta \|\ell\|,$$

that is S is bounded from below.

Now, for each ϵ define

$$K_\epsilon = \{x \ : \ \ell(x) \leq S_\epsilon(\ell) \text{ for all } \ell\},$$

where

$$S_\epsilon(\ell) = \inf_{\ell_1, \ell_2 \in X^\star \, : \, \ell_1 + \ell_2 = \ell} \left[S(\ell_1) + \frac{1}{\epsilon} \|\ell_2\| \right].$$

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

$$-\beta \|\ell\| \leq S_\epsilon(\ell) \leq \frac{1}{\epsilon} \|\ell\|, \quad \text{and } S_\epsilon(\ell) \leq S(\ell),$$

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

$$\ell(x) = \ell_1(x) + \ell_2(x) \leq S(\ell_1) + \frac{1}{\epsilon} \|\ell_2\|$$

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 $\frac{1}{\epsilon}\|\ell\|$ for all ℓ so $x \in K \cap B_{1/\epsilon}(0)$. It follows that $K = \bigcup_\epsilon K_\epsilon$, 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

$$S_\epsilon(\ell) \leq S(\ell - \ell_\epsilon) + \frac{1}{\epsilon} \|\ell_\epsilon\| \leq S_\epsilon(\ell) + \epsilon.$$

Since $S(\ell - \ell_\epsilon) \geq -\beta\|\ell\| - \beta\|\ell_\epsilon\|$ we see that

$$-\beta\|\ell\| + \left(\frac{1}{\epsilon} - \beta \right) \|\ell_\epsilon\| \leq S_\epsilon(\ell) + \epsilon.$$

Consider the following cases: (1) $\|\ell_\epsilon\|/\epsilon$ is bounded as $\epsilon \to 0$ or (2) $\|\ell_\epsilon\|/\epsilon$ 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

$$S(\ell) \le \liminf_{\epsilon\to 0} S(\ell - \ell_\epsilon) \le S_0(\ell) - \liminf_{\epsilon\to 0} \frac{1}{\epsilon}\|\ell_\epsilon\| \le S_0(\ell),$$

which completes the proof. □

2. Krein-Milman Theorem

Definition 18.2. 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 18.3 (Carathéodory 1911)**.** *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 18.4. The proof is left as an exercise. Use induction on N. The case $N = 1$ is easy!

Theorem 18.5 (Krein and Milman 1940)**.** *Let X be a locally convex space. If K is 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.

To prove the Krein-Milman theorem we will use a characterization of compactness in terms of intersections of closed sets. A family $\mathcal{F}$ of closed sets in a topological space 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: $F_1 \cap \cdots \cap F_n \neq \emptyset$.

Lemma 18.6. *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 $\mathcal{U} = \{F^c : F \in \mathcal{F}\}$. Thus $M = \cup_{j=1}^n F_j^c$ for some finite collection. Then $\cap_{j=1}^n F_j = \emptyset$, so $\mathcal{F}$ 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 $\mathcal{U} = \{F^c : F \in \mathcal{F}\}$ is an open cover of M with no finite subcover so M is not compact. □

PROOF OF THEOREM 18.5. 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 $\alpha \ge \beta$. Clearly $\cap\mathcal{T}$ is a candidate for an "upper bound." To see that it is, we must show that $\cap\mathcal{T} \in \mathcal{E}$, i.e. that it is nonempty, closed, and extreme.

Clearly $\cap\mathcal{T}$ is closed. Furthermore, the intersection of a family of extreme sets is easily seen to be extreme, provided it is non-empty. Thus it suffices to show that $\mathcal{T}$ is non-empty.

Here we use compactness of K in a crucial way. Note that the collection $\mathcal{T}$ has the FIP — since it is totally ordered, any finite collection can be ordered with $E_1 \supset E_2 \cdots \supset E_n$ so the intersection is just E_n. By Lemma 18.6, we see that $\cap\mathcal{T} \neq \emptyset$.

By Zorn's Lemma $\mathcal{E}$ has a "maximal element," namely a set $E \in \mathcal{E}$ with no proper subset contained in $\mathcal{E}$. We claim that any such "maximal" element is a one point set. Indeed, suppose $E \subset \mathcal{E}$ contains two distinct points — as E is convex 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 at which ℓ attains its maximum on E:

$$M = \{x \in E \;:\; \ell(x) = \max_{z\in E} \ell(z)\}.$$

Then M is a non-empty, proper, closed subset of E. It is clearly convex and is easily seen to be an extreme subset of E. It follows that M is an extreme subset of K and $M \subsetneq E$ so E is not minimal.

Exercise 18.1. Let K be a convex set and let E be an extreme subset of K. If M is an extreme subset of E show that M is an extreme subset of K.

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. By the above exercise, an extreme point of E is also an extreme point of K. 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 $\check{K}_e$ its closed convex hull. As in a Banach space one has

$$\check{K}_e \;=\; \text{closure of } \widehat{K}_e \;=\; \text{smallest closed convex set containing } K_e\,.$$

Since K is closed and convex and contains K_e, we have $\check{K}_e \subset K$. On the other hand if $z \notin \check{K}_e$ then there is an open set U with $z \in U \subset K_e^c$. Since X is a LCS, 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

$$\ell(x) < c \leq \ell(y) \quad \text{for all } x \in U \text{ and } y \in \check{K}_e.$$

(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. If $\frac{x}{t} \in U - z$ we have

$$\tfrac{1}{t}\ell(x) = \ell(\tfrac{x}{t} + z) - \ell(z) < c - \ell(z).$$

Thus $\ell(x) < t(c - \ell(z))$ and so

$$\ell(x) \leq (c - \ell(z))p_{U-z}(x).$$

Therefore ℓ is a continuous linear functional on X.

Since K is compact ℓ achieves its minimum on K. Let E be the set of minimizers. Then E is closed, convex and extreme. 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 $z \notin K$. □

LECTURE 19

The Stone Weierstrass Theorem and Choquet's Theorem

Reading: §13.4 and 14.10 in Lax.

1. The Stone Weierstrass Theorem

An interesting application of the Krein Milman Theorem 18.5 is the *Stone-Weiertstrass Theorem*, which is a vast generalization of the classical result of Weierstrass 1885 on approximation of continuous functions on an interval with polynomials.

Definition 19.1. Let S be a compact Hausdorff space and $C_{\mathbb{R}}(S)$ the set of real valued continuous functions on S. Let $E \subset C(S)$.

(1) E is a *sub-algebra* if E is a linear subspace and $fg \in E$ whenever f and g are elements of E. (Here fg denotes the function $fg(x) = f(x)g(x)$).
(2) E is said to *separate points* if for any pair $p, q \in S$ with $p \neq q$ there is $f \in E$ such that $f(p) \neq f(q)$.
(3) E is said to be *nowhere vanishing* if for any $p \in S$ there is $f \in E$ such that $f(p) \neq 0$.

Theorem 19.2 (Stone-Wierstrass)**.** *Let S be a compact Hausdorff and let $E \subset C_{\mathbb{R}}(S)$ be a sub-algebra. If E separates points and is nowhere vanishing, then E is dense in $C(S)$ in the max norm.*

Remark. This theorem is due to Stone 1937.

In the proof we will use the following

Exercise 19.1. Let S be a compact Hausdorff space and $E \subset C_{\mathbb{R}}(S)$. Consider the collection $\mathcal{N}$ of all finite Baire measures μ on S such that $\int f \mathrm{d}\mu = 0$ for all $f \in E$. Show that E is dense if and only if $\mathcal{N} = \{0\}$.

PROOF. Let $\overline{E}$ denote the closure of E in the max norm. We must show that $\overline{E} = C_{\mathbb{R}}(S)$. Note that $\overline{E}$ is, itself, a subalgebra that is nowhere vanishing and separates points.
Claim: *If $f \in \overline{E}$, then $|f| \in \overline{E}$.*

To see this, let $M = \max_x |f(x)|$ and let $\varepsilon > 0$. By the classical Weierstrass theorem (see Rudin 1976, Theorem 7.26), there is a polynomial P such that

$$\max_{x \in [-M,M]} \left| |x| - P(x) \right| < \epsilon .$$

Since $\overline{E}$ is an algebra, $P \circ f \in \overline{E}$ and we have $\|P \circ f - |f|\| < \epsilon$. As ϵ is arbitrary and $\overline{E}$ is closed, we have $|f| \in \overline{E}$.
Claim: *If $f, g \in \overline{E}$ then* $\max(f,g) \in \overline{E}$*.*

This follows from the previous claim since for any two real numbers a, b, we have $\max(a,b) = \frac{1}{2}(|a-b| + a + b)$.
Claim: *There is a function $G \in \overline{E}$ such that $G(x) > 0$ for all $x \in S$.*

To see this, note that for each $x \in S$ there is $g_x \in S$ with $g_x(x) \neq 0$. Multiplying by -1 if necessary, we may choose g_x so that $g_x(x) > 0$. Let $U_x = \{y : g_x(y) > 0\}$. Then $(U_x)_{x \in S}$ is an open cover of S. Since S is compact there is a finite subcover $x_1, \ldots, x_n$. Let $\widetilde{g}_j = \max(g_{x_j}, 0)$ and let $G = \sum_j \widetilde{g}_j$.

Now let $\mathcal{N}$ denote the set of all finite Baire measures μ on S such that $\int f \mathrm{d}\mu = 0$ for all $f \in \overline{E}$. We must show that $\mathcal{N} = \{0\}$. By construction $\mathcal{N}$ is weak* closed and convex. So $K = \mathcal{N} \cap B_1(0)$ is weak* compact and convex. Suppose K contains a non-zero measure μ. Then K must contain a non-zero extreme point μ. Since μ is extreme we have $\|\mu\| = 1$ (otherwise we could write μ as a linear combination of 0 and a multiple of μ).

Suppose such a non-zero extreme point μ exists. Since $\overline{E}$ is an algebra $\int fg \mathrm{d}\mu = 0$ for all $f, g \in \overline{E}$. Thus $g\mathrm{d}\mu \in \mathcal{N}$ for all $g \in E$. Now suppose $g \in \overline{E}$ and $0 < g(x) < 1$ for all $x \in S$. Let $a = \int g \mathrm{d}|\mu|$ and $b = \int (1-g)\mathrm{d}|\mu|$. So $a, b > 0$ and $a + b = 1$ and $g\mathrm{d}\mu/a$, $(1-g)\mathrm{d}\mu/b \in K$. Since

$$\mathrm{d}\mu = a\tfrac{g}{a}\mathrm{d}\mu + b\tfrac{(1-g)}{b}\mathrm{d}\mu$$

and μ is an extreme point, we must have $g\mathrm{d}\mu/a = \mathrm{d}\mu$.Consider the support of μ:

$$\operatorname{supp} \mu \;=\; \{x \;:\; |\mu|(U) > 0 \text{ for any open neighborhood of } x\}.$$

Since $\mathrm{d}\mu = g\mathrm{d}\mu / \int gd|\mu|$ we must have $g = \int gd|\mu|$ on $\operatorname{supp} \mu$.

Now suppose x and y are distinct points in S. Then there is a function $g \in E$ such that $0 < g < 1$ and $g(x) \neq g(y)$ — let h be any function that separates x and y and let $g = \frac{1}{\gamma}(h + \delta G)$ for suitable γ and δ. Thus at most one of the points x, y lies in the support of μ. That is the support of μ is a single point $\operatorname{supp} \mu = \{x_0\}$! Since $|\mu|(1) = \|\mu\| = 1$ we have

$$\int f(p)\mathrm{d}\mu(p) \;=\; f(x_0) \quad \text{or} \quad \int f(p)\mathrm{d}\mu(p) \;=\; -f(x_0).$$

However, there is a function $f \in \overline{E}$ with $f(x_0) \neq 0$ and thus $\int f \mathrm{d}\mu \neq 0$, contradicting the definition of $\mathcal{N}$. Thus $\mathcal{N} = \{0\}$ and $\overline{E} = C_{\mathbb{R}}(S)$. □

Following the above proof, we find

Theorem 19.3. *Let S be a compact Hausdorff space. The extreme points of the unit ball $\{\mu : \int \mathrm{d}|\mu| \leq 1\} \subset \mathcal{M}_{\mathbb{R}}(S)$ are the point masses $\pm\delta(p - p_0)$.*

and

Theorem 19.4. *Let S 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 Stone-Weierstrass Theorem, as stated, does not hold for complex valued functions. Here is a simple example. Let $S = \{\mathrm{e}^{i\theta} \;:\; \theta \in \mathbb{R}\}$, the unit circle in the complex plane. Let

$$A(S) \;=\; \{f : S \to \mathbb{C} \;:\; f \text{ has a continuous extension to } \overline{\mathbb{D}} \text{ that is analytic on } \mathbb{D}.\}$$

Exercise 19.2. Show that $A(S)$ is a norm closed proper sub-algebra of $C(S)$ and that

$$E \;=\; \left\{ \sum_{n=0}^{N} a_n \mathrm{e}^{in\theta} \;:\; a_0, \ldots, a_N \in \mathbb{C} \text{ and } N \in \mathbb{N} \right\}$$

is a nowhere vanishing sub-algebra of $A(S)$ that separates points of S. Show that E is dense in $A(S)$ (but not in $C(S)$, of course, since $A(S)$ is closed and a proper subset.)

There is an easy remedy to the above problem. We say that a subalgebra $E \subset C(S)$ is *a $\star$-algebra* if $f^* \in E$ whenever $f \in E$. (Here f^* denotes the function with $f^*(x) = (f(x))^* =$ complex conjugate of $f(x)$.)

Theorem 19.5. *Let S be a compact Hausdorff space and let $C(S)$ denote the set of all complex valued continuous functions on S. If $E \subset C(S)$ is a nowhere vanishing $\star$-algebra that separates points, then E is dense in $C(S)$ in the max norm.*

PROOF. Let $E_{\mathbb{R}} = \{\operatorname{Re} f \,:\, f \in E\}$. Note that if $f \in E$, then $\operatorname{Re} f = \frac{1}{2}(f + f^*) \in E$. Thus

$$E_{\mathbb{R}} \;=\; \{f \in E \,:\, f(S) \subset \mathbb{R}\}\;.$$

We have $E_{\mathbb{R}} \subset C_{\mathbb{R}}(S)$ is a linear subspace, and if $\operatorname{Re} f, \operatorname{Re} g \in E_{\mathbb{R}}$ we have $\operatorname{Re} f \operatorname{Re} g \in E_{\mathbb{R}}$. Thus $E_{\mathbb{R}}$ is a (real) sub-algebra of $C_{\mathbb{R}}(S)$.

Let $p \in S$. There is $f \in C(S)$ such that $f(p) \neq 0$. Scaling by a suitable complex number, we may choose f such that $\operatorname{Re} f(p) \neq 0$. Thus $E_{\mathbb{R}}$ is nowhere vanishing. Let $q \neq p$ be another point in S. If $\operatorname{Re} f(q) \neq \operatorname{Re} f(p)$ then this function separates p and q. If not, then pick another function $g \in C(S)$ such that $g(q) \neq g(p)$. Replacing g by $ag - bf$ for suitable complex numbers a, b, we may choose g that $g(p) = 0$ and $\operatorname{Re} g(q) \neq 0$. Thus $E_{\mathbb{R}}$ separates points. Therefore $E_{\mathbb{R}}$ is dense in $C_{\mathbb{R}}(S)$.

Now let $f \in C(S)$ and let $\varepsilon > 0$. We can choose $g, h \in E_{\mathbb{R}}$ such that $\|g - \operatorname{Re} f\| \leq \varepsilon$ and $\|h - \operatorname{Im} f\| \leq \varepsilon$. Then $g + ih \in E$ and $\|g + ih - f\| < 2\varepsilon$. Thus E is dense in $C(S)$. ☐

2. Applications of Stone Weierstrass

There are numerous applications of the Stone-Weierstrass theorems, 19.2 and 19.5. Here are a few.

Theorem 19.6. *The set $\mathcal{S} = \{\psi_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta} \;:\; n \in \mathbb{Z}\}$ are an orthonormal basis for $L^2([0, 2\pi])$.*

PROOF. Note that it is an easy exercise to check that $\mathcal{S}$ is an orthonormal set. Thus it suffices to prove that $\mathcal{A} = \operatorname{span} \mathcal{S}$ is dense in $L^2([0, 2\pi])$.

Note that

$$\mathcal{A} \;=\; \{\sum_{n=-N}^{N} a_n f_n \;:\; a_{-N}, \ldots a_N \in \mathbb{C} \text{ and } N \in \mathbb{N}\}$$

is a $\star$-algebra in $C([0, 2\pi])$. It is nowhere vanishing since $1 \in \mathcal{A}$. It doesn't quite separate points, since $f(0) = f(2\pi)$ for all $f \in \mathcal{A}$. However, if we think of it as a $\star$-algebra over $C(T^1)$, with T^1 the unit circle, then it does separate points ($e^{i\theta}$ alone does this). Thus $\mathcal{A}$ is dense in $C(T^1)$, which we can identify with

$$C(T^1) \;=\; \{f \in C([0, 2\pi]) \;:\; f(0) = f(2\pi)\}\;.$$

Now let $f \in L^2([0, 2\pi])$ and $\varepsilon > 0$. I claim there is $g \in C(T^1)$ such that $\|f - g\|_2 < \varepsilon$. Explicitly, we can take extend f periodically on $\mathbb{R}$ (so that $f(\theta \pm 2\pi) = f(\theta)$) and let

$$g(\theta) \;=\; \frac{1}{2\delta} \int_{\theta-\delta}^{\theta+\delta} f(\phi) d\phi$$

for small enough δ. We can choose $p \in \mathcal{A}$ such that $\sup_\theta |g(\theta) - p(\theta)| < \varepsilon$. Then

$$\|f - p\|_2 \;\leq\; \|f - g\|_2 + \|g - p\|_2 \;\leq\; \varepsilon + \sqrt{2\pi}\varepsilon\;.$$

Thus $\mathcal{A}$ is dense in L^2. □

A similar argument can be used to prove the following:

Theorem 19.7. *Let μ be a compactly supported finite Borel measure on $\mathbb{R}$. Let $p_0, p_1, \ldots$ be the polynomials obtained from the Gram-Schmidt process applied to $1, x, x^2, \ldots,$ with respect to the real inner product $\langle f, g, \rangle = \int fg d\mu$. Then $\{p_j : j = 0, \ldots\}$ is an orthonormal basis for $L^2_{\mathbb{R}}(\mu)$.*

Part 6

Spectral Theory of Linear Maps

LECTURE 20

Banach Algebras and Spectral Theory

Reading: §11.4 and Ch. 17 in Lax

Definition 20.1. A *Banach algebra* is a Banach space (completed normed space) $\mathcal{A}$ on which we have defined an associative product so that $\mathcal{A}$ is an *algebra* and such that

$$\|AB\| \leq \|A\|\|B\|, \quad \|cA\| = |c|\|A\|.$$

The motivating example of a Banach algebra is the space $\mathcal{B}(X)$ of bounded linear maps from a Banach space X into itself. It turns out that a good deal of the theory of linear maps can be carried out in the more general context of Banach algebras. A Banach algebra may or may not have a unit I, which is an element such that $IA = AI = A$ for all A.

Here are some basic facts and definitions about Banach algebras. The proofs are left as exercises:

(1) If $\mathcal{A}$ has a unit, then the unit is unique.
(2) Any Banach algebra is a closed sub-algebra of an algebra with a unit. If $\mathcal{A}$ has a unit there is nothing to prove. If not, consider the space $\mathcal{A} \oplus \mathbb{C}$ with product

$$(A, z)(B, w) = (AB + zB + wA, zw)$$

and norm

$$\|(A, z)\| = \|A\| + |z|.$$

(3) An element $A \in \mathcal{A}$ is *invertible* if there exists $B \in \mathcal{A}$ such that $BA = AB = I$.
(4) It can happen that A has either a left inverse ($BA = I$) or a right inverse ($AB = I$) but is not invertible.

Exercise 20.1. Find an example of an operator in $\mathcal{B}(\ell^2(\mathbb{N}))$ that has a left inverse but no right inverse. Can you find one in $\mathcal{B}(\mathbb{C}^n)$? Why?

(5) If A has a left inverse B and a right inverse C then $B = C$ and A is invertible.

Theorem 20.2. *The set of invertible elements in $\mathcal{A}$ is open. Specifically, if A is invertible then $A + K$ is invertible provided $\|K\| < 1/\|A^{-1}\|$.*

Exercise 20.2. Using the geometric series to prove this theorem.

Definition 20.3. Let $\mathcal{A}$ be a Banach algebra with a unit. The *resolvent set* of $\rho(A)$ of $A \in \mathcal{A}$ is the set of $\zeta \in \mathbb{C}$ such that

$$\zeta I - A$$

is invertible. The *spectrum* $\sigma(A)$ of A is the complement of the resolvent set. The *resolvent* of A is the map $R : \rho(A) \to \mathcal{A}$ given by

$$R(\zeta) = (\zeta I - A)^{-1}.$$

It turns out that $R(\zeta)$ is an analytic function. To make sense of this, we should first consider what it means for a Banach space valued function to be analytic.

1. Interlude: Analytic functions

Definition 20.4. Let X be a Banach space. A function $f : \Omega \to X$, with $\Omega \subset \mathbb{C}$ an open set, is *strongly analytic* (or just *analytic*) if

$$\lim_{h\to 0} \frac{1}{h}\left[f(\zeta+h) - f(\zeta)\right]$$

exists as a norm limit for every $\zeta \in \Omega$, in which case the limit is denoted $f'(\zeta)$ or $\frac{\mathrm{d}}{\mathrm{d}\zeta} f(\zeta)$.

If f is a strongly analytic function, the following familiar facts from complex analysis hold:

(1) If f is analytic so is f'.
(2) f is analytic if and only if
  (a) The power series for f at any point converges in a disc centered at that point.
  (b) f is continuous and $f(\zeta) = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} f(z) \frac{1}{z-\zeta} \mathrm{d}z$ for any rectifiable closed curve $\mathcal{C}$ that can be contracted to a point in Ω and with winding number 1 around z. The integral here can be taken to be a vector valued *Riemann integral* since f is continuous.

Exercise 20.3. Prove these facts. While you are at it, verify that the Riemann integral can be defined for norm continuous functions taking values in a Banach space.

One might also define what looks like a weaker notion of analyticity:

Definition 20.5. A function $f : \Omega \to \mathbb{C}$ is *weakly analytic* if for every $\ell \in X'$ = the dual of X, $\ell(f(\zeta))$ is a (scalar) analytic function.

Clearly if f is strongly analytic, then it is weakly analytic. Suprisingly, the converse is true:

Theorem 20.6. *Let X be a Banach space and $\Omega \subset \mathbb{C}$ an open set. If $f : \Omega \to X$ is weakly analytic, then it is strongly analytic.*

PROOF. By the Cauchy integral formula

$$\ell(f(\zeta)) \;=\; \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} \frac{\ell(f(z))}{z-\zeta} \mathrm{d}z$$

for a suitably chosen curve $\mathcal{C}$. This formula holds if ζ is moved a little bit, so

$$\begin{aligned}
&\ell\left(\frac{f(\zeta+h)-f(\zeta)}{h} - \frac{f(\zeta+k)-f(\zeta)}{k}\right) \\
&\quad = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} \left[\frac{1}{h}\left(\frac{1}{z-h-\zeta} - \frac{1}{z-\zeta}\right) - \frac{1}{k}\left(\frac{1}{z-k-\zeta} - \frac{1}{z-\zeta}\right)\right] \ell(f(z))\mathrm{d}z \\
&\quad = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} \left[\frac{1}{z-h-\zeta} - \frac{1}{z-k-\zeta}\right] \frac{\ell(f(z))}{z-\zeta}\mathrm{d}z \\
&\quad = \frac{(h-k)}{2\pi \mathrm{i}} \int_{\mathcal{C}} \frac{1}{(z-h-\zeta)(z-k-\zeta)} \frac{\ell(f(z))}{z-\zeta}\mathrm{d}z.
\end{aligned}$$

It follows that

$$\left|\ell\left(\frac{1}{h-k}\left[\frac{f(\zeta+h)-f(\zeta)}{h} - \frac{f(\zeta+k)-f(\zeta)}{k}\right]\right)\right| \;\le\; M(\ell) < \infty$$

uniformly for all h, k sufficiently close to 0. By the Principle of Uniform Boundedness (Thm. 15.5), there is a constant $C < \infty$ such that

$$\left\| \frac{f(\zeta+h)-f(\zeta)}{h} - \frac{f(\zeta+k)-f(\zeta)}{k} \right\| \leq C|h-k|.$$

Thus the limit defining $f'(\zeta)$ exists and f is strongly analytic. □

2. Back to the spectrum

Theorem 20.7. *Let $\mathcal{A}$ be a Banach algebra. For any $A \in \mathcal{A}$ the resolvent $R(\zeta)$ is analytic on the resolvent set $\rho(A)$ and the spectrum $\sigma(A)$ is a non-empty, compact subset of $\{\zeta \leq \|A\|\}$.*

PROOF. The Neumann series shows that $R(\zeta)$ has a convergent power series at each point:

$$R(\zeta+h) = ((\zeta+h)I - A)^{-1} = \sum_{n=0}^{\infty} h^n (\zeta I - Z)^{n+1}$$

for small enought h. Analyticity follows.

Furthermore, for $\zeta > \|A\|$ we have

$$R(\zeta) = \sum_{n=0}^{\infty} A^n \frac{1}{\zeta^{n+1}}, \tag{20.1}$$

so $\sigma(A) \subset \{\zeta \leq \|A\|\}$. Since $\rho(A)$ is open it follows that $\sigma(A)$ is compact.

Integrating (20.1) around a large circle, we obtain

$$\int_{|\zeta|=r} R(\zeta) \mathrm{d}\zeta = 2\pi \mathrm{i} I$$

for $r > \|A\|$. Suppose $\sigma(A)$ were empty. Then we could contract the circle down to a point and would obtain 0. Since the result is not 0, $\sigma(A)$ is not empty. □

LECTURE 21

Spectral Radius, Functional Calculus and Spectral Mapping

1. Spectral radius

Having seen that $\sigma(A) \subset \{z \leq \|A\|\}$, it is natural to ask if this estimate is sharp. In other words if we let the *spectral radius* of A be

$$\text{sp-rad}(A) = \max\{|z| \, : \, z \in \sigma(A)\},$$

then can it happen that $\text{sp-rad}(A) < \|A\|$?

This can indeed happen, even for 2×2 matrices:

Exercise 21.1. Consider the matrix $A = \left(\begin{smallmatrix} 0 & t \\ 0 & 0 \end{smallmatrix}\right)$. Show that $\|A\| = |t|$ and $\sigma(A) = \{0\}$.

Note that the example in the exercise shows that that the gap between $\sigma(A)$ and the resolvent set $\rho(A)$ can be arbitrarily large. However, we do have the following

Theorem 21.1. *Let $\mathcal{A}$ be a Banach algebra and let $A \in \mathcal{A}$. Then* $\text{sp-rad}\, A = \lim_{n\to\infty} \|A^n\|^{1/n}$.

Remark. Note that the matrix $A = \left(\begin{smallmatrix} 0 & t \\ 0 & 0 \end{smallmatrix}\right)$ in the above exercise has $A^2 = 0$, so we have $\text{sp-rad}\, A = \left\|A^2\right\|$ in that case.

PROOF. Consider the Laurent expansion of the resolvent around ∞:

$$R(\zeta) = \sum_{n=0}^{\infty} \zeta^{-n-1} A^n. \tag{21.1}$$

I claim that this converges whenever $|\zeta| > \left\|A^k\right\|^{\frac{1}{k}}$. Indeed we can write

$$R(\zeta) = \left[\sum_{m=0}^{k-1} \zeta^{-m-1} A^m\right] \left[\sum_{n=0}^{\infty} \zeta^{-nk} A^k\right].$$

The first factor is a finite sum and the second converges if the above condition holds. Thus

$$\text{sp-rad}(A) \leq \liminf_{k\to\infty} \left\|A^k\right\|^{\frac{1}{k}}.$$

On the other hand, let $\phi(t) = (\text{sp-rad}(A) + \delta)\mathrm{e}^{it}$, $t \in [0, 2\pi]$. Then ϕ is a curve in the resolvent set which winds once around the spectrum. Integrating and using (21.1) we find that

$$\frac{1}{2\pi \mathrm{i}} \int_{\phi} \zeta^n R(\zeta) \mathrm{d}\zeta \;=\; A^n.$$

Thus

$$\|A^n\| \leq \left[\sup_{\zeta \in \phi([0,2\pi])} \|R(\zeta)\|\right] (\text{sp-rad}(A) + \delta)^{n+1}.$$

Taking the n^{th} root and sending n to infinity, we obtain

$$\limsup_{n\to\infty} \|A^n\|^{\frac{1}{n}} \le \text{sp-rad}(A) + \delta.$$

Note that $\sup_{\zeta\in\phi([0,2\pi])} \|R(\zeta)\| < \infty$, since the resolvent is analytic, hence continuous, on the resolvent set. Since δ was arbitrary, we have

$$\limsup \|A^n\|^{\frac{1}{n}} \le \text{sp-rad}(A) \le \liminf \|A^n\|^{\frac{1}{n}}$$

and the result follows. □

2. Functional calculus

The Cauchy integral formula

$$f(\zeta) = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} \frac{f(z)}{z-\zeta} \mathrm{d}z$$

for *scalar* analytic functions suggests a way of defining $f(A)$ for A in a Banach algebra.

Definition 21.2. Let $\mathcal{A}$ be a Banach algebra and let $A \in \mathcal{A}$. Given an open set $\Omega \supset \sigma(A)$ and an analytic function $f : \Omega \to \mathbb{C}$, we define

$$f(A) := \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} f(z)R(z)\mathrm{d}z, \tag{21.2}$$

with R the resolvent of A and $\mathcal{C}$ is any chain of contours in Ω that has winding number 1 around $\sigma(A)$ and winding number 0 around any point in Ω^c.

Exercise 21.2. Verify that this definition does not depend on the choice of chain $\mathcal{C}$.

That's that. We have defined $f(A)$ for any analytic function. For instance

$$A^2 = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} z^2 R(z)\mathrm{d}z.$$

But, WAIT!!!!!! We can't just define A^2 — it's already defined! We need to check something. Does this definition make sense? Yes it does.

Theorem 21.3. *Let $\mathcal{A}$ be a Banach algebra and let $A \in \mathcal{A}$. Then*

(1) *For any polynomial p, $p(A)$ (evaluated by algebra) is equal to the r.h.s. of* (21.2).
(2) *More generally if f has a power series $f(z) = \sum_n a_n(z-z_0)^n$ convergent in a disk $\{|z-z_0| < r\}$ which contains $\sigma(A)$ then*

$$\sum_{n=0}^{\infty} a_n(A - z_0 I)^n$$

is norm convergent and agrees with the r.h.s. of (21.2).
(3) *If f and g are two analytic functions defined in a neighborhood of $\sigma(A)$ then*

$$f(A)g(A) = [fg](A).$$

Exercise 21.3. Prove (1) and (2).

PROOF. To prove (3) we will use the following

Lemma 21.4 (Resolvent identity). *Let $z, w \in \rho(A)$. Then*

$$R(z) - R(w) = (w - z)R(z)R(w).$$

Exercise 21.4. Prove the resolvent identity.

Now let f, g be analytic in a neighborhood of $\sigma(A)$:

$$f(A) = \frac{1}{2\pi i}\int_{\mathcal{C}} f(z)R(z)\mathrm{d}z, \quad g(A) = \frac{1}{2\pi i}\int_{\mathcal{D}} g(z)R(z)\mathrm{d}z.$$

Without loss of generality, assume $\mathcal{C}$ lies *inside* $\mathcal{D}$ — so the winding number of $\mathcal{D}$ around any point $z \in \mathcal{C}$ is one. Then

$$f(A)g(A) = \frac{1}{2\pi i}\int_{\mathcal{C}}\int_{\mathcal{D}} f(z)g(w)R(z)R(w)\mathrm{d}w\mathrm{d}z = \frac{1}{2\pi i}\int_{\mathcal{C}}\int_{\mathcal{D}} \frac{f(z)g(w)}{w-z}\left[R(z) - R(w)\right]\mathrm{d}w\mathrm{d}z.$$

Let's compute each term separately:

$$\frac{1}{2\pi i}\int_{\mathcal{C}}\int_{\mathcal{D}} \frac{f(z)g(w)}{w-z}R(z)\mathrm{d}w\mathrm{d}z \;=\; \frac{1}{2\pi i}\int_{\mathcal{C}} f(z)g(z)R(z)\mathrm{d}z \;=\; [fg](A),$$

$$\frac{1}{2\pi i}\int_{\mathcal{C}}\int_{\mathcal{D}} \frac{f(z)g(w)}{w-z}R(w)\mathrm{d}w\mathrm{d}z \;=\; \frac{1}{2\pi i}\int_{\mathcal{D}}\left[\int_{\mathcal{C}} \frac{f(z)}{w-z}\mathrm{d}z\right] g(w)R(w)\mathrm{d}w \;=\; 0. \quad \square$$

The map $A \mapsto f(A)$ is called the Riesz functional calculus. Part (2) of the functional calculus shows that $f \mapsto f(A)$ is an algebraic homomorphism of the algebra of functions analytic in a neighborhood of $\sigma(A)$ into the Banach algebra $\mathcal{A}$. Next we will consider some of the analytic properties of this homomorphism.

3. Spectral mapping theorem and Riesz Projections

Theorem 21.5. *Let $\mathcal{A}$ be a Banach Algebra, $A \in \mathcal{A}$, and f analytic in a neighborhood of $\sigma(A)$.*

(1) (The spectral mapping theorem)*: $\sigma(f(A)) = f(\sigma(A))$*
(2) If g is analytic in a neighborhood of $\sigma(f(A))$ then

$$g(f(A)) = [g \circ f](A).$$

PROOF. To show (1) we need to show that $\zeta I - f(A)$ is invertible if and only if $\zeta \notin f(\sigma(A))$. If $\zeta \notin f(\sigma(A))$ then $h(z) = (\zeta - f(z))^{-1}$ is analytic in a neighborhood of $\sigma(A)$. But then

$$h(A)(\zeta I - f(A)) = I$$

by the multiplicative property of the functional calculus. On the other hand, if $\zeta \in f(\sigma(A))$, say $\zeta = f(w)$ with $w \in \sigma(A)$. Let

$$k(z) = \frac{f(w) - f(z)}{w - z},$$

so k is analytic in a neighborhood of $\sigma(A)$, and

$$k(A)(wI - A) = (wI - A)k(A) \;=\; \zeta I - f(A).$$

Suppose $(\zeta I - f(A))$ were invertible, then we would have

$$(\zeta I - f(A))^{-1}k(A)(wI - A) \;=\; (wI - A)k(A)(\zeta I - f(A))^{-1} \;=\; I,$$

which would imply that $w \notin \sigma(A)$, a contradiction.

To show (2), since $\sigma(f(A)) = f(\sigma(A))$ we have

$$g(f(A)) = \frac{1}{2\pi i} \oint_{\mathcal{D}} (\zeta I - f(A))^{-1} g(\zeta) d\zeta$$

with $\mathcal{D}$ a suitable contour. But

$$(\zeta I - f(A))^{-1} = \frac{1}{2\pi i} \oint_{\mathcal{C}} (zI - A)^{-1} \frac{1}{\zeta - f(z)} dz,$$

so

$$\begin{aligned} g(f(A)) &= \frac{1}{(2\pi i)^2} \oint_{\mathcal{D}\times\mathcal{C}} (zI - A)^{-1} \frac{1}{\zeta - f(z)} g(\zeta) dz d\zeta \\ &= \frac{1}{2\pi i} \oint_{\mathcal{C}} (zI - A)^{-1} g(f(z)) dz = [g \circ f](A). \quad \square \end{aligned}$$

The functional calculus $f \mapsto f(A)$ is often known as the "Riesz functional calculus" to distinguish it from the functional calculus we will develop later for self-adjoint operators, which will allow the evaluation of $f(A)$ for measurable functions.

One of the key conclusions of the spectral mapping theorem is the association of *projections* to each component of the spectrum of A.

Definition 21.6. A *projection* P in a Banach algebra $\mathcal{A}$ is any element of $\mathcal{A}$ which satisfies (1) $P^2 = P$ and (2) $P \neq 0$.

Proposition 21.7. *If P is a projection in a Banach algebra $\mathcal{A}$, then $P\mathcal{A}P = \{PAP : A \in \mathcal{A}\}$ is a Banach algebra with unit P. If $P \neq I$ then $\sigma(P) = \{0,1\}$.*

PROOF. Since $(PAP)(PBP) = P(APB)P$,

$$\|PAPPBP\| \leq \|PAP\|\|PBP\| \quad \text{and } P(PAP) = (PAP)P = PAP,$$

it follows that $P\mathcal{A}P$ is a Banach algebra with unit P.

To see that $\sigma(P) = \{0,1\}$, first note that

$$P(I - P) = (I - P)P = P - P = 0.$$

Thus neither P nor $(I - P)$ can be invertible. Since $P, (I - P) \neq 0$, it follows that $\{0,1\} \subset \sigma(P)$.

It remains to show that any $\zeta \in \mathbb{C} \setminus \{0,1\}$ is in the resolvent set. For this purpose, consider the Laurent series for the resolvent

$$(\zeta I - P)^{-1} = \sum_n \zeta^{-n-1} P^n,$$

convergent for $|\zeta| > \text{sp-rad}(P)$. Since $P^n = P$, $n \geq 1$, we may sum the series to get

$$R(\zeta) := (\zeta I - P)^{-1} = \frac{1}{\zeta}(I - P) + \sum_n \zeta^{-n-1} P = \frac{1}{\zeta}(I - P) + \frac{1}{\zeta - 1} P, \tag{21.3}$$

which is well defined for all $\zeta \in \mathbb{C} \setminus \{0,1\}$.

Exercise 21.5. Check that the r.h.s. of (21.3) is equal to $(\zeta I - P)^{-1}$ for $\zeta \neq 0, 1$. $\square$

Now suppose $A \in \mathcal{A}$ and

$$\sigma(A) = \cup_{j=1}^{N} \sigma_j$$

with σ_j disjoint. Then we can define

$$P_j = \frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{C}_j} (\zeta I - A)^{-1} \mathrm{d}\zeta$$

with $\mathcal{C}_j$ any contour that winds once around σ_j and zero times around σ_i, $i \neq j$.
Theorem 21.8.

(1) *P_j are projections*
(2) *$P_j P_i = 0$ for $i \neq j$*
(3) *$\sum_j P_j = I$.*
(4) *The spectrum of $P_j A P_j = A P_j = P_j A$, as an element of the algebra $P_j \mathcal{A} P_j$, is*

$$\sigma_{P_j \mathcal{A} P_j}(P_j A P_j) = \sigma_j.$$

Remark. The spectrum of $P_j A P_j$ in $\mathcal{A}$ is $\sigma_j \cup \{0\}$.

PROOF. Note that $P_j = f_j(A)$ with f_j an analytic function that is 1 in a neighborhood of σ_j and 0 in a neighborhood of σ_i for $i \neq j$. Thus (1), (2) and (3) follow from the functional calculus. □

The projections P_j are known as "Riesz projections." For matrices, they give the projection onto generalized eigenspaces. To see this, let us compute an example. Consider the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The spectrum of A is $\{-1, +1\}$. If we write down the resolvent

$$(\zeta I - A)^{-1} = \frac{1}{\zeta^2 - 1} \begin{pmatrix} \zeta & 1 \\ 1 & \zeta \end{pmatrix},$$

then we may compute

$$P_{\pm} = \frac{1}{2\pi \mathrm{i}} \oint_{z = \pm 1 + \mathrm{e}^{\mathrm{i}\theta}} \frac{1}{\zeta^2 - 1} \begin{pmatrix} \zeta & 1 \\ 1 & \zeta \end{pmatrix} \mathrm{d}\zeta = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}.$$

Exercise 21.6. Verify that $P_{\pm}^2 = I$ and $AP_{\pm} = P_{\pm} A = \pm P_{\pm}$.

More generally, the matrix may have non-trivial blocks in it's Jordan form.

Exercise 21.7. Compute the resolvent and Riesz projections for

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Exercise 21.8. Show that

$$f\left(\underbrace{\begin{pmatrix}\lambda & 1 & 0 & \cdots & 0\\ & \lambda & 1 & & \vdots\\ & & \ddots & \ddots & 0\\ & & & \ddots & 1\\ & & & & \lambda\end{pmatrix}}_{n\times n}\right) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda)\\ & f(\lambda) & f'(\lambda) & \ddots & \vdots\\ & & \ddots & \ddots & \frac{1}{2}f''(\lambda)\\ & & & \ddots & f'(\lambda)\\ & & & & f(\lambda)\end{pmatrix}$$

In infinite dimensions the Riesz projections may not be related to generalized eigenvectors. For instance the shift operator

$$S(a_0, a_1, \cdots) = (a_0, a_1, \cdots)$$

on ℓ^2 has spectrum

$$\sigma(S) = \{|z| \leq 1\}.$$

Thus S has only one Riesz projection — the identity map.

Exercise 21.9. S corresponds to the infinite matrix

$$S \sim \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \end{pmatrix}.$$

Show that, if f is analytic in a neighborhood of $\{|z| \leq 1\}$ then $f(S)$ corresponds to the infinite matrix

$$S \sim \begin{pmatrix} f(0) & f'(0) & \frac{1}{2}f''(0) & \cdots & \frac{1}{n!}f^{(n)}(0) & \cdots & \\ & f(0) & f'(0) & \frac{1}{2}f''(0) & \cdots & \frac{1}{n!}f^{(n)}(0) & \cdots \\ & & \ddots & \ddots & \ddots & & \ddots \\ & & & \ddots & \ddots & \ddots & & \ddots \end{pmatrix}.$$

In other words, if $\boldsymbol{a} = (a_0, a_1, a_2, \cdots)$ then

$$[f(S)\boldsymbol{a}]_j = j^{th} \text{ entry of } f(S)\boldsymbol{a} = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) a_{j+n}.$$

LECTURE 22

Commutative Banach Algebras

Reading: Chapter 18 in Lax

We will now specialize to spectral theory in an algebra $\mathcal{A}$ with a unit I and such that the multiplication is *commutative*: $AB = BA$ for all $A, B \in \mathcal{A}$. The theory we will develop here is due to Gelfand 1941. Throughout this lecture all Banach algebras will be commutative and have a unit.

Definition 22.1. A *multiplicative functional* p on a Banach algebra $\mathcal{A}$ is a homomorphism of $\mathcal{A}$ into $\mathbb{C}$.

So $p : \mathcal{A} \to \mathbb{C}$ is a linear functional *and* $p(AB) = p(A)p(B)$. This definition is *purely algebraic*. In particular, it is *not* assumed that p is bounded. However we have

Theorem 22.2. *Every multiplicative functional p on a commutative Banach algebra is a contraction: $|p(A)| \leq \|A\|$.*

PROOF. Since $p(A) = p(IA) = p(I)p(A) \quad \forall A \in \mathcal{A}$ we have either $p(A) = 0$ for all A or $p(I) = 1$. In the first case it is clear that p is a contraction. In the second case, if A is invertible then

$$p(A^{-1})p(A) = p(I) = 1,$$

and so

Lemma 22.3. *If $p \neq 0$ is a multiplicative functional on a Banach algebra and A is invertible then $p(A) \neq 0$.*

Now suppose $|p(A)| > \|A\|$ for some A. Let

$$B = \frac{A}{p(A)}.$$

Then $\|B\| < 1$. Thus $I - B$ is invertible, and

$$p(I - B) = p(I) - \frac{p(A)}{p(A)} = 0,$$

which is a contradiction. □

There is a lot of interplay between algebraic and analytic notions in the context of commutative Banach algebras.

Definition 22.4. A subset $\mathcal{I}$ of a commutative Banach algebra $\mathcal{A}$ is called an *ideal* if

(1) $\mathcal{I}$ is a linear subspace of $\mathcal{A}$
(2) $A\mathcal{I} \subset \mathcal{I}$ for any $A \in \mathcal{A}$
(3) $\mathcal{I} \neq 0, \mathcal{A}$.

Again, this is a purely algebraic notion. The following is a standard algebraic fact:

Proposition 22.5. *Let $\mathcal{A}$ and $\mathcal{B}$ be commutative algebras with units and $q : \mathcal{A} \to \mathcal{B}$ a homomorphism. Suppose that q*

(1) is not an isomorphism, and
(2) is not the zero map.

Then the

$$\ker q = \{A \in \mathcal{A} \,:\, q(A) = 0\}$$

is an ideal in $\mathcal{A}$. Conversely any ideal in $\mathcal{A}$ is the kernel of a homorphism satisfying (1) and (2)

SKETCH OF PROOF. It is easy to see that $\ker q$ is an ideal. Given an ideal $\mathcal{I}$, to construct the homorphism, we let $\mathcal{B} = \mathcal{A}/\mathcal{I}$. That is

$$\mathcal{B} = \{\text{equivalence classes for } A \sim B \text{ iff } A - B \in \mathcal{I}\}.$$

Check that $\mathcal{B}$ is an algebra with addition or multiplication given by addition or multiplication of any pair of representatives. Now let $q : \mathcal{A} \to \mathcal{B}$ be the map

$$q(A) = [A] = \text{ equivalence class containing } A. \quad \square$$

An ideal cannot contain *any* invertible elements. Indeed if A is invertible and A were in $\mathcal{I}$, then $A^{-1}A = I$ would be in $\mathcal{I}$ which would imply $AI = A \in \mathcal{I}$ for all A, that is $\mathcal{I} = \mathcal{A}$. On the other hand

Lemma 22.6. *Every non-invertible element B of $\mathcal{A}$ belongs to an ideal.*

PROOF. If $B = 0$ it is in every ideal, since ideals are, in particular, vector spaces. If $B \neq 0$ then $B\mathcal{A} = \{BA : A \in \mathcal{A}\}$ is an ideal and contains B.

Exercise 22.1. Show that $B\mathcal{A}$ is an ideal if B is not invertible.

$\square$

Definition 22.7. A *maximal ideal* is an ideal that is not contained in a larger ideal.

The space of ideals in $\mathcal{A}$ can be partially ordered by inclusion. It is easy to see that the union of an *arbitrary collection* of ideals is itself an ideal. Thus Zorn's lemma gives

Lemma 22.8. *Every ideal is contained in a maximal ideal. In particular, every non-invertible element $B \in \mathcal{A}$ belongs to a maximal ideal.*

Lemma 22.9. *Let $\mathcal{M}$ be a maximal ideal in $\mathcal{A}$. Every non-zero element of $\mathcal{A}/\mathcal{M}$ is invertible.*

Remark. That is, $\mathcal{A}/\mathcal{M}$ is a division algebra.

PROOF. Suppose $[B] \in \mathcal{A}/\mathcal{M}$ is not invertible. Then $[B]\mathcal{A}/\mathcal{M} = (B\mathcal{A})\,/\mathcal{M}$ is an ideal. Let

$$\mathcal{I} = \{A \in \mathcal{A} \,:\, [A] \in [B]\mathcal{A}/\mathcal{M}\},$$

that is

$$\mathcal{I} = \{A \in \mathcal{A} \,:\, A = BK + M \text{ for some } K \in \mathcal{A} \text{ and } M \in \mathcal{M}\}.$$

Exercise 22.2. Show that $\mathcal{I}$ is an ideal.

Since $\mathcal{I}$ is an ideal and clearly $\mathcal{I} \supset \mathcal{M}$, we must have $\mathcal{M} = \mathcal{I}$. Since $B = BI + 0 \in \mathcal{I}$, it follows that $B \in \mathcal{M}$. That is, $[B] = 0$. $\square$

So far we have done no analysis on ideals. To proceed we need an analytic result:

Theorem 22.10 (Gelfand-Mazur). *Let $\mathcal{A}$ be a Banach algebra with unit that is a division algebra. Then $\mathcal{A}$ is isomorphic to $\mathbb{C}$.*

PROOF. Let $B \in \mathcal{A}$. The spectrum of B is non-empty. Thus there is $\zeta \in \mathbb{C}$ such that $\zeta I - B$ is non-invertible. Since $\mathcal{A}$ is a division algebra, $\zeta I = B$. Thus every element of $\mathcal{A}$ is a multiple of the identity. The map $B \to \zeta$ is the isomorphism onto $\mathbb{C}$. □

We would like to conclude from Thm. 22.10 that $\mathcal{A}/\mathcal{M} \cong \mathbb{C}$ for any maximal ideal $\mathcal{M}$. Indeed, we have seen that $\mathcal{A}/\mathcal{M}$ is a division algebra. However, we *are not done* as we have not shown it is a Banach algebra. (There *are* division algebras not isomorphic to $\mathbb{C}$. For example, the algebra of rational functions on $\mathbb{C}$.)

To show that $\mathcal{A}/\mathcal{M}$ is a Banach algebra, we must show in particular that it is a Banach space. That this is true follows because

Lemma 22.11. *Let $\mathcal{I}$ be an ideal in a commutative Banach algebra. Then the closure $\overline{\mathcal{I}}$ of $\mathcal{I}$ is an ideal. In particular, a maximal ideal $\mathcal{M}$ is closed.*

Exercise 22.3. Prove this lemma.

Thus $\mathcal{A}/\mathcal{M}$ is a quotient of Banach spaces. It follows that it is a Banach space in the following norm:

$$\|[B]\| = \inf_{M \in \mathcal{M}} \|B + M\|.$$

(See Lax 2002, Chapter 5.)

Lemma 22.12. *Let $\mathcal{I}$ be a closed ideal in a commutative Banach algebra $\mathcal{A}$. Then $\mathcal{A}/\mathcal{I}$ is a Banach algebra.*

Exercise 22.4. Prove this lemma

Thus, given a maximal ideal $\mathcal{M}$, the quotient $\mathcal{A}/\mathcal{M}$ is a Banach division algebra and, thus, naturally isomorphic to $\mathbb{C}$ by Mazur's theorem. In particular, the quotient map

$$p(B) = [B]$$

is a multiplicative functional. In fact,

Theorem 22.13. *Let $\mathcal{A}$ be a commutative Banach algebra. There is a one-to-one correspondence between non-zero multiplicative functionals and maximal ideals given by*

$$\mathcal{M} \mapsto p_{\mathcal{M}}(B) = [B],\ p_{\mathcal{M}} : \mathcal{A} \to \mathcal{A}/\mathcal{M} \cong \mathbb{C},$$

and

$$p \mapsto \ker p.$$

PROOF. We have already seen that the quotient map associated to any maximal ideal is a multiplicative functional, so it remains to show that $\ker p$ is a maximal ideal for any multiplicative functional. This is a general algebraic fact. Since p is a non-zero linear functional, $\ker p$ is a subspace of co-dimension 1. Thus any subspace $\mathcal{V} \supset \ker p$ satisfies $\mathcal{V} = \mathcal{A}$ or $\mathcal{V} = \ker p$. Since any ideal $\mathcal{M} \supset \ker p$ is a subspace with $\mathcal{M} \neq \mathcal{A}$, we conclude that $\mathcal{M} = \ker p$ is a maximal ideal. □

Corollary 22.14. *An element B of a commutative Banach algebra with unit is invertible if and only if $p(B) \neq 0$ for all multiplicative functionals.*

PROOF. We have already seen that B invertible $\implies p(B) \neq 0$ for all multiplicative functionals. Conversely, if B is singular it is contained in a maximal ideal $\mathcal{M}$. Then $p_{\mathcal{M}}(B) = 0$. □

1. Spectral theory in commutative Banach Algebras

Theorem 22.15. *Let $\mathcal{A}$ be a commutative Banach algebra and let $B \in \mathcal{A}$. Then*

$$\sigma(B) = \{p(B) \ : \ p \text{ is a multiplicative linear functional}\}.$$

PROOF. $\zeta \in \sigma(B)$ if and only if $\zeta I - B$ is not invertible. We have seen that this happens if and only if $p(\zeta I - B) = 0$ for some multiplicative functional p. That is if and only if $\zeta = p(B)$. □

The set $\mathcal{J} = \{\text{maximal ideals in } \mathcal{A}\}$ is called the *spectrum* of the algebra $\mathcal{A}$. Using the correspondence $\mathcal{M} \sim p_{\mathcal{M}}$ between maximal ideals and multiplicative functionals established last time, we have a natural correspondence between $\mathcal{A}$ and an algebra of functions on $\mathcal{J}$, namely

$$A \ \mapsto \ f_A(\mathcal{M}) = p_{\mathcal{M}}(A), \tag{22.1}$$

where $p_{\mathcal{M}}$ is the multiplicative functional with kernel $\mathcal{M}$. This map is called the *Gelfand representation* of $\mathcal{A}$.

Theorem 22.16.

1. *The Gelfand representation is a homomorphism of $\mathcal{A}$ into the algebra of bounded functions on $\mathcal{J}$.*
2. *$|f_A(\mathcal{M})| \leq \|A\|$ for all $A \in \mathcal{A}$ and $\mathcal{M} \in \mathcal{J}$.*
3. *The spectrum of A is the range of f_A.*
4. *The identity I is represented by $f_I = 1$.*
5. *The functions f_A separate points of $\mathcal{J}$: if $\mathcal{M} \neq \mathcal{M}'$ are maximal ideals, then there is $A \in \mathcal{A}$ such that*

$$f_A(\mathcal{M}) \neq f_A(\mathcal{M}').$$

PROOF. The proofs of (1), (2), (3), and (4) are left as exercises. To see (5) note that given $A \in \mathcal{M} \setminus \mathcal{M}'$ we have $f_A(\mathcal{M}) = 0$ and $f_A(\mathcal{M}') \neq 0$. □

Definition 22.17. The *natural topology* on $\mathcal{J}$ is the *weakest* topology in which all the functions f_A, $A \in \mathcal{M}$, are continuous. It is called the *Gelfand topology*.

Theorem 22.18. *$\mathcal{J}$ is a compact Hausdorff space in the Gelfand topology.*

PROOF. The proof is based on Tychonoff's theorem. Let

$$P = \prod_{A \in \mathcal{A}} D_{\|A\|},$$

with $D_{\|A\|}$ the closed disk of radius $\|A\|$ in $\mathbb{C}$. By Tychonoff's theorem P is compact in the product topology. By part (2) of the first theorem

$$f_A(\mathcal{M}) \in D_{\|A\|},$$

so

$$\Phi(\mathcal{M})_A = f_A(\mathcal{M})$$

defines a map from $\mathcal{J} \to P$. By (5) this map is injective.

Exercise 22.5. Check that the Gelfand topology is the same as the topology induced on $\mathcal{J}$ by this embedding.

Since P is compact, it suffices to show that $\Phi(\mathcal{J})$ is closed.

Exercise 22.6. Show that $\Phi(\mathcal{J})$ is closed. Namely, show that any point $t = t.$ in the closure of $\Phi(\mathcal{J})$ is a homomorphism

$$t_{A+cB} = t_A + ct_B \quad \text{and} \quad t_{AB} = t_A t_B.$$

The Hausdorff property for $\mathcal{J}$ follows since f_A separate points. □

The Gelfand representation *need not be injective.* For example

$$\mathcal{A} = \left\{ \begin{pmatrix} z & w \\ 0 & z \end{pmatrix} \ : \ z, w \in \mathbb{C} \right\}$$

is a commutative Banach algebra, with identity. (Use the matrix norm, or

$$\left\| \begin{pmatrix} z & w \\ 0 & z \end{pmatrix} \right\| \ = \ |z| + |w|,$$

which is also sub-multiplicative.) It has a *unique* maximal ideal namely

$$\mathcal{M} = \left\{ \begin{pmatrix} 0 & w \\ 0 & 0 \end{pmatrix} \ : \ w \in \mathbb{C} \right\}.$$

The Gelfand homomorphism is the map

$$\begin{pmatrix} z & w \\ 0 & z \end{pmatrix} \to z, \quad \mathcal{A} \to \mathbb{C}.$$

In general, the kernel of the Gelfand representation is

$$\mathcal{R} = \bigcap_{\mathcal{M} \in \mathcal{J}} \mathcal{M},$$

which is called the *radical of* $\mathcal{A}$.

Proposition 22.19. *$A \in \mathcal{R}$ if and only if $\sigma(A) = \{0\}$.*

PROOF. This follows from the identity

$$\sigma(A) = \{p_{\mathcal{M}}(A) \ : \ \mathcal{M} \in \mathcal{J}\}. \quad \square$$

In particular, $\mathcal{R}$ contains all the nilpotent elements (if there are any). More generally, if $\|A^n\|^{\frac{1}{n}} \to 0$, so sp-rad$(A) = 0$, then $A \in \mathcal{R}$.

Proposition 22.20. *$\mathcal{R}$ is closed, and is an ideal if $\mathcal{R} \neq 0$.*

Exercise 22.7. Prove this.

The radical is essentially the barrier to representing $\mathcal{A}$ as an algebra of functions. Since $\mathcal{R}$ is closed, we may consider the quotient Banach algebra $\mathcal{A}/\mathcal{R}$, which has trivial radical. The Gelfand representation shows that:

Theorem 22.21. *If $\mathcal{A}$ is a commutative Banach Algebra, then there is a compact Hausdorff space Ω and continuous (bounded) injective homomorphism of $\mathcal{A}/\mathcal{R}$ into $C(\Omega)$.*

Thus a commutative Banach Algebra with trivial radical may be thought of as a subalgebra of the continuous functions on a compact Hausdorff space, and in fact the algebra determines the space.

LECTURE 23

C^* algebras

The algebra of functions on a compact Hausdorff space has an additional structure —complex conjugation —which is not present in commutative Banach algebras. What happens if we put it there?

More generally, we can define

Definition 23.1. A *$*$-operation* on a Banach algebra $\mathcal{A}$ is a map $A \mapsto A^*$ from $\mathcal{A} \to \mathcal{A}$ satisfying

(1) $(A^*)^* = A$
(2) $(AB)^* = B^*A^*$
(3) $(A+B)^* = A^* + B^*$
(4) $(wA)^* = \overline{w}A^*$.

A *C^* algebra* $\mathcal{A}$ is a Banach algebra together with a $*$-operation such that $\|A\|^2 = \|A^*A\|$.

The *prime* example of a C^* algebra is the algebra of bounded operators on a Hilbert space. In fact, although we will not show this, any C^* algebra is isometrically isomorphic to a sub-algebra of the bounded operators on a Hilbert space. A second example is $C_0(\Omega)$ with Ω a locally compact Hausdorff space. This example is commutative. If Ω is compact then $C_0(\Omega) = C(\Omega)$ has an identity. If Ω is non-compact then $C_0(\Omega)$ does not have an identity.

Proposition 23.2. *If $\mathcal{A}$ is a C^* algebra, then $\|A\| = \|A^*\|$.*

PROOF. Note that $\|A\|^2 = \|A^*A\| \le \|A^*\|\|A\|$, so $\|A\| \le \|A^*\|$. □

An *isomorphism* of C^* algebras $\mathcal{A}$ and $\mathcal{B}$ is a bounded linear isomorphism $T : \mathcal{A} \to \mathcal{B}$ that is multiplicative $T(AB) = T(A)T(B)$ and a *star map* $T(A)^* = T(A^*)$. An *isometric isomorphism* satisfies $\|T(A)\| = \|A\|$ for all A.

Spectral Theory in C^* algebras

Definition 23.3. An element of a C^* algebra is *self-adjoint* if $A^* = A$, is *anti-self-adjoint* if $A^* = -A$ and is *unitary* if $A^*A = I$.

Theorem 23.4. *If p is a multiplicative functional on a C^* algebra then*

(1) $p(A) \in \mathbb{R}$ if A is self-adjoint.
(2) $p(A^*) = \overline{p(A)}$.
(3) $p(A^*A) \ge 0$.
(4) $|p(U)| = 1$ if U is unitary.

PROOF. We already have $\|p\| \le 1$. Let A be self adjoint and suppose $p(A) = a + ib$. With $T_t = A + itI$, we have $a^2 + (b+t)^2 = |p(T_t)|^2 \le \|T_t\|^2$. Since $T_t^* = A - itI$, we have

$T_t^* T_t = A^2 + t^2 I$, and thus $\|T_t\|^2 \leq \|A\|^2 + t^2$. Therefore

$$a^2 + (b+t)^2 \leq \|A\|^2 + t^2.$$

This inequality can hold for all t if and only if $b = 0$. So $p(A) = a \in \mathbb{R}$, and (1) follows.

For general A we may write

$$A = \frac{1}{2}(A + A^*) + \mathrm{i}\frac{1}{2}(A - A^*),$$

and

$$A^* = \frac{1}{2}(A + A^*) - \mathrm{i}\frac{1}{2}(A - A^*),$$

so (2) follows from (1). (3) and (4) follow from (2) since $p(A^*A) = p(A^*)p(A)$. □

Corollary 23.5. *If $\mathcal{A}$ is a commutative C^* algebra then*

(1) If A is self adjoint $\sigma(A) \subset \mathbb{R}$.
(2) If A is anti-self-adjoint $\sigma(A) \subset \mathrm{i}\mathbb{R}$
(3) If A is unitary $\sigma(A) \subset \{|z| = 1\}$.

PROOF. Use the Gelfand theory:

$$\sigma(A) = \{p(A) \ : \ p \text{ is a multiplicative functional}\}. \quad \square$$

Given two Banach algebras $\mathcal{A} \subset \mathcal{B}$, both with an identity, and $A \in \mathcal{A}$ we can consider the spectrum of A as an element of $\mathcal{A}$ or $\mathcal{B}$. In general, these may be distinct. First of all the identity elements of $\mathcal{A}$ and $\mathcal{B}$ may be distinct — we saw this above with the Banach algebras $P\mathcal{A}P$ and $\mathcal{A}$. But even if the algebras share the same identity, the spectrum can change.

Example 23.6. Let T denote the unit circle, $\mathcal{B} = C(T)$, and

$$\mathcal{A} = A(T) = \{\text{continuous functions on circle with an analytic extension to the interior}\}.$$

Both algebras have the same norm and $\mathcal{A}$ is a closed subalgebra of $\mathcal{B}$. Consider the function $f(z) = z \in \mathcal{A} \subset \mathcal{B}$. As an element of $\mathcal{B}$ (which is a C^*-algebra), the spectrum of $f(z)$ is $C_{\mathcal{B}}(f) = T$, which is the range of f. However, as an element of $\mathcal{A} = A(T)$, we have $\sigma_{\mathcal{A}}(f) = \overline{D} = \{|\zeta| \leq 1\}$. To see this, note that if g is analytic in the disc and $(\zeta - z)g(z) = 1$ on T then $(\zeta - z)g(z) = 1$ in $\overline{D}$. Thus we must have $|\zeta| > 1$. So $\overline{D} \subset \sigma_{\mathcal{A}}(f)$. That they are equal follows from Lemma 23.8 below.

For C^*-algebras, this pheonomenonon does not occur:

Theorem 23.7. *Let $\mathcal{A} \subset \mathcal{B}$ be C^* algebras with the same identity and norm. If $A \in \mathcal{A}$ then $\sigma_{\mathcal{A}}(A) = \sigma_{\mathcal{B}}(A)$.*

We will need

Lemma 23.8. *If $\mathcal{A} \subset \mathcal{B}$ are Banach algebras with a common identity and $A \in \mathcal{A}$ then 1) $\sigma_{\mathcal{B}}(A) \subset \sigma_{\mathcal{A}}(A)$ and 2) $\partial\sigma_{\mathcal{B}}(A) \subset \partial\sigma_{\mathcal{A}}(A)$*

PROOF. Let I denote the identity in $\mathcal{A}$ and $\mathcal{B}$. If $(\zeta I - A)$ is invertible in $\mathcal{A}$, then since the identity is the same in $\mathcal{B}$ and $\mathcal{A}$, it is also invertible in $\mathcal{B}$. (1) follows.

Now suppose that $z \in \partial\sigma_{\mathcal{A}}(A)$. We must show that $z \in \sigma_{\mathcal{B}}(A)$ — it follows from (1) that then $z \in \partial\sigma_{\mathcal{B}}(A)$. Suppose on the contrary that there is $R \in \mathcal{B}$ such that

$$R(zI - A) \ = \ (zI - A)R \ = \ I \, .$$

Since $z \in \partial\sigma_{\mathcal{A}}(A)$ there are $z_n \to z$ with $z_n \in \mathbb{C} \setminus \sigma_{\mathcal{A}}(A)$. Thus $(z_n I - A)^{-1} \in \mathcal{A}$. It follows that

$$(z_n I - A)^{-1} \ \to \ R \quad \text{in } \mathcal{B}\,.$$

Since $\mathcal{A}$ is a closed subspace of $\mathcal{B}$ we then have $R \in \mathcal{A}$. This contradicts the fact that $z \in \sigma_{\mathcal{A}}(A)$. (2) follows. □

PROOF OF THM. 23.7. First let A be self-adjoint and let $\mathcal{C} = C^*(A) =$ the algebra generated by A. So $\mathcal{C}$ is a *commutative* C^* algebra and $\mathcal{C} \subset \mathcal{A} \subset \mathcal{B}$. Since $\mathcal{C}$ is commutative and A is self-adjoint, we have $\sigma_{\mathcal{C}}(A) \subset \mathbb{R}$. Thus

$$\sigma_{\mathcal{B}}(A) \subset \sigma_{\mathcal{A}}(A) \subset \sigma_{\mathcal{C}}(A) = \partial\sigma_{\mathcal{C}}(A) \subset \partial\sigma_{\mathcal{A}}(A) \subset \partial\sigma_{\mathcal{B}}(A).$$

It follows that $\sigma_{\mathcal{B}}(A) = \sigma_{\mathcal{A}}(A) = \sigma_{\mathcal{C}}(A)$.

To prove the general statement, it suffices to show that if A is invertible in $\mathcal{B}$ it is invertible in $\mathcal{A}$. So suppose we have $B \in \mathcal{B}$ such that $BA = AB = I$. It follows that

$$(A^*A)(BB^*) = (BB^*)(A^*A) = I.$$

Since A^*A is self-adjoint, the first part of the proof implies that A^*A is invertible in $\mathcal{A}$. Thus $BB^* \in \mathcal{A}$. Thus

$$B = B(B^*A^*) = (BB^*)A^* \in \mathcal{A}. \quad \square$$

Thus in *any* C^* algebra we can use the Gelfand theory to compute the spectrum, because the spectrum of an element A is the same as its spectrum in the smallest C^* algebra containing it

$$C^*(A) = \text{ algebra generated by } A \text{ and } A^*.$$

Definition 23.9. An element A in a C^* algebra is called *normal* if $AA^* = A^*A$.

Theorem 23.10. *For normal A in a C^* algebra,*

$$\text{sp-rad}(A) = \|A\|. \tag{$*$}$$

In particular, in a commutative C^* algebra $(*)$ holds for every A.

PROOF. For self-adjoint A we have $\|A^2\| = \|A\|^2$. It follows that $\left\|A^{2^k}\right\| = \|A\|^{2^k}$. Thus

$$\text{sp-rad}(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \lim_{k\to\infty} \left\|A^{2^k}\right\|^{2^{-k}} = \|A\|.$$

If A is normal, then $C^*(A)$ is commutative, so by the Gelfand theory we have

$$\|A\|^2 = \|A^*A\| = \text{sp-rad}(A^*A) = \sup_p p(A^*A) = \sup_p |p(A)|^2 = \text{sp-rad}(A)^2,$$

where the sup is over multiplicative functionals on $C^*(A)$. □

Functional Calculus

Theorem 23.11. *If $\mathcal{A}$ is a commutative C^* algebra with unit then there is a compact Hausdorff space Ω and an isometric isomorphism $\Phi : \mathcal{A} \to C(\Omega)$.*

PROOF. Let Ω be the maximal ideal space of $\mathcal{A}$ in the Gelfand topology, and let $\Phi : \mathcal{A} \to C(\Omega)$ be the Gelfand representation. By the previous theorem, sp-rad$(A) = 0 \implies A = 0$ so the radical of $\mathcal{A}$ is $\{0\}$. Thus the Gelfand representation is injective. Also,

$$\|A\|^2 = \sup_{\mathcal{M}} |p_{\mathcal{M}}(A)|^2 = \sup_{\mathcal{M}} |\Phi(A)(\mathcal{M})|^2,$$

so the Gelfand representation is an isometry.

It remains to show that the range of Φ is all of $C(\Omega)$. This follows from Stone-Weirstrass since $\Phi(\mathcal{A})$ is closed, separates points, and is closed under conjugation. $\square$

LECTURE 24

Continuous functional Calculus and Spectral Theorem for Self-Adjoint Operators

1. The Continuous Functional Calculus

Theorem 24.1 (Continuous functional calculus). *Let $\mathcal{A}$ be a C^* algebra with unit. If $A \in \mathcal{A}$ is normal, then there is a unique isometric isomorphism $f \mapsto f(A)$ of $C(\sigma(A))$ with $C^*(A)$ satisfying*

$$1 \mapsto I \quad \text{and} \quad z \mapsto A\,,$$

where 1 denotes the constant function equal to 1 on $\sigma(A)$ and z denotes the identity map $z \mapsto z$ on $\sigma(A)$. The mapping satisfies

$$p(A) = \sum_{j,k} a_{j,k} A^j (A^*)^k \tag{24.1}$$

for any polynomial $p(A) = \sum_{j,k} a_{j,k} z^j \bar{z}^k$.

PROOF. Since A is normal $C^*(A)$ is commutative and thus isomorphic to $C(\Omega)$ for some compact Hausdorff space and A is isomorphic to some function $\phi : \Omega \to \mathbb{C}$ with $\operatorname{ran}\phi = \sigma(A)$. Given $f \in C(\sigma(A))$, let $f(A) \in C^*(A)$ be the element that corresponds to $f \circ \phi$. The map $f \mapsto f(A)$ defined in this way is clearly linear and multiplicative (i.e., an algebra homomorphism). Furthermore, it is a star map, i.e., $f^*(A) = f(A)^*$, and

$$\|f(A)\| = \|f \circ \phi\| = \sup_{x\in\Omega} |f(\phi(x))| = \sup_{z\in\sigma(A)} |f(z)|\,,$$

so the map is an isometry.

It remains to show that the map is onto, i.e., that any element of $C^*(A)$ can be expressed as $f(A)$ for some $f \in C(\sigma(A))$. To this end note that (24.1) holds for any polynomial in z and $\bar{z}$, since $f \mapsto f(A)$ is an $*$-algebra homomorphism. By Stone-Weierstrass polynomials are dense in $C(\sigma(A))$. But we also have

$$C^*(A) = \text{closure of}\{p(A)\}\,,$$

since the right hand side is $C*$ algebra contained in any C^* algebra containing A. □

The map $f \mapsto f(A)$ is called the *continuous functional calculus* for the normal operators.

Theorem 24.2 (Spectral Mapping Theorem). *Let $\mathcal{A}$ be a C^* algebra and let $A \in \mathcal{A}$ be normal. The functional calculus satisfies*

(1) $\|f(A)\| = \sup_{\lambda\in\sigma(A)} |f(\lambda)|$,
(2) $\sigma(f(A)) = f(\sigma(A))$,
(3) *$f(A)$ is normal for all $f \in C(\sigma(A))$,*
(4) *$f(A)$ is self-adjoint if and only if f is real valued, and*
(5) *$f(A)$ is unitary if and only if $|f(z)| = 1$ for all $z \in \sigma(A)$.*

PROOF. content... □

2. Self-adjoint operators

Reading: Chapter 31

Recall that an operator $A \in \mathcal{L}(H)$, H a complex Hilbert space, is self-adjoint (or Hermitian or symmetric) if

$$\langle Ax, y \rangle = \langle x, Ay \rangle \quad \text{for all } x, y \in H.$$

We will assume unless mentioned otherwise that H is *separable*. So H is isomorphic either to $\mathbb{C}^n$ or ℓ^2.

If A is a self-adjoint matrix on $\mathbb{C}^N$ then there is a orthonormal basis of eigenvectors for A:

$$Ae_n = \lambda_n e_n, \quad \langle e_n, e_m \rangle = \delta_{n,m}, \quad \overline{\operatorname{span}\{e_n : n = 1, \ldots, N\}} = H.$$

Thus given $x \in H$ we have

$$x = \sum_n a_n e_n, \quad Ax = \sum_n a_n \lambda_n e_n,$$

with $a_n = \langle x_n, e_n \rangle$.

An expansion into eigenvectors does not exist for an arbitrary self-adjoint operator. For instance multiplication by x in $L^2(-1,1)$,

$$Mf(x) = xf(x)$$

has no eigenvectors in L^2. In this case there are eigenvectors in the sense of distributions:

$$M\delta(x - \lambda) = \lambda, \quad \lambda \in (-1, 1).$$

In some sense we have

$$f(x) = \int_{-1}^{1} f(\lambda)\delta(x - \lambda)\mathrm{d}\lambda, \quad Mf(x) = \int_{-1}^{1} f(\lambda)\lambda\delta(x - \lambda)\mathrm{d}\lambda,$$

analogous to the above expression.

Thus for a general self-adjoint operator the eigenvectors may be quite singular. A slightly better object to work with are projections onto the spaces spanned by eigenvectors with eigenvalues in some set S. In the case of $Mf(x) = xf(x)$, we have

$$E_M(S)f(x) = \chi_S(x)f(x),$$

which is a non-zero projection if S is a set of positive Lebesgue measure in $L^2(-1,1)$.

Let us rewrite, the expressions for a matrix A in this form. Let $E(\{\lambda_n\})$ for the projection onto the subspace of eigenvectors with eigenvalue λ_n. For a general set $S \subset \mathbb{R}$, let

$$E(S) = \sum_{\lambda_n \in S} E(\{\lambda_n\}).$$

Proposition 24.3.

(1) *$E(S)$ is a projection for each $S \subset \mathbb{R}$.*
(2) *$E(S)E(S') = E(S \cap S')$. In particular, if $S \cap S' = \emptyset$ then the* $\operatorname{ran} E(S) \perp \operatorname{ran} E(S')$.
(3) *If $S_1, \ldots, S_n$ are disjoint then*

$$E(S_1 \cup \cdots \cup S_n) = E(S_1) + \cdots E(S_n).$$

(4) *If S_j, $j = 1, \dots, \infty$, are disjoint then*

$$E(\cup_j S_j)x = \sum_{j=1}^{\infty} E(S_j)x$$

for each $x \in H$.

(5) *The same properties hold for $E_M(S)$ provided we restrict our attention to Lebesgue measurable sets.*

Exercise 24.1. Prove this.

The maps $S \mapsto E(S)$, $E_M(S)$ are *projection valued measures*. Note that

$$A = \int \lambda \mathrm{d}E((-\infty, \lambda]) \quad M = \int \lambda \mathrm{d}E_M((-\infty, \lambda]),$$

with the integrals understood as Stieltje's integrals in the strong operator topology. That is, for every $x \in H$,

$$Ax = \lim_{n \to \infty} \sum_{j=1}^{n} \lambda_j^{(n)} E(-\lambda_{j-1}^{(n)}, \lambda_j^{(n)})x,$$

with $\lambda_j^{(n)}$ a partition of the interval $[-\|A\|, \|A\|]$, say, with mesh size $\to 0$ as $n \to \infty$.

Definition 24.4. A *projection valued measure* over H is a map $E : \Sigma \to \mathcal{L}(H)$ defined on a sigma algebra of sets on some measurable space with the following properties.

(1) $E(S)$ is an orthogonal projection for every S.
(2) (finite additivity) $E(S_1) + E(S_2) + \cdots E(S_n) = E(S_1 \cup \cdots S_n)$ if S_j are disoint.
(3) (strong countable additivity) If S_j, $j = 1, \dots, \infty$, are disjoint then

$$E(\cup_j S_j)x = \sum_{j=1}^{\infty} E(S_j)x$$

for each $x \in H$.

Exercise 24.2. Derive $E(S)E(S') = E(S \cap S')$ from (1) and (2).

Associated to any projection valued measure on $\mathbb{R}$, $E(S)$, with compact support ($E(\mathbb{R} \setminus [-r, r]) = 0$ for some r) there is a bounded self-adjoint operator

$$A = \int \lambda \mathrm{d}E((-\infty, \lambda]).$$

Our ultimate goal is to show that the converse is true. This is the "spectral theorem." (If E does not have compact support, there is still a self-adjoint operator, but it is unbounded. We will get to this.)

That is where we are headed, but it will take a little while.

Theorem 24.5. *The spectrum of a bounded, self-adjoint operator M on a Hilbert space is a compact subset of the real line and* sp-rad$(M) = \|M\|$.

PROOF. This follows from the results on C^* algebras. □

Given a self-adjoint operator A on a Hilbert space, the functional calculus $f \mapsto f(A)$ is a bounded linear map from $C(\sigma(A)) \to \mathcal{L}(H)$. As such, the maps

$$\ell_{x,y}(f) = \langle f(A)x, y \rangle ,$$

defined for every pair $x, y \in H$, are bounded linear functionals. According to the Riesz representation theorem, then, to each pair $x, y \in H$ there corresponds a complex regular Borel measure on $\sigma(A)$ such that

$$\langle f(A)x, y \rangle = \int_{\sigma(A)} f(\lambda) \mathrm{d}m_{x,y}(\lambda).$$

Theorem 24.6.

(1) $m_{x,y}$ is sesquilinear *in x, y (linear in x and conjugate linear in y).*
(2) $m_{y,x} = \overline{m_{x,y}}$.
(3) $\|m_{y,x}\| \leq \|x\| \|y\|$.
(4) The measures $m_{x,x}$ are non-negative.

Remark 24.7. $\|m_{x,y}\|$ denotes the total variation norm of $m_{x,y}$:

$$\|m_{x,y}\| = \sup_{f \in C(\sigma(A)) \,:\, |f(\lambda)| \leq 1} \left| \int_{\sigma(A)} f(\lambda) \mathrm{d}m_{x,y}(\lambda) \right|.$$

PROOF. (1), (2), and (3) are left as exercises. To prove (4) note that if $f(\lambda) \geq 0$ on $\sigma(A)$, then $\sqrt{f} \in C(\sigma(A))$, so

$$\langle f(A)x, x \rangle = \left\langle \sqrt{f}(A) \sqrt{f}(A)x, x \right\rangle = \left\langle \sqrt{f}(A)x, \sqrt{f}(A)x \right\rangle \geq 0.$$

Thus

$$\int_{\sigma(A)} f(\lambda) \mathrm{d}m_{x,x}(\lambda) \;\geq\; 0$$

if $f \geq 0$. It follows from the Riesz theorem that $m_{x,x}$ is a non-negative measure. ☐

Thus for every Borel set $S \subset \sigma(M)$ we have a quadratic form

$$Q_S(x, y) = m_{x,y}(S),$$

which is bounded ($|Q_S(x, y)| \leq \|x\| \|y\|$), sesquilinear and skew-symmetric under interchange of x and y ($Q_S(x, y) = |Q_S(y, x)|$).

Theorem 24.8. *Associated to any function $B(x, y)$ on $H \times H$ which is bounded, sesquilinear and skew-symmetric, there is a bounded self-adjoint operator M such that*

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

PROOF. Fix y and consider $B(x, y)$ as a function of x. This a bounded linear functional on H, so by the Riesz-Frechet theorem there is $w \in H$ such that

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

Let $My = w$. Since $\|w\| \leq c\|y\|$, so M is bounded. Self-adjointness of M follows form the skew-symmetry:

$$\langle x, My \rangle = B(x, y) = \overline{B(y, x)} = \overline{\langle y, Mx \rangle} = \langle Mx, y \rangle . \quad ☐$$

Thus to each Borel subset $S \subset \sigma(A)$ is associated a bounded self-adjoint operator $E(S)$ so that

$$m_{x,y}(S) = \langle E(s)x, y\rangle \, .$$

Theorem 24.9.

(1) $E(\emptyset) = 0$ *and* $E(\sigma(A)) = I$.
(2) *If* $S \cap T = \emptyset$ *then* $E(S \cup T) = E(S) + E(T)$.
(3) $E(S \cap T) = E(S)E(T)$
(4) *Each* $E(S)$ *is an orthogonal projection, and* $\operatorname{ran} E(S) \perp \operatorname{ran} E(T)$ *if* S, T *are disjoint.*
(5) $[E(S), E(T)] = 0$ *and* $[E(S), A] = 0$ *for all* S, T.
(6) $E(S)$ *is countably additive in the strong operator topology.*

Remark. So $E(S)$ is a projection valued measure as defined in Lecture 9.

PROOF. Clearly $m_{x,y}(\emptyset) = 0$ for all x, y, so $E(\emptyset) = 0$. Likewise

$$m_{x,y}(\sigma(A)) = \int_{\sigma(A)} \mathrm{d}m_{x,y}(\lambda) = \langle Ix, y\rangle = \delta_{x,y},$$

so $E(\sigma(A)) = I$. Part (2) follows from the additivity of the measures $m_{x,y}$.

Note that (3) is equivalent to

$$m_{x,y}(S \cap T) = m_{E(T)x,y}(S) \quad \forall x, y \in H,\ S, T \subset \sigma(A).$$

This in turn is equivalent to

$$\int_T f(\lambda)\mathrm{d}m_{x,y}(\lambda) = \langle f(A)E(T)x, y\rangle \quad \forall x, y \in H,\ f \in C(\sigma(A)),\ T \subset \sigma(A),$$

which is equivalent to

$$\int_{\sigma(A)} g(\lambda) f(\lambda)\mathrm{d}m_{x,y}(\lambda) = \langle f(A)g(A)x, y\rangle \, \forall x, y \in H,\ f, g \in C(\sigma(A)),$$

which holds. Since $E(S)$ is self-adjoint and $E(S)^2 = E(S)$, we see that $E(S)$ is an orthogonal projection. Since $E(S)E(T) = 0$ if $S \cap T = \emptyset$, we conclude that $\operatorname{ran} E(S) \perp \operatorname{ran} E(T)$. The first part of (4) follows from (3).

Exercise 24.3. Show that $[E(S), A] = 0$.

Exercise 24.4. Show that $E(S)$ is countably additive in the strong operator topology.

□

Thus,

Theorem 24.10. *To each self-adjoint operator* A*, there corresponds a unique projection valued measure* E *on* $\sigma(A)$ *such that*

$$f(A) = \int_{\sigma(A)} f(\lambda)\mathrm{d}E,$$

for all $f \in C(\sigma(A))$*, with the integral on the r.h.s. a norm convergent Riemann-Stieltjes integral.*

PROOF. We have already constructed the P.V.M. The uniqueness follows form the uniqueness in the Riesz theorem. If $I_1, \ldots, I_n$ are subsets of $\sigma(A)$ with $\cup_j I_j = \sigma(A)$ and I_j pairwise disjoint then

$$\left\| \sum_j a_j E(I_j) \right\| \le \max_j |a_j|$$

orthogonality of the ranges of $E(I_j)$. It follows that, for continuous f, the oscillation of f on the partition $I_1, \ldots, I_n$

$$\operatorname{Osc}(f, \{I_j\}) = \left\| \sum_j \sup_{x,y \in I_j} |f(x) - f(y)| E(I_j) \right\|$$

converges to zero as the mesh

$$\Delta = \max_j \operatorname{diam}(I_j) \to 0.$$

In the standard way, one concludes the existence of the integral. □

In particular, we have

$$I = \int_{\sigma(A)} \mathrm{d}E \,, \quad A = \int_{\sigma(A)} \lambda \mathrm{d}E.$$

One can refine the spectral resolution a bit further. Every Borel measure on μ on the line $\mathbb{R}$ can be written as a sum of three part:

$$\mu = \mu^{(p)} + \mu^{(sc)} + \mu^{(ac)},$$

where

(1) $\mu^{(p)}$ is the *point measure*: $\mu^{(p)}(S) = \sum_{x \in S} \mu(\{x\})$.
(2) $\mu^{(ac)}$ is the *absolutely continuous measure*: $\mathrm{d}\mu^{(ac)}(\lambda) = \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \mathrm{d}\lambda$.
(3) $\mu^{(sc)}$ is everything else, and is *singular continuous* — it has no atoms (so $F(\lambda) = \mu^{(sc)}(-\infty, \lambda)$ is a continuous function) but is supported on a set of measure 0 (so $F'(\lambda) = 0$ almost everywhere).

Applying this decomposition to the measures $m_{x,y}$ we obtain three distinct projection valued measures $E^{(p)}$, $E^{(sc)}$, and $E^{(ac)}$. The measures are orthogonal to one another, that is

$$H = H^{(p)} \oplus H^{(sc)} \oplus H^{(ac)} \quad H^{(\sharp)} = \operatorname{ran} E^{(\sharp)}(\sigma(A)).$$

Exercise 24.5. Show that $H^{(p)}$ = closed linear span of the eigenvectors of A and that $E^{(p)}(S) = \sum_{\lambda \in S} E(\{\lambda\})$, where the sum runs over the eigenvectors of A and $E(\{\lambda\})$ is the projection onto the corresponding eigenspace.

LECTURE 25

Compact Operators

A subset $S \subset X$ of a metric space X is *pre-compact* if $\overline{S}$ is compact.

Proposition 25.1. *Let X be a complete metric space and $S \subset X$. The following are equivalent.*

(1) *S is pre-compact.*
(2) *every sequence in S has a Cauchy subsequence.*
(3) *S is* totally bounded, *i.e., for every $\varepsilon > 0$, S can be covered by fintely many ε-balls.*

Corollary 25.2. *Let X be a Banach space. Then*

(1) *If S_1 and S_2 are precompact subsets of X, then $S_1 + S_2$ is precompact.*
(2) *If S is precompact in X, then so is its convex hull.*
(3) *If $S \subset X$ is precompact, and $T \in \mathcal{L}(X, Y)$ with Y a Banach space, then TS is precompact in Y.*

Definition 25.3. Let X and Y denote Banach spaces. A linear map $T : X \to Y$ is called *compact* if $CB_1(X)$ is precompact in Y, where $B_1(X)$ is the unit ball in X. We denote the set of compact maps from X to Y by $\mathcal{C}(X, Y)$.

Theorem 25.4. *Let X and Y be Banach spaces.*

(1) *$\mathcal{C}(X, Y)$ is a closed linear sub-space of $\mathcal{L}(X, Y)$.*
(2) *Let $T \in \mathcal{C}(X, Y)$ and let Z be a Banach space. If $M \in \mathcal{L}(Y, Z)$, then $MT \in \mathcal{C}(X, Z)$. If $M \in \mathcal{L}(Z, X)$, then $TM \in \mathcal{C}(Z, Y)$,*

PROOF. ... □

We will write $\mathcal{C}(X)$ for $\mathcal{C}(X, X)$, the set of compact maps from a Banach space X to itself.

Theorem 25.5. *Let X be a Banach space, let $M \in \mathcal{C}(X)$, and let $T = I - M$, with I the identity map. Then*

(1) $\dim \ker T$ *is finite.*
(2) *Let $N_j = \ker T^j$. Then there is an integer i such that $N_k = N_i$ for all $k \geq i$.*
(3) *The range of T is closed*

PROOF. If $Tu = 0$, then $u = Mu$. Thus $\ker T \in \operatorname{ran} M$. Since $M\overline{B_1(X)}$ is pre-compact, it follows that $\overline{B_1(\ker T)} = \overline{B_1(X)} \cap \ker T$ is compact. Thus, by Thm. 3.9, $\ker T$ is finite dimensional. Part (2) follows immediately, since $N_{j+1} \subset N_j \subset \ker T$ for all j. (Note that Thm. 3.9 does not imply that $\operatorname{ran} M$ is finite dimensional, since this space is not closed — so it is not a Banach space.)

To prove that $\operatorname{ran} T$ is closed, let $y_k \in \operatorname{ran} T$ be a Cauchy sequence converging to a limit y. So $y_k = x_k - Mx_k$. Let $d_k = \operatorname{dist}(x_k, \ker T)$. Let $z_k \in \ker T$ with $\|w_k\| < 2d_k$, $w_k = x_k - x_k$. Since $Tz_k = 0$, we have

$$Tw_k \;=\; Tx_k = y_k\,.$$

To obtain a contradiction, suppose that $d_k \to \infty$ (perhaps along a subsequence). Then,

$$\frac{1}{d_k} T w_k \;\to\; 0\,.$$

Let $u_k = \frac{1}{d_k} w_k$. Then $\|u_k\| \le 2$ and $Tu_k \to 0$. Since $(u_k)_k$ is bounded, there is a subsequence along which Mu_k converges. Thus, along this subsequence, $u_k \to u$ with $Tu = 0$ since $Tu_k = u_k - Mu_k \to 0$. But $\|u_k - z\| \ge 1$ for all $z \in \ker T$. This is a contradiction, so $\sup_k d_k < \infty$.

Using the sequence w_k constructed above, we have $Tw_k = w_k - Mw_k = y_k \to y$. Since d_k is bounded, and $\|w_k\| < 2d_k$, we see that the sequence $(w_k)_k$ is bounded. Thus by compactness of M, there is a subsequence of Mw_k that converges. But then, along this subsequence, $(w_k)_k$ also converges to a limit w which is easily seen to satisfy $w - Mw = Tw = y$. Thus $\operatorname{ran} T$ is closed. □

Theorem 25.6. *Let $M \in \mathcal{C}(X)$. Then $T = I - C$ satisfies*

$$\operatorname{ind} T \;:=\; \dim\ker T - \operatorname{codim}\operatorname{ran} T \;=\; 0\,.$$

PROOF. Suppose that $\ker T = \{0\}$. We must show $\operatorname{ran} T = X$. Suppose, on the contrary, that $\operatorname{ran} T = X_1 \subsetneq X$. Let $X_j = T^j X$. Then we must have $X_j \subsetneq X_{j-1}$ for all j (since $\ker T = \{0\}$). The map

$$T^j \;=\; I + \sum_{k=1}^{j} \binom{j}{k} C^k$$

is of the form identity + compact. Thus $X_j = \operatorname{ran} T^j$ is closed by the prior theorem. By Riesz's lemma 3.10, there is $x_j \in X_j$ with $\|x_j\| = 1$ and $\operatorname{dist}(x_j, X_{j+1}) > \frac{1}{2}$. If $k > j$, then

$$Mx_j - Mx_k \;=\; x_j - Tx_j - x_k + Tx_k\,,$$

where the last three terms on the right hand side all belong to X_{j+1}. Thus $\|M(x_j - x_k)\| > \frac{1}{2}$, which contradicts the compactness of M. Thus $\operatorname{ran} T = X_1 = X$ and $\operatorname{ind} T = 0$.

In the general case, with $\ker T \neq \{0\}$, let $N_j = \ker T^j$. As shown above, there is an index i such that $N_j = N_i$ for all $j \ge i$. Let $N = N_i$. Then $TN \subset N$ and so $MN \subset N$. Let $\widetilde{M} \in L(X/N)$ be defined by $M(x + N) = Mx + N$ (this makes sense because N is an invariant subspace for M).

Exercise 25.1. Prove that $\widetilde{M}$ is a compact map.

Let $\widetilde{T} = I - \widetilde{M}$. It follows that $\ker \widetilde{T} = 0$. Indeed, suppose not. Then there is $x \notin N$ such that $Tx \in N = \ker T^i$. Then $T^{i+1}x = 0$, so $x \in N_{i+1} = N_i = N$. Thus we can apply the trivial kernel case to conclude that $\operatorname{codim}\operatorname{ran} \widetilde{T} = 0$. Thus for every $y \in X$ there are $x \in X$ and $z \in N$ such that $y = Tx + z$. Thus $X = \operatorname{ran} T + N$. If $i = 1$ then $\operatorname{ran} T \cap N = \{0\}$. However, if $i > 1$, then $N \cap \operatorname{ran} T = N_{i-1}$. By basic linear algebra, we have

$$\dim N = \dim\operatorname{ran} T \cap N + \dim\ker T\,.$$

Thus

$$\operatorname{codim}\operatorname{ran} T \;=\; \operatorname{codim}\operatorname{ran}\widetilde{T} + \dim N - \dim\operatorname{ran} T \cap N \;=\; \dim\ker T\,. \quad \square$$

1. Spectral Theory

Theorem 25.7 (F. Riesz). *Let X be an infinite dimensional Banach space and $M \in \mathcal{C}(X)$. Then*

(1) *$\sigma(M)$ consists of a finite or countable set of complex numbers $\{\lambda_n\}$ whose only possible accumulation point is* 0.
(2) *Each non-zero $\lambda \in \sigma(M)$ is an eigenvalue of M, of finite algebraic and geometric multiplicity. That is*
   (a) $\dim\ker(M - \lambda I)$ *is finite, and*
   (b) *there is an integer i such that* $\ker(M - \lambda I)^k = \ker(M - \lambda I)^i$ *for all $k \geq i$.*
(3) *The resolvent $(\zeta I - M)$ has a pole at each nonzero $\lambda \in \sigma(M)$.*

Remarks. 1) If X is finite dimensional, then the theorem remains true with the exception that the spectrum need not have 0 as a limit point. 2) If X is infinite dimensional, it follows that $0 \in \sigma(M)$.

Proof. Let $T(\zeta) = I - \zeta^{-1}M$ for $\zeta \neq 0$. Then we have either $\operatorname{ran} T(\zeta) = X$ or $\ker T(\zeta) \neq \{0\}$ — this is the so-called "Fredholm alternative." Thus every non-zero point in the spectrum is an eigenvalue. Part (2) follows from the previous theorem.

If $\sigma(M)$ is finite, then it has no accumulation points. Suppose that $(\lambda_n)_{n=1}^\infty$ is a sequence of distinct eigenvalues with eigenvectors x_n. Define $Y_n = \operatorname{span}(x_1, \ldots, x_n)$. Then $Y_n \subsetneq Y_{n+1}$. Using Riesz's lemma 3.10, we may find $y_n \in Y_n \setminus Y_{n-1}$ with $\|y_n\| = 1$ and $\operatorname{dist}(y_n, Y_{n-1}) > \frac{1}{2}$. Now $y_n = \sum_{j=1}^n a_j^n x_j$. Thus

$$My_n - \lambda_n y_n \;=\; \sum_{j=1}^{n-1} (\lambda_j - \lambda_n) a_j^n x_j \;\in\; Y_{n-1}\,.$$

Thus for $n > m$, we have

$$My_n - My_m \;=\; \lambda_n y_n - y \quad y \in Y_{n-1}\,.$$

Thus, by definition of the y_n's, we have

$$\|My_n - My_m\| \;\geq\; \tfrac{1}{2}|\lambda_n|\,.$$

Since $(My_n)_n$ is pre-compact, it follows that $\lambda_n \to 0$.

Because 0 is the only possible accumulation point of $\sigma(M)$ it follows that this set is either finite or countable.

To show that the resolvent has a pole at non-zero $\lambda \in \sigma(M)$, take $\zeta \in \{0 < |z - \lambda| < \varepsilon\}$ with ε small enough that no other eigenvalue is in this punctured disc. If $(\zeta I - M)^{-1}x = u$ then $x = \zeta T(\zeta)u$. Let i be large enough that $N_{i+1} = N_i = N$ with $N_i = \ker(\lambda - M)^i$. Then N is an invariant subspace and we may consider the reduced operator $\widetilde{M}$ on X/N. We claim that λ is in the resolvent set of this operator. If it weren't, then it would be an eigenvalue since $\widetilde{M}$ is compact. So we would have $x \in X$ with $Mx = \lambda x + y$ with $y \in N$. But then $x \in N_{i+1} = N_i$. It follows that $(\lambda - \widetilde{M})$ is invertible and therefore that $(\zeta - \widetilde{M})$ is invertible for $|\zeta - \lambda|$ small enough.

We have shown that for $x \in X$ and ζ in the puncture disc, we can find $v_\zeta \in X$ and $y_\zeta \in N$ such that

$$\zeta v_\zeta - Mv_\zeta \;=\; x - y_\zeta\,.$$

Furthermore, we may choose the solution such that $\|v_\zeta\| \leq C\|x\|$, uniformly in the punctured disc.

We claim that we can find $u \in N$ such that

$$\zeta u - Mu \;=\; y \,.$$

Because N is a finite dimensional invariant subspace, this amounts to a finite dimensional linear algebra problem. As such the solution may be represent, e.g., by Kramer's rule, as a *rational function* of ζ with a pole of order at most $\dim N$ at $\zeta = \lambda$ (where the determinant vanishes). □

Let us look at the last theorem in light of the spectral theory of commutative Banach algebras. We can generate the Banach algebra $\mathcal{B}(M)$ as the smallest Banach sub-algebra of $\mathcal{L}(X)$ containing M. This is precisely the limits of polynomials $p(M)$. The radical of this algebra

Homework

Homework I

Exercise I.1.

(1) Let $\mathcal{M}$ be a (not necessarily closed) subspace of a Hilbert space $\mathcal{H}$. Prove that
$$\mathcal{M}^{\perp} = \{x : \langle x,\ y\rangle = 0 \text{ for } y \in \mathcal{M}\}$$
is a closed subspace and that $\overline{\mathcal{M}} = (\mathcal{M}^{\perp})^{\perp}$. (Reed and Simon 1980, Ch. II, Problem 6)

(2) Let $\mathcal{M}$ be a (not necessarily closed) subspace of a Banach space $\mathcal{B}$. Prove that
$$\mathcal{M}^{\perp} = \{\ell \in \mathcal{B}^{\star}\ :\ \ell(x) = 0 \text{ for } x \in \mathcal{M}\}$$
is a closed subspace of $\mathcal{B}^{\star}$ and that
$$\overline{\mathcal{M}}\ =\ \{x \in \mathcal{B}\ :\ \ell(x) = 0 \text{ for } \ell \in \mathcal{M}^{\perp}\}\ .$$

Exercise I.2. Let X and Y be normed linear spaces and let $\mathcal{B}(X,Y)$ denote the set of all bounded linear maps from X to Y.

(1) Prove that the operator norm (see Def. 4.4) is a norm and that $\mathcal{B}(X,Y)$ is a Banach space.

(2) Let $T \in \mathcal{B}(X,Y)$ and let $\overline{X}, \overline{Y}$ denote the completions of X and Y. Prove that there is a unique linear map $\overline{T} : \overline{X} \to \overline{Y}$ such that $\overline{T}|_X = T$.

(3) Prove that $\|T\| = \left\|\overline{T}\right\|$ for any $T \in \mathcal{B}(X,Y)$ and thus that $\mathcal{B}(\overline{X},\overline{Y})$ and $\mathcal{B}(X,Y)$ are isometric Banach spaces.

Exercise I.3. Let (X,μ) be a measure space and let $1 < p < \infty$. Prove that $L^p(\mu)$ is uniformaly convex.

Exercise I.4. Prove Prop. 10.13: *If X is a locally compact Hausdorff space, then $\mathcal{M}(X)$ and $\mathcal{M}_{\mathbb{R}}(X)$ are Banach spaces in the total variation norm.*

Exercise I.5. Let $\Omega \subset \mathbb{R}^d$ be an open set.

(1) Give $L^1_{\text{loc}}(\Omega)$ the LCS topology generated by the semi-norms $p_K(f) = \int_K |f(x)|dx$ for compact $K \subset \Omega$. Show that $L^1_{\text{loc}}(\Omega)$ is a Fréchet space.

(2) Show that $(L^1_{\text{loc}}(\Omega))^{\star} \cong L^{\infty}_c(\Omega)$, where
$$L^{\infty}_c(\Omega)\ =\ \{f \in L^{\infty}(\Omega)\ :\ \exists \text{ compact } K \subset \Omega \text{ such that } f(x) = 0 \text{ for a.e. } x \notin K\}\ .$$

(3) Define a suitable inductive limit topology on $L^{\infty}_c(\Omega)$ and prove that it is a complete LCS in this topology.

Exercise I.6. (Reed and Simon 1980, Ch. V, Problem 46)

(a) Suppose X is the strict inductive limit of $(X_n)_{n=1}^\infty$ and that $(Y_n)_{n=1}^\infty$ is an increasing family of subspaces of X so that for any n, there is N with $X_n \subset Y_N$. Prove that X is the strict inductive limit of $(Y_n)_{n=1}^\infty$.

(b) Let $K \subset \Omega \subset \mathbb{R}^d$ with K compact and Ω open. Prove that if $C_c^\infty(\Omega)$ has the inductive limit topology given by $(C_c^\infty(K_n))_{n=1}^\infty$ for some increasing family $(K_n)_{n=1}^\infty$ of comapct sets with $\bigcup_n K_n = \Omega$, then the restriction of this topology to $C_0^\infty(K^\circ)$ is the Fréchet topology for $C_0^\infty(K^\circ)$ defined in Lecture 13.

(c) Prove that the inductive limit topology on $C_c^\infty(\Omega)$ is independent of the choice of the increasing family $(K_n)_{n=1}^\infty$ of compact sets.

Exercise I.7. In this exercise you will prove the following generalization of Thm. 3.9: *If X is a locally compact, locally convex space, then X is finite dimensional.*

(1) Let $U \ni 0$ be an open set with $\overline{U}$ compact. Prove that there are $x_1, \ldots, x_n \in U$ such that $\overline{U} \subset \bigcup_{i=1}^n (x_i + \frac{1}{2}U)$. Conclude that there is a finite dimensional space M such that $U \subset M + \frac{1}{2}U$.

(2) If U is also convex, prove that $U \subset M + \frac{1}{2^n}U$ for any n and thus that $U \subset \overline{M}$. Conclude that $\overline{M} = X = M$.

Exercise I.8. Let X be a locally convex space. A subset $E \subset X$ is called *bounded* if for any open neighborhood $U \ni 0$ we have $E \subset nU$ for some n.

(1) Prove that E is bounded if and only if $\sup_{x\in E} p(x) < \infty$ for any continuous semi-norm $p : X \to [0, \infty)$

(2) Prove that X is a normed space (has a topology generated by a single norm) if and only if the topology on X is generated by finitely many semi-norms.

(3) Prove that X is a normed space if and only if there is a bounded open neighborhood of 0.

Appendices

APPENDIX A

The Zorn-Kuratowski Lemma

The proof of the Zorn-Kuratowski Lemma (Thm. 1.16) uses *ordinals*. Informally, an ordinal is the order type of a well-ordered set. (Recall that an ordered set is *well-ordered* if every subset has a minimal element.) Von-Neumann proposed an alternative, more formal, definition of ordinals:

Definition A.1. A set S is an *ordinal* if every element of S is a subset of S and S is well-ordered by with respect to inclusion.

Thus any element $R \in S$ is also a subset $R \subset S$. The order relation on elements is $R \leq R'$ if and only if $R \subset R'$. Finally given any subset $T \subset S$ (which may not be an element of S), there is a unique element $R \in T$ such that $R \subset R'$ for all $R' \in T$. The smallest ordinal is the emptyset $\emptyset$, which is typically denoted 0. The ordinal 1 is

$$1 = \{0\} = \{\emptyset\},$$

the ordinal 2 is

$$2 = \{0, 1\} = \{\emptyset, \{\emptyset\}\},$$

the ordinal 3 is

$$3 = \{0, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}.$$

Note that $2 \in 3$ and $2 \subset 3$, since $2 = \{0, 1\}$. More generally, a natural number n is the ordinal

$$n = \{0, 1, \ldots, n-1\}.$$

The set of all natural numbers is an ordinal, typically denoted ω:

$$\omega = \{0, 1, \ldots\}.$$

The ordinals do not stop here, there is an ordinal

$$\omega + 1 = \{0, 1, \ldots\} \cup \{\omega\}.$$

More generally, given *any* ordinal α we can define the *successor* of α by

$$\alpha + 1 = \alpha \cup \{\alpha\}.$$

The key facts we need about ordinals are the following:

(1) Given ordinals α and β, either $\alpha \subset \beta$ or $\beta \subset \alpha$. We define $\alpha \leq \beta$ if $\alpha \subset \beta$.
(2) Given ordinals α and β, we have $\alpha < \beta$, so $\alpha \leq \beta$ and $\alpha \neq \beta$, if and only if $\alpha \in \beta$.
(3) The class of all ordinals is not a set.[1]

[1]Assuming it to be a set leads to the Burali-Forti paradox: If there were a set S containing all ordinals, it would be an ordinal. Then the successor of S would be an ordinal. Thus, by definition of S, we would have $S + 1 \in S$. But then we must have $S < S + 1$ and $S + 1 < S$, a contradiction.

Any ordinal α satisfies:

$$\alpha = \bigcup_{\beta<\alpha} \beta.$$

An ordinal α is a *limit ordinal* if $\alpha \neq 0$ and α is not a successor to any ordinal. For example, ω is a limit ordinal.

Key tools in the proof of Zorn-Kuratowski lemma are *transfinite recursion* and *transfinite induction*. *Transfinite recursion* is used to define an object O_α for each ordinal α by first defining O_0 and then showing how O_α can be defined in terms of $O_{\alpha'}$ for $\alpha' < \alpha$. *Transfinite induction* demonstrates that a collection of statements S_α indexed by ordinals is true for every ordinal by proving S_0 and proving that $\forall \alpha' < \alpha, S'_\alpha$ implies S_α. This proof may *look* complicated because of words like *transfinite*. However, it does not involve any hard analysis.

PROOF OF THE ZORN-KURATOWSKI LEMMA. We prove the contrapositive: *If S is a partially ordered set with no maximal element, then S has a totally ordered subset without an upper bound.*

Suppose S is a poset with no maximal element. Let $\mathcal{T}$ denote the collection of totally ordered subsets of S. If $R \in \mathcal{T}$ is bounded, then there is $a \in S$ such that $a \notin R$ and $x \leq a$ for all $x \in R$. (To see this, first choose an upper bound b for R. If $b \notin R$ we are done. If $b \in R$, then there is a with $b < a$ since b is not maximal.) Thus, by the axiom of choice, we may define a function F on $\mathcal{T}$ as follows:

(1) If $R \in \mathcal{T}$ is not bounded, then $F(R) = \emptyset$ if $R \in \mathcal{T}$.
(2) If $R \in \mathcal{T}$ is bounded, then $F(R) = \{a\}$ where $x < a$ for all $x \in R$.

Note that $R \cup F(R)$ is totally ordered for each $R \in \mathcal{T}$ and that $R \cup F(R) \neq R$ if and only if R is bounded.

We now use *transfinite recursion* to define, for each ordinal α, an element $R_\alpha \in \mathcal{T}$ such that $R_{\alpha'} \subset R_\alpha$ if $\alpha' \leq \alpha$. First, note that $\emptyset \in \mathcal{T}$, so we may define

$$R_0 := \emptyset.$$

Next, given R_α, let

$$R_{\alpha+1} := R_\alpha \cup F(R_\alpha).$$

Finally, for a limit ordinal α, let

$$R_\alpha := \bigcup_{\alpha'<\alpha} R_{\alpha'}.$$

Fix α. One can verify, using transfinite induction, that $R_\alpha \subset R_{\alpha'}$ whenever $\alpha \leq \alpha'$. If R_α, then $R_\alpha \neq R_{\alpha'}$ for $\alpha < \alpha'$.

Now let β be an ordinal with cardinality strictly larger than the cardinality of $\mathcal{T}$. Consider the function $\phi : \beta \to \mathcal{T}$ given by $\phi(\alpha) = R_\alpha$. By the pigeon-hole principle, ϕ is not injective. Thus we have $\alpha < \alpha' < \beta$ such that $R_\alpha = R_{\alpha'}$. It follows that R_α is not bounded! □

Bibliography

Texts

Conway, J. B. *A Course in Functional Analysis*. Vol. 96. Graduate Texts in Mathematics. New York, NY: Springer New York, 2007.

Folland, G. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1999.

Lax, P. D. *Functional Analysis*. New York: Wiley, 2002.

Reed, M. and B. Simon. *Methods of Modern Mathematical Physics I: Functional Analysis*. San Diego, CA: Academic Press, 1980.

Rudin, W. *Principles of Mathematical Analysis*. 3rd ed. International Series in Pure and Applied Mathematics. McGraw-Hill, 1976.

— *Real and Complex Analysis*. 3rd ed. New York: McGraw-Hill, 1987.

— *Functional Analysis*. 2nd ed. International Series in Pure and Applied Mathematics. New York: McGraw-Hill, 1991.

Sarason, D. *Complex Function Theory*. Second. Providence, RI: American Mathematical Society, 2007.

Simon, B. *Real Analysis*. Providence, Rhode Island: American Mathematical Society, 2015.

Tao, T. *An Introduction to Measure Theory*. Graduate Studies in Mathematics 126. Providence, RI: American Mathematical Soc., 2011.

Other References

Alaoglu, L. "Weak Topologies of Normed Linear Spaces". *Ann. Math.* 41.1 (1940), pp. 252–267.

Baire, R. "Sur les fonctions de variables réelles". *Annali di Matematica, Serie III* 3.1 (1899), pp. 1–123.

Banach, S. "Sur Les Fonctionnelles Linéaires". *Studia Math.* 1.1 (1929), pp. 211–216.

Banach, S. "Sur Les Opérations Dans Les Ensembles Abstraits et Leur Application Aux Équations Intégrales". *Fund. Math.* 3 (1922), pp. 133–181.

— "Théorème Sur Les Ensembles de Première Catégorie". *Fundam. Math.* 16 (1930), pp. 395–398.

Banach, S. and H. Steinhaus. "Sur le principe de la condensation de singularités". *Fundam. Math.* 9.1 (1927), pp. 50–61.

Bing, R. H. "Metrization of Topological Spaces". *Can. J. Math.* 3 (1951), pp. 175–186.

Bohnenblust, H. F. and A. Sobczyk. "Extensions of Functionals on Complex Linear Spaces". *Bull. Amer. Math. Soc.* 44.2 (1938), pp. 91–93.

Carathéodory, C. "Über den variabilitätsbereich der fourier'schen konstanten von positiven harmonischen funktionen". *Rend. Circ. Matem. Palermo* 32.1 (1911), pp. 193–217.

Cech, E. "On Bicompact Spaces". *Ann. Math.* 38.4 (1937), pp. 823–844.

Chernoff, P. R. “A Simple Proof of Tychonoff’s Theorem Via Nets”. *Am. Math. Mon.* 99.10 (1992), pp. 932–934.

Clarkson, J. A. “Uniformly Convex Spaces”. *Trans. Amer. Math. Soc.* 40.3 (1936), pp. 396–414.

Dudley, R. M. *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics 74. Cambridge ; New York: Cambridge University Press, 2002.

Eberlein, W. F. “Weak Compactness in Banach Spaces: I”. *Proc. Natl. Acad. Sci. U.S.A.* 33.3 (1947), pp. 51–53.

Fischer, E. “Sur La Convergence En Moyenne”. *C. r. Académie Sci. Paris* 144 (1907), pp. 1022–1024.

Fréchet, M. “Sur Les Opérations Linéaires. III”. *Trans. Amer. Math. Soc.* 8.4 (1907), pp. 433–446.

— “Les Espaces Abstraits Topologiquement Affines”. *Acta Math.* 47.1-2 (1926), pp. 25–52.

Gelfand, I. “Normierte Ringe”. *Rec Math Moscou N Ser* 9 (1941), pp. 3–23.

Hahn, H. “Über Lineare Gleichungssysteme in Linearen Räumen.” *J. Für Reine Angew. Math. Crelles J.* 1927.157 (1927), pp. 214–229.

Hausdorff, F. *Grundzüge der Mengenlehre*. Leipzig: Veit & Company, 1914. reprinted in translation as *Set Theory* by AMS-Chelsea, Providence, 2021.

Helly, E. “Über lineare Funktionaloperationen.” *Wien. Ber.* 121 (1912), pp. 265–297.

Herglotz, G. “Über Potenzreihen Mit Positivem, Reelen Teil Im Einheitskreis”. *Berichte Über Verhandlungen K. Sächs. Ges. Wiss. Zu Leipz. Math.-Phys. Cl.* 63 (1911), pp. 501–511.

Hilbert, D. “Grundzüge Einer Allgemeinen Theorie Der Linearen Integralgleichungen. Fünfte Mitteilung”. *Nachrichten Von Ges. Wiss. Zu Gött. Math.-Phys. Kl.* 1906 (1906), pp. 439–480.

Jordan, P. and J. von Neumann. “On Inner Products in Linear, Metric Spaces”. *Ann. Math.* 36.3 (1935), pp. 719–723.

Kakutani, S. “Concrete Representation of Abstract (M)-Spaces (A Characterization of the Space of Continuous Functions)”. *Ann. Math.* 42.4 (1941), pp. 994–1024.

Krein, M. and D. Milman. “On Extreme Points of Regular Convex Sets”. *Stud. Math.* 9 (1940), pp. 133–138.

Kuratowski, C. “Une Méthode d’élimination Des Nombres Transfinis Des Raisonnements Mathématiques”. *Fund. Math.* 3 (1922), pp. 76–108.

— “La Propriété de Baire Dans Les Espaces Métriques”. *Fundam. Math.* 16 (1930), pp. 390–394.

Lax, P. D. and A. N. Milgram. “Parabolic Equations”. *Contributions to the Theory of Partial Differential Equations*. Annals of Mathematics Studies, No. 33. Princeton University Press, Princeton, N.J., 1954, pp. 167–190.

Markov, A. “On Mean Values and Exterior Densities.” *Mat. Sb.* 4(46).1 (1938), pp. 165–191.

Mazur, S. and S. Ulam. “Sur Les Transformation Isométriques d’espace Vectoriel Normés”. *C. r. Académie Sci. Paris* 194 (1932), pp. 946–948.

Milman, D. “On Some Criteria for the Regularity of Spaces of the Type (B)”. *C. R. (Dokl.) Acad. Sci. URSS, n. Ser.* 20 (1938), pp. 243–246.

Munkres, J. R. *Topology; a First Course*. Englewood Cliffs, N.J: Prentice-Hall, 1974.

Nagata, J.-i. “On a Necessary and Sufficient Condition of Metrizability”. *J. Inst. Polytech. Osaka City Univ. Ser. Math.* 1.2 (1950), pp. 93–100.

Nikodym, O. "Sur une généralisation des intégrales de M. J. Radon". *Fundam. Math.* 15.1 (1930), pp. 131–179.

Poincare, H. "Sur Les Equations Aux Derivees Partielles de La Physique Mathematique". *American Journal of Mathematics* 12.3 (1890), p. 211.

Radon, J. "Theorie Und Anwendungen Der Absolut Additiven Mengenfunktionen". *Sitzungsberichte Kais. Akad. Wiss.* Mathematisch-Naturwissenschaftliche Klasse 122.2a (1913), pp. 1295–1438.

Riesz, F. "Sur Les Systèmes Orthogonaux de Fonctions". *C. r. Académie Sci. Paris* 144 (1907), pp. 615–619.

— "Sur Certains Systèmes Singuliers d'équations Intégrales". *Ann. Sci. École Norm. Sup.* 28 (1911), pp. 33–62.

Riesz, F. "Über Lineare Funktionalgleichungen". *Acta Math.* 41.0 (1916), pp. 71–98.

— "Zur Theorie Des Hilbertschen Raumes". *Acta Sci. Math. Szeged.* 7 (1934), p. 34038.

Riesz, F. "Sur Une Espèce de Géométrie Analytique Systèmes de Fonctions of Sommables". *C. r. Académie Sci. Paris* 144 (1907), pp. 1409–1411.

Simon, B. *Harmonic Analysis*. Providence, Rhode Island: American Mathematical Society, 2015.

Smirnov, Y. "A Necessary and Sufficient Condition for Metrizability of Topological Spaces". *Dokl. Acad Nau SSSR NS* 77 (1951).

Smulian, V. "Über lineare topologische Räume". *Rec Math Moscou N Ser* 7 (1940), pp. 425–444.

Soukhomlinoff, G. "Über Fortsetzung von linearen Funktionalen in linearen komplexen Räumen und linearen Quaternionenräumen". *Rec Math Moscou N Ser* 3 (1938), pp. 353–358.

Stone, M. H. "Applications of the Theory of Boolean Rings to General Topology". *Trans. Amer. Math. Soc.* 41.3 (1937), pp. 375–481.

Tietze, H. "Über Funktionen, die auf einer abgeschlossenen Menge stetig sind." 1915.145 (1915), pp. 9–14.

Tychonoff, A. "Über die topologische Erweiterung von Räumen". *Math. Ann.* 102.1 (1930), pp. 544–561.

Urysohn, P. "Zum Metrisationsproblem". *Math. Ann.* 94.1 (1925), pp. 309–315.

von Neumann, J. "Mathematische Begründung Der Quantenmechanik". *Nachrichten Von Ges. Wiss. Zu Gött. Math.-Phys. Kl.* 1927 (1927), pp. 1–57.

Weierstrass, K. "Über Die Analytische Darstellbarkeit Sogenannter Willkürlicher Functionen Einer Reellen Veränderlichen". *Sitzungsberichte K. Preuss. Akad. Wiss. Zu Berl.* 1885 (1885), 633-639 and 789–805.

Zaremba, S. "Sur Le Calcul Numérique Des Fonctions Demandées Dans Le Problème de Dirichlet et Le Problème Hydrodynamique". *Bull. Int. L'Académie Sci. Crac. Cl. Sci. Mathématiques Nat.* 1909.2 (1909), pp. 125–264.

Zorn, M. "A Remark on Method in Transfinite Algebra". *Bull. Amer. Math. Soc.* 41.10 (1935), pp. 667–670.