FUNCTIONAL ANALYSIS

Lecture notes by Răzvan Gelca

Contents

1 Topological Vector Spaces **5**

1.1 What is functional analysis? . . . 5
1.2 The definition of topological vector spaces . . . 6
1.3 Basic properties of topological vector spaces . . . 7
1.4 Hilbert spaces . . . 9
1.5 Banach spaces . . . 16
1.6 Fréchet spaces . . . 18
1.6.1 Seminorms . . . 18
1.6.2 Fréchet spaces . . . 21

2 Linear Functionals **23**

2.1 The Riesz representation theorem . . . 23
2.2 The Riesz Representation Theorem for Hilbert spaces . . . 27
2.3 The Hahn-Banach Theorems . . . 29
2.4 A few results about convex sets . . . 32
2.5 The dual of a topological vector space . . . 35
2.5.1 The weak*-topology . . . 35
2.5.2 The dual of a normed vector space . . . 37

3 Bounded Linear Operators **43**

3.1 Continuous linear operators . . . 43
3.1.1 The case of general topological vector spaces . . . 43
3.2 The three fundamental theorems . . . 44
3.2.1 Baire category . . . 44
3.2.2 Bounded linear operators on Banach spaces . . . 44
3.3 The adjoint of an operator between Banach spaces . . . 50
3.4 The adjoint of an operator on a Hilbert space . . . 53
3.5 The heat equation . . . 59

4 Banach Algebra Techniques in Operator Theory **63**

4.1 Banach algebras . . . 63
4.2 Spectral theory for Banach algebras . . . 65
4.3 Functional calculus with holomorphic functions . . . 68
4.4 Compact operators, Fredholm operators . . . 71
4.5 The Gelfand transform . . . 74

5 C^* algebras **77**
5.1 The definition of C^*-algebras . . . 77
5.2 Commutative C^*-algebras . . . 79
5.3 C^*-algebras as algebras of operators . . . 80
5.4 Functional calculus for normal operators . . . 82

6 Distributions **93**
6.1 The motivation for using test functions . . . 93
6.2 Test functions . . . 95

A Background results **99**
A.1 Zorn's lemma . . . 99

Chapter 1

Topological Vector Spaces

1.1 What is functional analysis?

Functional analysis is the study of vector spaces endowed with a topology, and of the maps between such spaces.

Linear algebra in infinite dimensional spaces.

It is a field of mathematics where linear algebra and geometry+topology meet.

Origins and applications:

- The study of spaces of functions (continuous, integrable) and of transformations between them (differential operators, Fourier transform).
- The study of differential and integral equations (understanding the solution set).
- Quantum mechanics (the Heisenberg formalism).

Our goals:

- Understand the properties of linear spaces endowed with topologies. This can be applied to answering questions such as for which functions should the integral be defined, in what space should we look for the solution to a differential equation, etc.
- Understand subspaces and convex sets, finding bases. Fourier series expansions can be viewed as expansions in an orthonormal basis, and many special functions provide examples, too. Convex sets are related to optimization problems and knowing the extremal points of such sets is useful.
- Understand linear functionals. They are common in mathematics, an example is the integral of a function.

- Understand linear transformations (operators), find their spectra, learn how to do functional calculus with them. The spectrum of a linear differentiable operator is used when solving a differential equation via the method of stationary states. Functional calculus (actually the exponential) is also useful when solving differential equations such as the Schrödinger equation.

- Understand algebras of linear operators. They show up in quantum theory.

1.2 The definition of topological vector spaces

The field of scalars will always be either $\mathbb{R}$ or $\mathbb{C}$, the default being $\mathbb{C}$.

Definition. A *vector space* over $\mathbb{C}$ (or $\mathbb{R}$) is a set V endowed with an addition and a scalar multiplication with the following properties

- to every pair of vectors $x, y \in V$ corresponds a vector $x + y \in V$ such that
  $x + y = y + x$ for all x, y
  $x + (y + z) = (x + y) + z$ for all x, y, z
  there is a unique vector $0 \in V$ such that $x + 0 = 0 + x = x$ for all x
  for each $x \in V$ there is $-x \in V$ such that $x + (-x) = 0$.

- for every $\alpha \in \mathbb{C}$ (respectively $\mathbb{R}$) and $x \in V$, there is $\alpha x \in V$ such that
  $1x = x$ for all x
  $\alpha(\beta x) = (\alpha\beta)x$ for all α, β, x
  $\alpha(x + y) = \alpha x + \alpha y$, $(\alpha + \beta)x = \alpha x + \beta x$.

A linear map between two vector spaces is a map that preserves addition and scalar multiplication. An isomorphism between two vector spaces is a bijective homomorphism.

A set $C \in V$ is called convex if $tC + (1 - t)C \subset C$ for every $t \in [0, 1]$.

A set $B \subset V$ is called balanced if for every scalar α with $|\alpha| \leq 1$, $\alpha B \subset B$.

Definition. If V and W are vector spaces, a map $T : V \to W$ is called a *linear map* (or *linear operator*) if for every scalars α and β and every vectors $x, y \in V$,

$$T(\alpha x + \beta y) = \alpha T x + \beta T y.$$

Definition. A *topological space* is a set X together with a collection $\mathcal{T}$ of subsets of X with the following properties

- $\emptyset$ and X are in $\mathcal{T}$

- The union of arbitrarily many sets from $\mathcal{T}$ is in $\mathcal{T}$

- The intersection of finitely many sets from $\mathcal{T}$ is in $\mathcal{T}$.

The sets in $\mathcal{T}$ are called open.

If X and Y are topological spaces, the $X \times Y$ is a topological space in a natural way, by defining the open sets in $X \times Y$ to be arbitrary unions of sets of the form $U_1 \times U_2$ where U_1 is open in X and U_2 is open in Y.

Definition. A map $f : X \to Y$ is called *continuous* if for every open set $U \in Y$, the set $f^{-1}(U)$ is open in X.

A map is called a homeomorphism if it is invertible and both the map and its inverse are continuous. A topological space X is called *Hausdorff* if for every $x, y \in X$, there are open sets U_1 and U_2 such that $x \in U_1$, $y \in U_2$ and $U_1 \cap U_2 = \emptyset$.

A neighborhood of x is an open set containing x. A system of neighborhoods of x is a family of neighborhoods of x such that for every neighborhood of x there is a member of this family inside it.

A subset $C \in X$ is closed if $X \backslash C$ is open. A subset $K \in X$ is called compact if every covering of K by open sets has a finite subcover. The closure of a set is the smallest closed set that contains it. If A is a set then $\overline{A}$ denotes its closure. The interior of a set is the largest open set contained in it. We denote by $int(A)$ the interior of A.

Definition. A *topological vector space* over the field K (which is either $\mathbb{C}$ or $\mathbb{R}$) is a vector space X endowed with a topology such that every point is closed and with the property that both addition and scalar multiplication,

$$+ : X \times X \to X \text{ and } \cdot : K \times X \to X,$$

are continuous.

Given two topological vector spaces X and Y, we are mostly interested in maps between them that are both linear and continuous. As a caveat, because of the nature of practical applications sometimes we have to deal with noncontinuous maps. Two topological vector spaces are identified if there is a linear bijection between them that is continuous and has continuous inverse.

Example. $\mathbb{R}^n$ is an example of a finite dimensional topological vector space, while $C([0, 1])$ is an example of an infinite dimensional vector space.

A subset E of a topological vector space is called bounded if for every neighborhood U of 0 there is a number $s > 0$ such that $E \subset tU$ for every $t > s$.

A topological vector space is called locally convex if every point has a system of neighborhoods that are convex.

1.3 Basic properties of topological vector spaces

Let X be a topological vector space.

Proposition 1.3.1. For every $a \in X$, the translation operator $x \mapsto x + a$ is a homeomorphism.

As a corollary, the topology on X is completely determined by a system of neighborhoods at the origin; the topology is translation invariant.

Proposition 1.3.2. Let W be a neighborhood of 0 in X. Then there is a neighborhood U of 0 which is symmetric ($U = -U$) such that $U + U \subset W$.

Proof. Because addition is continuous there are open neighborhoods of 0, U_1 and U_2, such that $U_1 + U_2 \subset W$. Choose $U = U_1 \cap U_2 \cap (-U_1) \cap (-U_2)$. □

Proposition 1.3.3. Suppose K is a compact and C is a closed subset of X such that $K \cap C = \emptyset$. Then there is a neighborhood U of 0 such that

$$(K + U) \cap (C + U) = \emptyset.$$

Proof. Applying Proposition 1.3.2 twice we deduce that for every neighborhood W of 0 there is an open symmetric neighborhood U of 0 such that $U+U+U+U \subset W$. Since the topology is translation invariant, it means that for every neighborhood W of a point x there is an open symmetric neighborhood of 0, U_x, such that $x + U_x + U_x + U_x + U_x \subset W$.

Now let $x \in K$ and $W = X \backslash C$. Then $x + U_x + U_x + U_x + U_x \subset X \backslash C$, and since U_x is symmetric, $(x + U_x + U_x) \cap (C + U_x + U_x) = \emptyset$.

Since K is compact, there are finitely many points $x_1, x_2, \ldots, x_k$ such that $K \subset (x_1 + U_{x_1}) \cup (x_2 + U_{x_2}) \cup \cdots \cup (x_k + U_{x_k})$. Set $U = U_{x_1} \cap U_{x_2} \cap \cdots \cap U_{x_k}$. Then

$$\begin{aligned} K + U &\subset (x_1 + U_{x_1}) + U) \cup (x_2 + U_{x_2} + U) \cup \cdots \cup (x_k + U_{x_k} + U) \\ &\subset (x_1 + U_{x_1} + U_{x_1}) \cup (x_2 + U_{x_2} + U_{x_2}) \cup \cdots \cup (x_k + U_{x_k} + U_{x_k}) \\ &\subset X \backslash [(C + U_{x_1} + U_{x_1}) \cap \cdots \cap (C + U_{x_k} + U_{x_k}) \subset X \backslash (C + U)], \end{aligned}$$

and we are done. □

Corollary 1.3.1. Given a system of neighborhoods of a point, every member of it contains the closure of some other member.

Proof. Set K equal to a point. □

Corollary 1.3.2. Every topological vector space is Hausdorff.

Proof. Let K and C be points. □

Proposition 1.3.4. Let X be a topological vector space.
a) If $A \subset X$ then $\overline{A} = \cap(A + U)$, where U runs through all neighborhoods of 0.
b) If $A \subset X$ and $B \subset X$, then $\overline{A} + \overline{B} \subset \overline{A + B}$.
c) If Y is a subspace of X, then so is $\overline{Y}$.
d) If C is convex, then so are $\overline{C}$ and $int(C)$.
e) If B is a balanced subset of X, then so is $\overline{B}$, if $0 \in int(B)$, then $int(B)$ is balanced.
f) If E is bounded, then so is $\overline{E}$.

Theorem 1.3.1. In a topological vector space X,
a) every neighborhood of 0 contains a balanced neighborhood of 0,
b) every convex neighborhood of 0 contains a balanced convex neighborhood of 0.

Proof. a) Because multiplication is continuous, for every neighborhood W of 0 there are a number $\delta > 0$ and a neighborhood U of 0 such that $\alpha U \subset W$ for all α such that $|\alpha| < \delta$. The balanced neighborhood is the union of all αU for $|\alpha| < \delta$.

b) Let W be a convex neighborhood of 0. Let $A = \cap \alpha W$, where α ranges over all scalars of absolute value 1. Let U be a balanced neighborhood of 0 contained in W. Then

$U = \alpha U \subset \alpha W$, so $U \subset A$. It follows that $int(A) \neq \emptyset$. Because A is the intersection of convex sets, it is convex, and hence so is $int(A)$. Let us show that A is balanced, which would imply that $int(A)$ is balanced as well. Every number α such that $|\alpha| < 1$ can be written as $\alpha = r\beta$ with $0 \leq r \leq 1$ and $|\beta| = 1$. If $x \in A$, then $\beta x \in A$ and so $(1-r)0 + r\beta x = \alpha x$ is also in A by convexity. This proves that A is balanced. □

Proposition 1.3.5. a) Suppose U is a neighborhood of 0. If r_n is a sequence of positive numbers with $\lim_{n\to\infty} r_n = \infty$, then

$$X = \cup_{n=1}^{\infty} r_n U.$$

b) If δ_n is a sequence of positive numbers converging to 0, and if U is bounded, then $\delta_n U$, $n \geq 0$ is a system of neighborhoods at 0.

Proof. a) Let $x \in X$. Since $\alpha \mapsto \alpha x$ is continuous, there is n such that $1/r_n x \in U$. Hence $x \in r_n U$.

b) Let W be a neighborhood of 0. Then there is s such that if $t > s$ then $U \subset tW$. Choose $\delta_n < 1/s$. □

Corollary 1.3.3. Every compact set is bounded.

1.4 Hilbert spaces

Let V be a linear space (real or complex). An inner product on V is a function

$$\langle\ ,\ \rangle : V \times V \to \mathbb{C}$$

that sastisfies the following properties

- $\langle y, x\rangle = \overline{\langle x, y\rangle}$
- $\langle ax_1 + bx_2, y\rangle = a\langle x_1, y\rangle + b\langle x_2, y\rangle$
- $\langle x, x\rangle \geq 0$, with equality precisely when $x = 0$.

Example. The space $\mathbb{R}^n$ endowed with the inner product

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^T\mathbf{y}.$$

Example. The space $\mathbb{C}^n$ endowed with the inner product

$$\langle \mathbf{z}, \mathbf{w}\rangle = \mathbf{z}^T\overline{\mathbf{w}}.$$

Example. The space $C([0,1])$ of continuous functions $f : [0,1] \to \mathbb{C}$ with the inner product

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

The norm of an element x is defined by

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

and the distance between two elements is defined to be $\|x - y\|$. Two elements, x and y, are called orthogonal if

$$< x, y >= 0.$$

The norm completely determines the inner product by the polarization identity which in the case of vector spaces over $\mathbb{R}$ is

$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

and in the case of vector spaces over $\mathbb{C}$ is

$$\langle x, y\rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$

Note that we also have the parallelogram identity

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

Proposition 1.4.1. The norm induced by the inner product has the following properties:
a) $\|\alpha x\| = |\alpha|\|x\|$,
b) (the Cauchy-Schwarz inequality) $|\langle x, y\rangle| \leq \|x\|\|y\|$,
c) (the Minkowski inequality aka the triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$.

Proof. Part a) follows easily from the definition. For b), choose α of absolute value 1 such that $\langle \alpha x, y\rangle > 0$. Let also t be a real parameter. We have

$$\begin{aligned} 0 &\leq \|\alpha x t - y\|^2 = \langle \alpha x t - y, \alpha x t - y\rangle \\ &= \|\alpha x\|^2 t^2 - (\langle \alpha x, y\rangle + \langle y, \alpha x\rangle)t + \|y\|^2 \\ &= \|x\|^2 t^2 - 2|\langle x, y\rangle| t + \|y\|^2. \end{aligned}$$

As a quadratic function in t this is always nonnegative, so its discriminant is nonpositive. The discriminant is equal to

$$4(|\langle x, y\rangle|^2 - \|x\|^2\|y\|^2),$$

and the fact that this is less than or equal to zero is equivalent to the Cauchy-Schwarz inequality.

For c) we use the Cauchy-Schwarz inequality and compute

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \\ &\leq \|x\|^2 + |\langle x, y\rangle| + |\langle y, x\rangle| + \|y\|^2 = \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Hence the conclusion. □

Proposition 1.4.2. The vector space V endowed with the inner product has a natural topological vector space structure in which the open sets are arbitrary unions of balls of the form

$$B(x,r) = \{y \,|\, \|x-y\| \le r\}, \quad x \in V, r > 0.$$

Proof. The continuity of addition follows from the triangle inequality. The continuity of the scalar multiplication is straightforward. □

Definition. A Hilbert space is a vector space H endowed with an inner product, which is complete, in the sense that if x_n is a sequence of points in H that satisfies the condition $\|x_n - x_m\| \to 0$ for $m, n \to \infty$, then there is an element $x \in H$ such that $\|x_n - x\| \to 0$.

We distinguish two types of convergence in a Hilbert space.

Definition. We say that x_n converges strongly to x if $\|x_n - x\| \to 0$. We say that x_n converges weakly to x if $\langle x_n, y\rangle \to \langle x, y\rangle$ for all $y \in H$.

Using the Cauchy-Schwarz inequality, we see that strong convergence implies weak convergence.

Definition. The dimension of a Hilbert space is the smallest cardinal number of a set of elements whose finite linear combinations are everywhere dense in the space.

We will only be concerned with Hilbert spaces of either finite or countable dimension.

Definition. An orthonormal basis for a Hilbert space is a set of unit vectors that are pairwise orthogonal and such that the linear combinations of these elements are dense in the Hilbert space.

Proposition 1.4.3. Every separable Hilbert space has an orthonormal basis.

Proof. Consider a countable dense set in the Hilbert space and apply the Gram-Schmidt process to it. □

From now on we will only be concerned with separable Hilbert spaces.

Theorem 1.4.1. If e_n, $n \ge 1$ is an orthonormal basis of the Hilbert space H, then:
a) Every element $x \in H$ can be written uniquely as

$$x = \sum_n c_n e_n,$$

where $c_n =< x, e_n >$.
b) The inner product of two elements $x = \sum_n c_n e_n$ and $y = \sum_n d_n e_n$ is given by the Parseval formula:

$$\langle x, y\rangle = \sum_n c_n \overline{d_n},$$

and the norm of x is computed by the Pythagorean theorem:

$$\|x\|^2 = \sum_n |c_n|^2.$$

Proof. Let us try to approximate x by linear combinations. Write

$$\begin{aligned}\|x-\sum_{n=1}^{N} c_n e_n\|^2 &= \left\langle x-\sum_{n=1}^{N} c_n e_n, x-\sum_{n=1}^{N} c_n e_n\right\rangle\\ &= \langle x,x\rangle - \sum_{n=1}^{N}\overline{c_n}\langle x,e_n\rangle - \sum_{n=1}^{N} c_n\langle e_n,x\rangle + \sum_{n=1}^{N} c_n\overline{c_n}\\ &= \|x\|^2 - \sum_{n=1}^{N}|\langle x,e_n\rangle|^2 + \sum_{n=1}^{N}|\langle x,e_n\rangle - c_n|^2.\end{aligned}$$

This expression is minimized when $c_n = \langle x, e_n\rangle$. As a corollary of this computation, we obtain Bessel's identity

$$\|x-\sum_{n=1}^{N}\langle x,e_n\rangle e_n\|^2 = \|x\|^2 - \sum_{k=1}^{N}|\langle x,e_n\rangle|^2.$$

and then Bessel's inequality

$$\sum_{n=1}^{N}|\langle x,e_n\rangle|^2 \leq \|x\|^2.$$

Note that

$$\|\sum_{n=1}^{N}\langle x,e_n\rangle e_n\|^2 = \sum_{n=1}^{N}|\langle x,e_n\rangle|^2,$$

and so Bessel's inequality shows that $\sum_{n\geq 1}\langle x,e_n\rangle e_n$ converges.

Given that the set of vectors of the form $\sum_{n=1}^{N} c_n e_n$ is dense in the Hilbert space, and that such a sum best approximates x if $c_n =< x, e_n >$, we conclude that

$$x = \sum_{n=1}^{\infty}\langle x,e_n\rangle e_n.$$

This proves a).

The identities form b) are true for finite sums, the general case folows by passing to the limit. □

Remark 1.4.1. Because strong convergence implies weak convergence, if $x_n \to x$ in norm, then $\langle x_n, e_k\rangle \to \langle x, e_k\rangle$ for all k. So if $x_n \to x$ in norm then the coefficients of the series of x_n converge to the coefficients of the series of x.

Example. An example of a finite dimensional complex Hilbert space is $\mathbb{C}^n$ with the inner product $\langle \mathbf{z}, \mathbf{w}\rangle = \mathbf{z}^T\overline{\mathbf{w}}$. The standard orthonormal basis consists of the vectors $\mathbf{e}_k$, $k = 1, 2, \ldots, n$ where $\mathbf{e}_k$ has all entries equal to 0 except for the kth entry which is equal to 1.

Example. An example of a separable infinite dimensional Hilbert space is $L^2([0,1])$ which consists of all square integrable functions on $[0,1]$. This means that

$$L^2([0,1]) = \left\{ f : [0,1] \to \mathbb{C} \mid \int_0^1 |f(t)|^2 dt < \infty \right\}.$$

The inner product is defined by

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

An orthonormal basis for this space is

$$e^{2\pi i n t}, \quad n \in \mathbb{Z}.$$

The expansion of a function $f \in L^2([0,1])$ as

$$f(t) = \sum_{n=-\infty}^{\infty} \left\langle f, e^{2\pi i n t} \right\rangle e^{2\pi i n t}$$

is called the Fourier series expansion of f.

Note also that the polynomials with rational coefficients are dense in $L^2([0,1])$, and hence the Gram-Schmidt procedure applied to $1, x, x^2, \ldots$ yields another orthonormal basis. This basis consists of the Legendre polynomials

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n].$$

Example. The Hermite polynomials are defined as

$$H_n(x) = (-1)^n e^{x^2} \frac{d}{dx^n} e^{-x^2}$$

They form an orthogonal basis of the space $L^2(\mathbb{R}, e^{-x^2} dx)$. The polynomials $\pi^{-1/4} 2^{-n/2} (n!)^{-1/2} H_n(x)$ form an orthonormal basis of this space.

Example. The Hardy space on the unit disk $H^2(\mathbb{D})$. It consists of the holomorphic functions on the unit disk for which

$$\sup_{0<r<1} \left(\frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta \right)^{1/2}$$

is finite. This quantity is the norm of $H^2(\mathbb{D})$; it comes from an inner product. An orthonormal basis consists of the monomials $1, z, z^2, z^3, \ldots$.

Example. The Segal-Bargmann space

$$\mathcal{H}L^2(\mathbb{C}, \mu_\hbar)$$

which consists of the holomorphic functions on $\mathbb{C}$ for which

$$\int_{\mathbb{C}} |f(z)|^2 e^{-|z|^2/\hbar} dxdy < \infty$$

(here $z = x + iy$). The inner product on this space is

$$\langle f, g \rangle = (\pi\hbar)^{-1} \int_{\mathbb{C}} f(z)\overline{g(z)} e^{-|z|^2/\hbar} dxdy.$$

An orthonormal basis for this space is

$$\frac{z^n}{\sqrt{n!\hbar^n}}, \quad n = 0, 1, 2, \ldots.$$

Here is the standard example of an infinite dimensional separable Hilbert space.

Example. Let $K = \mathbb{C}$ or $\mathbb{R}$. The space $l^2(K)$ consisting of all sequences of scalars

$$x = (x_1, x_2, x_3, \ldots)$$

with the property that

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty.$$

We set

$$< x, y >= \sum_{n=1}^{\infty} x_n \overline{y_n}.$$

Then $l^2(K)$ is a Hilbert space (prove it!). The norm of an element is

$$\|x\| = \sqrt{\sum_{n=1}^{\infty} |x_n|^2}.$$

Theorem 1.4.2. Every two Hilbert spaces (over the same field of scalars) of the same dimension are isometrically isomorphic.

Proof. Let $(e_n)_n$ and $(e'_n)_n$ be orthonormal bases of the first, respectively second space. The map

$$\sum_n c_n e_n \mapsto \sum_n c_n e'_n$$

preserves the norm. The uniqueness of writing an element in an orthonormal basis implies that this map is linear. □

Corollary 1.4.1. Every separable Hilbert space over $\mathbb{C}$ is isometrically isomorphic to either $\mathbb{C}^n$ for some n or to $l^2(\mathbb{C})$. Every separable Hilbert space over $\mathbb{R}$ is isometrically isomorphic to either $\mathbb{R}^n$ for some n or to $l^2(\mathbb{R})$.

Subspaces of a Hilbert space

Proposition 1.4.4. A finite dimensional subspace is closed.

Proof. Let $E \subset H$ be an N-dimensional subspace. Using Gram-Schmidt, produce a basis $e_1, e_2, e_3, \ldots$ of H such that $e_1, e_2, \ldots, e_N$ is a basis for E. Then every element in E is of the form

$$x = \sum_{k=1}^{N} c_k e_k,$$

and because convergence in norm implies the convergence of coefficients, it follows that the limit of a sequence of elements in E is also a linear combination of $e_1, e_2, \ldots, e_N$, hence is in E. □

However, if the Hilbert space H is infinite dimensional, then there are subspaces which are not closed. For example if $e_1, e_2, e_3, \ldots$ is an orthonormal basis, then the linear combinations of these basis elements define a subspace which is dense, but not closed because it is not the whole space.

Definition. We say that an element x is orthogonal to a subspace E if $x \perp e$ for every $e \in E$. The orthogonal complement of a subspace E is

$$E^\perp = \{x \in H \mid \langle x, e\rangle = 0 \text{ for all } e \in E\}.$$

Proposition 1.4.5. $E^\perp$ is a closed subspace of H.

Proof. If $x, y \in E^\perp$ and $\alpha, \beta \in C$, then for all $e \in E$,

$$\langle \alpha x + \beta y, e\rangle = \alpha \langle x, e\rangle + \beta \langle y, e\rangle = 0,$$

which shows that E is a subspace. The fact that it is closed follows from the fact that strong convergence implies weak convergence. □

Theorem 1.4.3. (The decomposition theorem) If E is a closed subspace of the Hilbert space H, then every $x \in H$ can be written uniquely as $x = y + z$, where $y \in E$ and $z \in E^\perp$.

Proof. (proof from the book of Riesz and Nagy) Consider $y \in E$ as variable and consider the distances $\|x - y\|$. Let d be their infimum, and let y_n be a sequence such that

$$\|x - y_n\| \to d.$$

Now we use the parallelogram identity to write

$$\|(x - y_n) + (x - y_m)\|^2 + \|(x - y_n) - (x - y_m)\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

Using it we obtain

$$\begin{aligned} &\|y_n - y_m\|^2 = 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4\|x - \frac{y_n + y_m}{2}\|^2 \\ &\leq 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4d^2. \end{aligned}$$

The last expression converges to 0 when $m, n \to \infty$. This implies that y_n is Cauchy, hence convergent. Let $y \in E$ be its limit. Then $\|x - y\| = d$.

Set $z = x - y$. We will show that z is orthogonal to E. For this, let y_0 be an arbitrary element of E. Then for every $\lambda \in \mathbb{C}$,

$$\|x-y\|^2 = d^2 \le \|x - y - \lambda y_0\|^2 = \|x-y\|^2 - \overline{\lambda}\langle x-y, y_0\rangle - \lambda\langle y_0, x-y\rangle + \lambda\overline{\lambda}\langle y_0, y_0\rangle .$$

Set $\lambda = \langle x-y, y_0\rangle / \langle y_0, y_0\rangle$ to obtain

$$\frac{|\langle x-y, y_0\rangle|^2}{\|y_0\|^2} \le 0.$$

(Adapt this proof to prove Cauchy-Schwarz!)

It follows that $\langle x-y, y_0\rangle = 0$, and so $z = x - y \in E^\perp$.

If there are other $y' \in E$, $z' \in E^\perp$ such that $x = y' + z'$, then $y + z = y' + z'$ so $y - y' = z' - z \in E \cap E^\perp$. This implies $y - y' = z' - z = 0$, hence $y = y'$, $z = z'$ proving uniqueness.

□

This result yields the notation $H = E \oplus E^\perp$, where E is a closed subspace. In particular, for every closed subspace E, there is an orthonormal basis of H that is the union of an orthonormal basis of E and an orthonormal basis of $E^\perp$.

Corollary 1.4.2. If E is a subspace of H then $(E^\perp)^\perp = \overline{E}$.

Proof. Clearly $\overline{E}^\perp = E^\perp$ and $\overline{E} \subset (E^\perp)^\perp$, because if $x_n \in E$, $n \ge 1$ and $x_n \to x$, and if $y \in E^\perp$, then $0 = \langle x_n, y\rangle \to \langle x, y\rangle$. We have

$$H = \overline{E} \oplus \overline{E}^\perp = \overline{E}^\perp \oplus (\overline{E}^\perp)^\perp.$$

Hence $\overline{E}$ cannot be a proper subspace of $(E^\perp)^\perp$. □

Exercise. Show that every nonempty closed convex subset of H contains a unique element of minimal norm.

1.5 Banach spaces

Definition. A *norm* on a vector space X is a function

$$\|\cdot\| : X \to [0, \infty)$$

with the following properties

- $\|ax\| = |a|\|x\|$ for all scalars a and all $x \in X$.
- (the triangle inequality) $\|x + y\| \le \|x\| + \|y\|$.
- $\|x\| = 0$ if and only if $x = 0$.

The norm induces a translation invariant metric (distance) $d(x,y) = \|x-y\|$.

A vector space X endowed with a norm is called a normed vector space. Like in the case of Hilbert spaces, X can be given a topology that turns it into a topological vector space. The open sets are arbitrary unions of balls of the form

$$B_{x,r} = \{y \in X \mid \|x-y\| < r\}, \quad x \in X, r \in (0,\infty).$$

Definition. A *Banach space* is a normed vector space that is complete, namely in which every Cauchy sequence of elements converges.

Example. Every Hilbert space is a Banach space. In fact, the necessary structure for a Banach space to have an underlying Hilbert space structure (prove it!) is that the norm satisfies

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

Example. The space $\mathbb{C}^n$ with the norm

$$\|(z_1, z_2, \ldots, z_n)\|_\infty = \sup_k |z_k|$$

is a Banach space.

Example. Let $p \geq 1$ be a real number. The space $\mathbb{C}^n$ endowed with the norm

$$\|(z_1, z_2, \ldots, z_n)\|_p = (|z_1|^p + |z_2|^p + \cdots + |z_n|^p)^{1/p}$$

is a Banach space.

Example. The space $C([0,1])$ of continuous functions on $[0,1]$ is a Banach space with the norm

$$\|f\| = \sup_{t\in[0,1]} |f(t)|.$$

Example. Let $p \geq 1$ be a real number. The space

$$L^p(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{C} \mid \int_{\mathbb{R}} |f(t)|^p dt < \infty\},$$

with the norm

$$\|f\|_p = \left(\int_{\mathbb{R}} |f(t)|^p dt\right)^{1/p},$$

is a Banach space. It is also separable. In general the L^p space over any measurable space is a Banach space.

The space $L^\infty(\mathbb{R})$ of functions that are bounded almost everywhere is also Banach. Here two functions are identified if they coincide almost everywhere. The norm is defined by

$$\|f\|_\infty = \inf\{C \geq 0 \mid |f(x)| \leq C \text{ for almost every } x\}.$$

The space L^∞ is not separable.

Example. The Hardy space on the unit disk $H^p(\mathbb{D})$. It consists of the holomorphic functions on the unit disk for which

$$\sup_{0<r<1}\left(\frac{1}{2\pi}\int_0^{2\pi}|f(re^{i\theta})|^p d\theta\right)^{1/p}$$

is finite. This quantity is the norm of $H^p(\mathbb{D})$. The Hardy space $H^p(\mathbb{D})$ is separable.

Also $H^\infty(\mathbb{D})$, the space of bounded holomorphic functions on the unit disk with the sup norm is a Banach space.

Example. Let D be a domain in $\mathbb{R}^n$. Let also k be a positive integer, and $1 \le p < \infty$. The Sobolev space $W^{k,p}(D)$ is the space of all functions $f \in L^p(D)$ such that for every multi-index $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ with $|\alpha_1| + |\alpha_2| + \cdots + |\alpha_n| \le k$, the weak partial derivative $D^\alpha f$ belongs to $L^p(D)$.

Here the weak partial derivative of f is a function g that satisfies

$$\int_D f D^\alpha \phi dx = (-1)^{|\alpha|}\int_D g\phi dx,$$

for all real valued, compactly supported smooth functions ϕ on D.

The norm on the Sobolev space is defined as

$$\|f\|_{k,p} := \sum_{|\alpha|\le k}\|D^\alpha f\|_p.$$

The Sobolev spaces with $1 \le p < \infty$ are separable. However, for $p = \infty$, one defines the norm to be

$$\max_{|\alpha|\le k}\|D^\alpha f\|_\infty,$$

and in this case the Sobolev space is not separable.

1.6 Fréchet spaces

This section is taken from Rudin's Functional Analysis book.

1.6.1 Seminorms

Definition. A seminorm on a vector space is a function

$$\|\cdot\| : X \to [0, \infty)$$

satisfying the folowing properties

- $\|x\| \ge 0$ for all $x \in X$,
- $\|x + y\| \le \|x\| + \|y\|$ for all $x, y \in X$,

- $\|\alpha x\| = |\alpha| \|x\|$ for all scalars α and $x \in X$.

We will also denote seminorms by p to avoid the confusion with norms.

A convex set A in X is called absorbing if for every $x \in X$ there is $s > 0$ such that $sx \in A$. Every absorbing set contains 0. The Minkowski functional defined by an absorbing set is

$$\mu_A : X \to [0, \infty), \quad \mu_A(x) = \inf\{t > 0, t^{-1}x \in A\}.$$

For the intuitive picture (see also part c) of the proposition below), think about the example where the seminorm is actually a norm and A is the unit ball. Then this definition yields the norm.

Proposition 1.6.1. Suppose p is a seminorm on a vector space X. Then
a) $\{x \,|\, p(x) = 0\}$ is a subspace of X,
b) $|p(x) - p(y)| \leq p(x - y)$
c) The set $B_{0,1} = \{x : \,|\, p(x) < 1\}$ is convex, balanced, absorbing, and $p = \mu_{B_{0,1}}$.

Proof. a) For x, y such that $p(x) = p(y) = 0$, we have

$$0 \leq p(\alpha x + \beta y) \leq |\alpha| p(x) + |\beta| p(y) = 0,$$

so $p(\alpha x + \beta y) = 0$.

b) This is just a rewriting of the triangle inequality.

c) The fact that is balanced follows from $\|\alpha x\| = |\alpha| \|x\|$. For convexity, note that

$$p(tx + (1 - t)y) \leq tp(x) + (1 - t)p(y). \quad \square$$

Proposition 1.6.2. Let A be a convex absorbing subset of X.
a) $\mu_A(x + y) \leq \mu_A(x) + \mu_A(y)$,
b) $\mu_A(tx) = t\mu_A(x)$, for all $t \geq 0$. In particular, if A is balanced then μ_A is a seminorm,
c) If $B = \{x \,|\, \mu_A < 1\}$ and $C = \{x \,|\, \mu_C \leq 1\}$, then $B \subset A \subset C$ and $\mu_A = \mu_B = \mu_C$.

Proof. a) Consider $\epsilon > 0$ and let $t = \mu_A(x) + \epsilon$, $s = \mu_A(y) + \epsilon$. Then x/t and y/s are in A and so is their convex combination

$$\frac{x + y}{s + t} = \frac{t}{s + t} \cdot \frac{x}{t} + \frac{s}{s + t} \cdot \frac{y}{s}.$$

It follows that $\mu_A(x + y) \leq s + t = \mu_A(x) + \mu_A(y) + 2\epsilon$. Now pass to the limit $\epsilon \to 0$.

b) follows from the definition.

For c) note that the inclusions $B \subset A \subset C$ show that $\mu_C \leq \mu_A \leq \mu_B$. For the converse inequalities, let $x \in X$ and choose t, s such that $\mu_C(x) < s < t$. Then $x/s \in C$ so $\mu_A(x/s) \leq 1$ and $\mu_A(x/t) < 1$. Hence $x/t \in B$, so $\mu_B(x) \leq t$. Vary t to obtain $\mu_B \leq \mu_C$. $\square$

A family of seminorms $\mathcal{P}$ on a vector space is called separating if for every $x \neq y$, there is a seminorm $p \in \mathcal{P}$ such that $p(x - y) > 0$.

Proposition 1.6.3. Suppose X has a system of neighborhoods of 0 that are convex and balanced. Associate so each open set V in this system of neighborhoods its Minkowski functional μ_V. Then $V = \{x \in X \mid \mu_V(x) < 1\}$, and the family of functionals μ_V defined for all such V's is a separating family of continuous functionals.

Proof. If $x \in V$ then x/t is still in V for some $t < 1$, so $\mu_V(x) < 1$. If $x \notin V$, then $x/t \in V$ implies $t \geq 1$ because V is balanced and convex. This proves that $V = \{x \in X \mid \mu_V(x) < 1\}$.

By Proposition 1.6.2, μ_V is a seminorm for all V. Applying Proposition 1.6.1 b) we have that for every $\epsilon > 0$ if $x - y \in \epsilon V$ then

$$|\mu_V(x) - \mu_V(y)| \leq \mu_V(x-y) < \epsilon,$$

which proves the continuity of μ_V at x. Finally, μ_V is separating because X is Hausdorff. □

Theorem 1.6.1. Suppose $\mathcal{P}$ is a separating family of seminorms on a vector space X. Associate to each $p \in \mathcal{P}$ and to each positive integer n the set

$$B_{1/n,p} = \{x \mid p(x) < 1/n\}.$$

Let $\mathcal{V}$ be the set of all finite intersections of such sets. Then $\mathcal{V}$ is a system of convex, balanced, absorbing neighborhoods of 0, which defines a topology on X and turns X into a topological vector space such that every $p \in \mathcal{P}$ is continuous and a set A is bounded if and only if $p|A$ is bounded for all p.

Proof. Proposition 1.6.1 implies that each set $B_{1/n,p}$ is convex and balanced, and hence so are the sets in $\mathcal{V}$. Consider all translates of sets in $\mathcal{V}$, and let the open sets be arbitrary unions of such translates. We thus obtain a topology on X. Because the family is separating, the topology is Hausdorff. We need to check that addition and scalar multiplication are continuous.

Let U be a neighborhood of 0 and let

$$B_{1/n_1,p_1} \cap B_{1/n_2,p_2} \cap \cdots \cap B_{1/n_k,p_k} \subset U.$$

Set

$$V = B_{1/2n_1,p_1} \cap B_{1/2n_2,p_2} \cap \cdots \cap B_{1/2n_k,p_k}.$$

Then $V + V \subset U$, which shows that addition is continuous.

Let also V be as above. Because V is convex and balanced, $\alpha V \subset U$ for all $|\alpha| \leq 1$. This shows that multiplication is continuous. We see that every seminorm is continous at 0 and so by Proposition 1.6.1 it is continuous everywhere.

Let A be bounded. Then for each $B_{1/n,p}$, there is $t > 0$ such that $A \subset tB_{1/n,p}$. Hence $p < t/n$ on A showing that p is bounded on A. Conversely, if $p_j < t_j$ on A, $j = 1, 2, \ldots, n$, then $A \subset t_j B_{1,p_j}$, and so

$$A \subset \max(t_j) \cap_{j=1}^n B_{1,p_j}.$$

Since every open neighborhood of zero contains such an open subset, A is bounded. □

1.6.2 Fréchet spaces

Let us consider a vector space X together with a countable family of seminorms $\|\cdot\|_k$, $k = 1, 2, 3, \ldots$. We define a topology on X such that a set set is open if it is an arbitrary union of sets of the form

$$B_{x,r,n} = \{y \in X \mid \|x-y\|_k < r \text{ for all } k \leq n\}.$$

If the family is separating then X is a topological vector space.

The topology on X is Hausdorff if and only if for every $x, y \in X$ there is k such that $\|x-y\|_k > 0$, namely if the family of seminorms is separating.

Definition. A Fréchet space is a topological vector space with the properties that

- it is Hausdorff
- the topology is induced by a countable family of seminorms
- the topology is complete, meaning that every Cauchy sequence converges.

The topology is induced by the metric $d : X \times X \to [0, \infty)$,

$$d(x,y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\|x-y\|_k}{1 + \|x-y\|_k}.$$

This metric is translation invariant.

Recall that a metric is a function $d : X \times X \to [0, \infty)$ such that

- $d(x, y) = 0$ if and only if $x = y$,
- $d(x, y) = d(y, x)$,
- $d(x, y) + d(y, z) \geq d(x, z)$.

Example. Every Banach space is a Fréchet space.

Example. The space of smooth functions $C^\infty([0, 1])$ becomes a Fréchet space with the seminorms

$$\|f\|_k = \sup_{x \in [0,1]} |f^{(k)}(x)|.$$

Example. The space of continuous functions $C(\mathbb{R})$ is a Fréchet space with the seminorms

$$\|f\|_k = \sup_{\|x\| \leq k} |f(x)|.$$

Example. Let D be an open subset of the complex plane. There is a sequence of compact sets $K_1 \subset K_2 \subset K_3 \subset \cdots \subset D$ whose union is D. Let $\mathcal{H}(D)$ be the space of holomorphic functions on D endowed with the seminorms

$$\|f\|_k = \sup\{|f(z)| \mid z \in K_j\}.$$

Then $\mathcal{H}(D)$ endowed with these seminorms is a Fréchet space.

Theorem 1.6.2. A topological vector space X has a norm that induces the topology if and only if there is a convex bounded neighborhood of the origin.

Proof. If a norm exists, then the open unit ball centered at the origin is convex and bounded.

For the converse, assume V is such a neighborhood. By Theorem 1.3.1, V contains a convex balanced neighborhood U, which is also bounded. By Proposition 1.3.5, the sets rU, $r \geq 0$, form a family of neighborhoods of 0. Moreover, because U is bounded, for every x there is $r > 0$ such that $x \notin rU$. Let $\|\cdot\|$ be the Minkowski functional of this neighborhood. Then $x \notin rU$ implies $\|x\| \geq r$, so $\|x\| = 0$ if and only if $x = 0$. Thus $\|\cdot\|$ is a norm and the topology is induced by this norm. □

Chapter 2

Linear Functionals

In this chapter we will look at linear functionals

$$\phi : X \to \mathbb{C}(\text{or } \mathbb{R}),$$

where X is a vector space.

2.1 The Hamburger moment problem and the Riesz representation theorem on spaces of continuous functions

This section is based on a series of lectures given by Hari Bercovici in 1990 in Perugia.

The Hamburger Moment Problem: Given a sequence s_n, $n \geq 0$, when does there exist a positive function f such that

$$s_n = \int_{-\infty}^{\infty} t^n f(t) dt$$

for all $n \geq 0$?

These integrals are called "moments", a name motivated by mechanics where the second moment is the actual moment of inertia. Such integrals are quite useful when studying probability distributions.

We ask the more general problem, if there is a measure σ on $\mathbb{R}$ such that $t^n \in L^1(d\sigma)$ for all n and

$$s_n = \int_{-\infty}^{\infty} t^n d\sigma(t).$$

It is easy to see that not all such sequences are moments. For arbitrary complex numbers $a_0, a_1, ..., a_n$ set

$$p(t) = \sum_{j=0}^{n} a_j t^j.$$

Then

$$0 \leq \int_{-\infty}^{\infty} |p(t)|^2 d\sigma(t) = \sum_{j,k=0}^{n} \int_{-\infty}^{\infty} t^{j+k} d\sigma(t) a_j \overline{a_k}$$
$$= \sum_{j,k=0}^{n} s_{j+k} a_j \overline{a_k}.$$

This shows that for all n, the matrix

$$S_n = \begin{pmatrix} s_0 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots, & s_{n+1} \\ \dots & \dots & \dots & \dots \\ s_n & s_{n+1} & \dots & s_{2n} \end{pmatrix} \tag{2.1.1}$$

is positive semidefinite. So this is a necessary condition.

We will show that this is also a sufficient condition.

Theorem 2.1.1. (M. Riesz) Let X be a linear space over $\mathbb{R}$ and $C \subset X$ a cone, meaning that if $x, y \in C$ and $t > 0$ then $x + y \in C$ and $tx \in C$. Assume moreover that the cone is proper, meaning that $C \cap (-C) = \{0\}$ and define the order $x \leq y$ if and only if $y - x \in C$. Let $Y \subset X$ be a subspace and let $\phi_0 : Y \to \mathbb{R}$ be a linear functional such that $\phi_0(y) \geq 0$ for all $y \in Y \cap C$. Supose that for every $x \in X$ there is $u \in Y \cap C$ such that $u - x \in C$. Then there is a linear functional $\phi : X \to \mathbb{R}$ such that $\phi | Y = \phi_0$ and $\phi(x) \geq 0$ for all $x \in C$.

Proof. First, let us assume $X = Y + \mathbb{R}x$ with $x \notin Y$. Let us first consider the set

$$A = \{\phi_0(y) \,|\, y \in Y, x - y \in C\}.$$

We claim that A is bounded from above. Indeed, we can write $x = u - c$ with $u \in Y \cap C$, $c \in C$. Write also $x - y = c(y)$. Then $u - c - y = c(y)$, so $u - y \in C$. This implies that $u \geq y$, so $\phi_0(u) \geq \phi_0(y)$. We conclude that A is bounded from above. Define $\phi(x) = \sup A$, then extend linearly to X so that $\phi | Y = \phi_0$.

We have to show that if $z = \pm tx + y \in C$, then $\phi(z) \geq 0$. This is equivalent to showing that $\phi(z/t) \geq 0$ for $t > 0$, so we only have to check the cases where $z = y \pm x$.

In the first case,

$$\phi(x + y) = \phi(x - (-y)) = \phi(x) - \phi_0(-y) \geq 0$$

because $\phi_0(-y) \in A$.

In the second case, choose $y_1 \in Y$ such that $z_1 = x - y_1 \in C$ and $\phi_0(y_1) \geq \phi(x) - \epsilon$ (here we use the definition of the supremum). Then $y - y_1 = z + z_1 \geq 0$, so $\phi_0(y) - \phi_0(y_1) \geq 0$. Then

$$\phi(z) = \phi_0(y) - \phi(x) \geq \phi_0(y) - \phi_0(y_1) - \epsilon.$$

Now make $\epsilon \to 0$ to obtain $\phi(z) \geq 0$.

For the general case of a space X, use transfinite induction. In other words, we apply Zorn's Lemma. Consider the set M of functionals $\phi : Z \to \mathbb{R}$ such that $Y \subset Z \subset X$, ϕ positive, and $\phi|Y = \phi_0$. Order it by

$$\phi \geq \phi' \text{ if and only if } Z \subset Z' \text{ and } \phi'|Z = \phi.$$

If $(\phi_a)_{a\in A}$ is totally ordered, then $Z = \cup_a Z_a$ is a subspace, and $\phi = \phi_a$ on Z_a for all a is a functional that is larger that all ϕ_a. Hence the conditions of Zorn's Lemma are satisfied. If $\phi : Z \to \mathbb{R}$ is a maximal functional, then $Z = X$, for if $x \in X$ but not in Z, then we can extend ϕ to $Z + \mathbb{R}x$ as seen above. □

Theorem 2.1.2. (F. Riesz) Let $\phi : C([0,1]) \to \mathbb{R}$ be a positive linear functional. Then there is a unique positive measure σ on $[0,1]$ such that

$$\phi(f) = \int_0^1 f(t)d\sigma(t). \tag{2.1.2}$$

Proof. We use the theorem of M. Riesz. Let $B([0,1])$ be the space of bounded functions on $[0,1]$. Set $X = B([0,1])$ and $Y = C([0,1])$. The conditions of Theorem 2.1.1 are satisfied, because every bounded function is the difference between a continuous bounded function and a positive function. Hence there is a positive linear functional $\psi : X \to \mathbb{R}$ such that $\psi|C([0,1]) = \phi$. Define the monotone increasing function $F : [0,1] \to \mathbb{R}$ such that

$$F(t) = \psi(\chi_{[0,t]}).$$

Let $\sigma = dF$. To prove (2.1.2) consider an approximation of f by step functions

$$a\chi_{\{0\}} + \sum a_i\chi_{(x_i,x_{i+1}]} \leq f \leq a\chi_{\{0\}} + \sum (a_i + \epsilon)\chi_{(x_i,x_{i+1}]}.$$

Because ψ is positive, it preserves inequalities, hence

$$a\psi(\chi_{\{0\}}) + \sum a_i\psi(\chi_{(x_i,x_{i+1}]}) \leq \phi(f) \leq a\psi(\chi_{\{0\}}) + \sum a_i\psi(\chi_{(x_i,x_{i+1}]}) + \epsilon\phi(1).$$

This can be rewritten as

$$aF(0) + \sum a_i(F(x_{i+1}) - F(x_i)) \leq \phi(f) \leq aF(0) + \sum a_i(F(x_{i+1} - F(x_i)) + \epsilon\phi(1).$$

The conclusion follows. □

For those with more experience in measure theory, here is the general statement of this result.

Theorem 2.1.3. Let X be a compact space, in which the Borel sets are the σ-algebra generated by open sets. Let $\phi : C(X) \to \mathbb{R}$ be a positive linear functional. Then there is a unique regular (positive) measure σ on $\mathbb{R}$ such that

$$\phi(f) = \int_X f d\sigma. \tag{2.1.3}$$

Proof. The same proof works, mutatis mutandis. The measure is defined as

$$\sigma(A) = \psi(\chi_A),$$

where χ_A is the characteristic function of the Borel set A. □

Now we are in position to prove the Hamburger moment problem.

Theorem 2.1.4. (Hamburger) Let s_n, $n \geq 0$, be a sequence such that for all n, the matrix (2.1.1) is positive semidefinite. Then there is a regular positive finite measure σ on $\mathbb{R}$ such that for all $n \geq 0$, $t^n \in L^1(\sigma)$ and

$$s_n = \int_{-\infty}^{\infty} t^n d\sigma(t).$$

Proof. Denote by $\mathbb{R}[x]$ the real valued polynomial functions on $\mathbb{R}$ and by $C_c(\mathbb{R})$ the continuous functions with compact support. Consider

$$X = \mathbb{R}[x] + C_c(\mathbb{R}), \quad Y = \mathbb{R}[x],$$

and

$$C = \{f \in X \mid f(t) \geq 0 \text{ for all } t\}.$$

For a polynomial $u(t) = \sum_{n=0}^{N} a_n t^n$, let

$$\phi_0(u) = \sum_{n=0}^{N} a_n s_n.$$

Let us show that ϕ_0 is positive on C. We have $u \geq 0$ if and only if $u = p^2 + q^2$. If $\mathbf{p}$ and $\mathbf{q}$ are the vectors with coordinates the coefficients of p and q, then

$$\phi_0(u) = \phi_0(p^2) + \phi_0(q^2) = \mathbf{p}^T S_N \mathbf{p} + \mathbf{q}^T S_N \mathbf{q} \geq 0.$$

The conditions of Theorem 2.1.1 are verified. Then there is a linear positive functional $\phi : X \to \mathbb{R}$ such that $\phi|Y = \phi_0$. By Theorem 2.1.2, on every interval $[-m, m]$, $m \geq 1$ there is a measure σ_m such that if f is continuous with the support in $[-m, m]$, then $\phi(f) = \int_{-m}^{m} f(t) d\sigma_m(t)$. Uniqueness implies that for $m_1 > m_2$, $\sigma_{m_1}|[-m_2, m_2] = \sigma_{m_2}$. Hence we can define σ on $\mathbb{R}$ such that $\sigma|[-m, m] = \sigma_m$. Then for all $f \in C_c(\mathbb{R})$,

$$\phi(f) = \int_{-\infty}^{\infty} f(t) d\sigma(t).$$

The fact that σ is a finite measure is proved as follows. Given an interval $[-m, m]$, let f be compactly supported, such that

$$\chi_{[-m,m]} \leq f \leq 1.$$

Then

$$\sigma([-m,m]) = \phi(\chi_{[-m,m]}) \leq \phi(f) \leq \phi_0(1) = s_0.$$

Hence σ is finite.

Let us now show that

$$\phi(p) = \int_{-\infty}^{\infty} p(t)d\sigma(t).$$

If p is an even degree polynomial with positive dominant coefficient, then it can be approximated from below by compactly supported continuous functions, and so using the positivity of ϕ we conclude that for every such function f

$$\phi(p) \geq \phi(f) = \int_{-\infty}^{\infty} f(t)d\sigma(t).$$

By passing to the limit we find that

$$\phi(p) \geq \int_{-\infty}^{\infty} p(t)d\sigma(t).$$

Let q be a polynomial of even degree with dominant coefficient positive, whose degree is less than the degree of p. Then for every $a > 0$,

$$\phi(p - aq) \geq \int_{-\infty}^{\infty} (p - aq)(t)d\sigma(t).$$

Said differently

$$\phi(p) - \int_{-\infty}^{\infty} p(t)d\sigma(t) \geq a\left(\phi(q) - \int_{-\infty}^{\infty} q(t)d\sigma(t)\right).$$

This can only happen if

$$\phi(q) = \int_{-\infty}^{\infty} q(t)d\sigma(t).$$

Varying p and q we conclude that this is true for every q with even degree and with positive dominant coefficient. Since every polynomial can be written as the difference between two even degree polynomials with positive dominant coefficients, the property is true for all polynomials. □

2.2 The Riesz Representation Theorem for Hilbert spaces

Theorem 2.2.1. (The Riesz Representation Theorem) Let H be a Hilbert space and let $\phi : H \to \mathbb{C}$ be a continuous linear functional. Then there is $z \in H$ such that

$$\phi(x) = \langle x, z \rangle, \text{ for all } x \in H.$$

Proof. Let us assume that ϕ is not identically equal to zero, for otherwise we can choose $z = 0$.

Because ϕ is continuous, $\text{Ker}\phi = \phi^{-1}(0)$ is closed. Let $Y = \text{Ker}\phi^{\perp}$. Then Y is one dimensional, because if y_1, y_2 were linearly independent in Y, then $\phi(\phi(y_2)y_1 - \phi(y_1)y_2) = 0$, but $\phi(y_2)y_1 - \phi(y_1)y_2$ is a nonzero vector orthogonal to the kernel of ϕ.

Next, let y be a nonzero vector in Y, so that $\phi(y) \neq 0$. Replace y by $y' = y/\phi(y)$. Let $z = y'/\|y'\|^2$. Then

$$\phi(z) = 1/\|y'\|^2 = \langle z, z \rangle .$$

Every vector $x \in H$ can be written uniquely as $x = u + \alpha z$ with $u \in \text{Ker}\phi$ and α a scalar. Then

$$\begin{aligned}\phi(x) &= \phi(u + \alpha z) = \alpha\phi(z) = \alpha \langle z, z \rangle \\ &= \langle u + \alpha z, z \rangle = \langle x, z \rangle .\end{aligned}$$

□

Example. If $\phi : L^2(\mathbb{R}) \mapsto \mathbb{C}$ is a continuous linear functional, then there is an L^2 function g such that

$$\phi(f) = \int_{-\infty}^{\infty} f(t)\overline{g(t)}dt,$$

for all $f \in L^2(\mathbb{R})$.

Example. Consider the Hardy space $H^2(\mathbb{D})$, and the linear functionals $\phi_z(f) = f(z)$, $z \in \mathbb{D}$. Then for all z, ϕ_z is continuous, and so there is a function $K_z(w) \in H^2(\mathbb{D})$ such that

$$f(z) = \langle f, K_z \rangle .$$

The function $(z, w) \mapsto \overline{K_z(w)}$ is called the reproducting kernel of the Hardy space.

Example. Consider the Segal-Bargmann space $\mathcal{H}L^2(\mathbb{C}, \mu_\hbar)$. The linear functionals $\phi_z(f) = f(z)$, $z \in \mathbb{C}$ are continuous. So we can find $K_z(w) \in \mathcal{H}L^2(\mathbb{C}, \mu_\hbar)$ so that

$$f(z) = \int_{\mathbb{C}} f(w)\overline{K_z(w)}d\mu_\hbar.$$

Again, $(z, w) \mapsto \overline{K_z(w)}$ is called the reproducing kernel of the Segal-Bargmann space.

Remark 2.2.1. Using the Cauchy-Schwarz inequality, we see that

$$|\phi(x)| \leq \|z\|\|x\|.$$

In fact it can be seen that the continuity of ϕ is equivalent to the existence of an inequality of the form $|\phi(x)| \leq C\|x\|$ that holds for all x, where C is a fixed positive constant.

2.3 The Hahn-Banach Theorems

Theorem 2.3.1. (Hahn-Banach) Let X be a real vector space and let $p : X \to \mathbb{R}$ be a functional satisfying

$$p(x+y) \leq p(x) + p(y), \quad p(tx) = tp(x)$$

if $x, y \in X$, $t \geq 0$. Also, let Y be a subspace and let $\phi_0 : Y \to \mathbb{R}$ be a linear functional such that $\phi_0(y) \leq p(y)$ for all $y \in Y$. Then there is a linear functional $\phi : X \to \mathbb{R}$ such that $\phi|Y = \phi_0$ and $\phi(x) \leq p(x)$ for all $x \in X$.

Remark 2.3.1. The functional p can be a seminorm, or more generally, a Minkowski functional.

Proof. First choose $x_1 \in X$ such that $x_1 \notin Y$ and consider the space

$$Y_1 = \{y + tx_1 \,|\, y \in Y, t \in \mathbb{R}\}.$$

Because

$$\phi_0(y) + \phi_0(y') = \phi_0(y + y') \leq p(y + y') \leq p(y - x_1) + p(x_1 + y')$$

we have

$$\phi_0(y) - p(y - x_1) \leq p(y' + x_1) - \phi_0(y')$$

for all $y, y' \in Y$. Then there is $\alpha \in \mathbb{R}$ such that

$$\phi_0(y) - p(y - x_1) \leq \alpha \leq p(y' + x_1) - \phi_0(y')$$

for all $y, y' \in Y$. Then for all $y \in Y$,

$$\phi_0(y) - \alpha \leq p(y - x_1) \text{ and } \phi_0(y) + \alpha \leq p(y + x_1).$$

Define $\phi_1 : Y_1 \to \mathbb{R}$, by

$$\phi_1(y + tx_1) = \phi_0(y) + t\alpha.$$

Then ϕ_1 is linear and coincides with ϕ on Y. Also,

$$\begin{aligned}\phi_1(y + tx_1) = |t|\phi_1(y/|t| \pm x_1) &= |t|\left(\phi_0(y/|t|) \pm \alpha\right) \\ &\leq |t|p(y/|t| \pm x_1) = p(y + tx_1).\end{aligned}$$

To finish the proof, apply Zorn's lemma to the set of functionals $\phi : Z \to \mathbb{R}$, with $Y \subset Z \subset X$ and $\phi|Y = \phi_0$, $\phi(x) \leq p(x)$, ordered by $\phi < \phi'$ if the domain Z of ϕ is a subspace of the domain of Z' of ϕ' and $\phi'|Z = \phi$. □

Theorem 2.3.2. (Hahn-Banach) Suppose Y is a subspace of the vector space X, p is a seminorm on X, and ϕ_0 is a linear functional on Y such that $|\phi_0(y)| \leq p(y)$ for all $y \in Y$. Then there is a linear functional ϕ on X that extends ϕ_0 such that $|\phi(x)| \leq p(x)$ for all $x \in X$.

Proof. This is an easy consequence of the previous result. If we work with real vector spaces, then because ϕ is linear, by changing x to $-x$ if necessary, we get $|\phi(x)| \leq p(x)$ for all x.

If X is a complex vector space, note that a linear functional can be decomposed as $\phi = \text{Re}\phi + i\text{Im}\phi$. Then $\text{Re}\phi(ix) = -\text{Im}\phi(x)$, so $\phi(x) = \text{Re}\phi(x) + i\text{Re}\phi(ix)$. So the real part determines the functional.

Apply the theorem to $\text{Re}\phi_0$ to obtain $\text{Re}\phi$, and from it recover ϕ. Note that for every $x \in X$, there is $\alpha \in \mathbb{C}$, $|\alpha| = 1$ such that $\alpha\phi(x) = |\phi(x)|$. Then

$$|\phi(x)| = |\phi(\alpha x)| = \phi(\alpha x) = \text{Re}\phi(\alpha x) \leq p(\alpha x) = p(x).$$

The theorem is proved. □

Definition. Let X be a vector space, $A, A' \subset X$, $\phi : X \to \mathbb{R}$. We say that a nontrivial functional ϕ separates A from A' if $\phi(x) \leq \phi(x')$ for all $x \in A$ and $x' \in A'$.

Definition. Let B be a convex set, $x_0 \in B$. We say that x_0 is internal to B if $B - x_0 = \{b - x_0 \,|\, b \in B\}$ is absorbing.

Theorem 2.3.3. (Hahn-Banach) Let X be a vector space, A, A' convex subsets of X, $A \cap A' = \emptyset$ and A has an internal point. Then A and A' can be separated by a nontrivial linear functional.

Proof. Fix $a \in A$, $a' \in A'$ such that a is internal to A. Consider the set

$$B = \{x - x' - a + a' \,|\, x \in A, x' \in A'\}.$$

It is not hard to check that B is convex, it is also absorbing. Consider the Minkowski functional μ_B. Because A and A' are disjoint, $a' - a \notin B$. Hence $\mu_B(a' - a) \geq 1$.

Set $Y = \mathbb{R}(a' - a)$ and define $\phi_0(\lambda(a' - a)) = \lambda$. Then $\phi_0(y) \leq \mu_B(y)$ for all $y \in Y$. By the first Hahn-Banach theorem, there is $\phi : X \to \mathbb{R}$ such that $\phi(x) \leq \mu_B(x)$ for all $x \in X$, and also $\phi(a' - a) = \phi_0(a' - a) = 1$.

If $x \in B$, then $\mu_B(x) \leq 1$, so $\phi(x) \leq 1$. Hence if $x \in A$, $x' \in A'$, then

$$\phi(x - x' - a + a') \leq 1.$$

In other words

$$\phi(x - x') + \phi(a' - a) \leq 1, \text{ for all } x \in A, x' \in A'.$$

Since $\phi(a' - a) = 1$, it follows that $\phi(x - x') \leq 0$, so $\phi(x) \leq \phi(x')$ for all $x \in A$, $x' \in A'$. □

Theorem 2.3.4. (Hahn-Banach) Suppose A and B are disjoint, nonempty, convex sets in a locally convex topological vector space X over the real numbers.
a) If A is open then there is a continuous linear functional ϕ on X and $\gamma \in \mathbb{R}$ such that

$$\phi(x) < \inf_{y \in B} \phi(y) \text{ for all } x \in A.$$

b) If A is compact and B is closed then there is a continuous linear functional ϕ such that

$$\sup_{x \in A} \phi(x) < \inf_{y \in B} \phi(y).$$

Proof. We consider only the real case. Because A is open, every point of A is internal. So there is a linear functional ϕ and a real number α such that $A \subset \{x \,|\, \phi(x) \leq \alpha\}$ and $B \subset \{x \,|\, \phi(x) \geq \alpha\}$. Let x_0 be a point in A. Then for every $x \in A$, $\phi(x-x_0) = \phi(x)-\phi(x_0) \leq \alpha - \phi(x_0)$. So there is an open neighborhood $A - x_0$ of 0 such that $\phi(y) \leq \beta$ if $y \in A - x_0$, where $\beta = \alpha - \phi(x_0)$. Choosing $V \subset A - x_0$ a balanced neighborhood of 0, we conclude that $|\phi(y)| \leq \beta$ for all $y \in V$. But this is the condition that ϕ is continuous.

We claim that because A is open $\phi(x) < \alpha$ for all $x \in A$. If not, let x be such that $\phi(x) = \alpha$. Consider a neighborhood $V \subset$ of x such that $V - x$ is a balanced convex neighborhood of 0. Then for every $y \in V$ there is $z \in V$ such that x is the midpoint of the segment yz. We have $\phi(x) = \frac{1}{2}\phi(y) + \frac{1}{2}\phi(z)$, and because $\phi(y)$ and $\phi(z)$ are both less than or equal to α, $\phi(y) = \phi(z) = \phi(x) = \alpha$. Hence ϕ is constant in a neighborhood of x. Consequently ϕ is constant in a neighborhood of 0, and because the neighborhoods of 0 are absorbing, it is constant everywhere. This is impossible. Hence a) is true.

For b) we use Proposition 1.3.3 to conclude that there is an open set U such that $(A + U) \cap (B + U) = \emptyset$. By shrinking, we can make U balanced and convex. Then $A + U$ and $B + U$ are open and convex. Now use part a) to construct a continuous linear functional that separates $A + U$ from $B + U$. Let $x_0 \in U\backslash\{0\}$ such that $\phi(x_0) = \gamma > 0$. Then

$$\sup_{x\in A} \phi(x) + \gamma \leq \sup_{x\in A+U} \phi(x), \quad \inf_{y\in B} \phi(y) - \gamma \geq \inf_{y\in B+U} \phi(y).$$

The conclusion follows. □

Here is a practical application of the Hahn-Banach Theorem.

Theorem 2.3.5. (Farkas Lemma) Let A be an $m \times n$ matrix with real entries and let $b \in \mathbb{R}^n$ be a vector. Then exactly one of the following two situations holds
(i) There exists $x \geq 0$ such that $Ax = b$.
(ii) There exists y such that $A^T y \geq 0$ and $y^T b < 0$.

For a vector v, $v \geq 0$ means that all its entries are nonnegative.

Proof. Both outcomes cannot happen simultaneously because such x, y would then satisfy

$$0 \leq (A^T y)^T x = y^T Ax = y^T b < 0.$$

Let $C = \{Ax \,|\, x \geq 0\} \in \mathbb{R}^m$. If $b \in C$, then (i) holds. Otherwise, C is closed; consider the compact set $K = \{b\}$. Then $C \cap K = \emptyset$, so we can apply the second part of the fourth version of the Hahn-Banach theorem to conclude that there is a linear functional $\phi : \mathbb{R}^m \to \mathbb{R}$ and a real number α such that $\phi(b) < \alpha$ and $\phi(a) \geq \alpha$ for all $a \in C$. But $\mathbb{R}^m$ is a Hilbert space, so we can apply the Riesz representation theorem to conclude that there is y such that $\phi(x) = y^T x$ for all $x \in \mathbb{R}^m$. Thus

$$y^T b < \alpha \text{ and } y^T Ax \geq \alpha \text{ for all } x \geq 0.$$

But now since the zero vector is in C, α is necessarily nonpositive, in particular $y^T b < 0$. On the other hand, if $(y^T A)x \geq \alpha$ for all $x \geq 0$, then all entries of $y^T A$ have to be nonnegative, for if the kth entry is negative, we can take x to have all entries equal to 0 but the kth, and let this entry go to infinity (which would then make $y^T Ax$ go to negative infinity). Hence $y^T A \geq 0$, and so also $A^T y \geq 0$. Done. □

2.4 A few results about convex sets

For a subset A of a vector space X, we denote by $\mathrm{Ext}(A)$ the extremal points of A, namely the points $x \in A$ which cannot be written as $x = ty + (1-t)z$ with $y, z \in A \backslash \{x\}$, $t \in (0,1)$. This definition can be extended from points to sets by saying that a subset B of A is extremal if for $x, y \in A$ and $t \in (0,1)$ such that $tx + (1-t)y \in B$, it automatically follows that $x, y \in B$. Note that a point x is extremal if an only if $\{x\}$ is an extremal set.

Also, for a subset A of the vector space X, we denote by $\mathrm{co}(A)$ the convex hull of A, namely the convex set consisting of all points of the form $tx + (1-t)y$ where $x, y \in A$ and $t \in [0,1]$.

Theorem 2.4.1. (Krein-Milman) Suppose X is a locally convex topological vector space, and let K be a subset of X that is compact and convex. Then

$$K = \overline{\mathrm{co}(\mathrm{Ext}(K))}.$$

Proof. Let us define the family of sets

$$\mathcal{F} = \{K' \subset K \mid K' : \text{ closed, convex, nonempty, and extremal in } K\}.$$

If $\mathcal{G}$ is a subfamily that is totally ordered by inclusion, then because K is compact,

$$K_0 = \cap_{K' \in \mathcal{G}} K' \neq \emptyset.$$

It is not hard to see that K_0 is also extremal. Hence we are in the conditions of Zorn's Lemma. We deduce that $\mathcal{F}$ has minimal elements.

Let K_m be a minimal element; we claim it is a singleton. Arguing by contradiction, let us assume that K_m has two distinct points. By the Hahn-Banach Theorem there is a continuous linear functional $\phi : X \to \mathbb{R}$ such that $\phi(x) \neq \phi(y)$. Let

$$\alpha = \max_{x \in K_m} \phi(x).$$

Define

$$K_1 = \{y \in K_0 \mid \phi(y) = \alpha\}.$$

Then K_1 is a nonempty extremal subset of K_m and consequently an extremal subset of K. It is also compact and convex, which contradicts the minimality of K_m. Hence K_m contains only one point. This proves

$$\mathrm{Ext}(K) \neq \emptyset.$$

Let $K_e = \overline{\mathrm{co}(\mathrm{Ext}(K))}$. Note that K_e is compact. Assume $K_e \neq K$. Then there is $x \in K \backslash K_e$. By the Hahn-Banach theorem, there is a continuous linear functional $\phi : X \to \mathbb{R}$ such that $\max_{y \in K_e} \phi(y) < \phi(x)$. Set

$$K_2 = \{z \in K \mid \phi(z) = \max_{y \in K} \phi(y)\}.$$

It is not hard to see that K_2 is extremal in K. Hence $K_2 \in \mathcal{F}$. Applying again Zorn's Lemma, we conclude that there is a minimal extremal set in K that is included in K_2. This minimal set is a singleton. So there is $y \in K_2 \cap \mathrm{Ext}(K)$. But K_2 is disjoint from $\mathrm{Ext}(K)$, which is a contradiction. The conclusion follows. □

Theorem 2.4.2. (Milman) Let X be a locally convex topological vector space and let K be a compact set such that $\overline{\text{co}(K)}$ is also compact. Then every extreme point of $\overline{\text{co}(K)}$ lies in K.

Proof. Assume that some extreme point $x_0 \in \overline{\text{co}(K)}$ is not in K. Then there is a convex balanced neighborhood V of 0 in X such that

$$(x_0 + \overline{V}) \cap K = \emptyset,$$

which is equivalent to

$$x_0 \notin K + \overline{V}.$$

Choose $x_1, x_2, \ldots, x_n \in K$ such that $K \subset \cup_{j=1}^n (x_j + V)$. Each of the sets

$$K_j = \overline{\text{co}(K \cap (x_j + V))}, \quad 1 \le j \le n$$

is compact and convex. Note in particular that, because V is convex,

$$K_j \subset \overline{x_j + V} = x_j + \overline{V}.$$

Also $K \subset K_1 \cup \cdots \cup K_n$. We claim that $\text{co}(K_1 \cup K_2 \cup \cdots \cup K_n)$ is compact as well. To prove this, let $\sigma = \{(t_1, t_2, \ldots, t_n) \in [0,1]^n \mid \sum_j t_j = 1\}$, and consider the function $f : \sigma \times K_1 \times K_2 \ldots \times K_n \to X$,

$$f(t_1, \ldots, t_n, k_1, \ldots, k_n) = \sum_j t_j k_j.$$

Let C be the image of f. Note that $C \subset \text{co}(K_1 \cup K_2 \cdots \cup K_n)$. Clearly C is compact, being the image of a compact set, and is also convex. It contains each K_j, and hence it coincides with $\text{co}(K_1 \cup K_2 \cdots \cup K_n)$.

So

$$\overline{\text{co}(K)} \subset \text{co}(K_1 \cup \cdots \cup K_n).$$

The opposite inclusion also holds, because $K_j \subset \overline{\text{co}(K)}$ for every j. Hence

$$\overline{\text{co}(K)} = \text{co}(K_1 \cup \cdots \cup K_n).$$

In particular,

$$x_0 = t_1 y_1 + t_2 y_2 + \cdots + t_n y_n,$$

where $y_j \in K_j$ and $t_j \ge 0$, $\sum t_j = 1$. But x_0 is extremal in $\overline{\text{co}(K)}$, so x_0 coincides with one of the y_j. Thus for some j,

$$x_0 \in K_j \subset x_j + \overline{V} \subset K + \overline{V},$$

a contradiction. The conclusion follows. $\square$

We will show an application of these results. Given a convex subset K of a vector space X, and a vector space Y, a map $T : K \to Y$ is called affine if for every $x, y \in K$ and $t \in [0, 1]$,

$$T(tx + (1-t)y) = tT(x) + (1-t)T(y).$$

The result is about groups of affine transformations from X into itself. If X has a topological vector space structure, a group G of affine transformations of K is called equicontinuous if for every neighborhood V of 0 in X, there is a neighborhood U of 0 such that $T(x) - T(y) \in V$ for every $x, y \in K$ such that $x - y \in U$ and for every $T \in G$.

Theorem 2.4.3. (Kakutani's Fixed Point Theorem) Suppose that K is a nonempty compact convex subset of a locally convex topological vector space X and that G is an equicontinuous group of affine transformations taking K into itself. Then there is $x_0 \in K$ such that $T(x_0) = x_0$ for all $T \in G$.

Proof. Let

$$\mathcal{F} = \{K' \subset K \mid K' : \text{ nonempty, compact, convex }, T(K') \subset K' \text{ for all } T \in G\}.$$

Note that $K \in \mathcal{F}$, so this family is nonempty. Order $\mathcal{F}$ by inclusion and note that if $\mathcal{G}$ is a subfamily that is totally ordered, then because K is compact,

$$K_0 = \cap_{K' \in \mathcal{G}} K' \neq \emptyset.$$

Clearly $T(K_0) \subset K_0$, so the conditions of Zorn's lemma are satisfied. It follows that $\mathcal{F}$ has minimal elements. Let K_0 be such a minimal element. We claim that it is a singleton.

Assume, to the contrary, that K_0 contains x, y with $x \neq y$. Let V be a neighborhood of 0 such that $x - y \notin V$, and let U be the neighborhood of 0 associated to V by the definition of equicontinuity. Then for every $T \in G$, $T(x) - T(y) \notin U$, for else, because $T^{-1} \in G$,

$$x - y = T^{-1}(T(x)) - T^{-1}(T(y)) \in V.$$

Set $z = \frac{1}{2}(x + y)$. Then $z \in K_0$. Let

$$G(z) = \{T(z) \mid T \in G\}.$$

Then $G(z)$ is G-invariant, hence so is its closure $K_1 = \overline{G(z)}$. Consequently, $\overline{\text{co}(K_1)}$ is a G-invariant, compact convex subset of K_0. The minimality of K_0 implies

$$K_0 = \overline{\text{co}(K_1)}.$$

By the Krein-Milman Theorem (Theorem 2.4.1), K_0 has extremal points. Applying Milman's Theorem (Theorem 2.4.2), we deduce that every extremal point of K_0 lies in K_1. Let x_0 be an extremal point.

Consider the set

$$S = \{(Tz, Tx, Ty) \mid T \in G\} \subset K_0 \times K_0 \times K_0.$$

Since $x_0 \in K_1 = \overline{G(z)}$, and $K_0 \times K_0$ is compact, there is a point $(x_1, y_1) \in K_0 \times K_0$ such that $(x_0, x_1, y_1) \in \overline{S}$. Indeed, if this were not true, then every $(x_1, y_1) \in K_0 \times K_0$ would have a neighborhood $W_{(x_1,y_1)}$ for which there would exist a neighborhood $V_{(x_1,y_1)}$ of x_0 such that $V_{(x_1,y_1)} \times W_{(x_1,y_1)} \cap S = \emptyset$. Then $K_0 \times K_0$ is covered by finitely many of the $W_{(x_1,y_1)}$ and the intersection of the corresponding $V_{(x_1,y_1)}$'s is a neighborhood of p that does not intersect K_1.

Because $2Tz = Tx + Ty$ for all T, we get $2x_0 = x_1 + y_1$, hence $x_0 = x_1 = y_1$, because x_0 is an extremal point. But $Tx - Ty \notin V$ for all $T \in G$, hence $x_1 - y_1 \notin V$, and so $x_1 \neq y_1$. This is a contradiction, which proves our initial assumption was false, and the conclusion follows. □

2.5 The dual of a topological vector space

2.5.1 The weak*-topology

Let X be a topological vector space.

Definition. The space X^* of *continuous linear functionals* on X is called the *dual* of X.

X^* is a vector space. We endow it with the weak* topology, in which a system of neighborhoods of the origin is given by

$$V(x_1, x_2, \ldots, x_n, \epsilon) = \{\phi \in X^* \,|\, |\phi(x_j)| < \epsilon, \quad j = 1, 2, \ldots, n\},$$

where $x_1, x_2, \ldots, x_n$ range in X and $\epsilon > 0$.

Proposition 2.5.1. The space X^* endowed with the weak* topology is a locally convex topological vector space.

The Hahn-Banach Theorem implies automatically that the weak*-continuous linear functionals on X^* separate the points of this space. Each point $x \in X$ defines a weak*-continuous linear functional x^* on X^* defined by

$$x^*(\phi) = \phi(x).$$

In fact we have the following result.

Theorem 2.5.1. Every weak*-continuous linear functional on X^* is of the form x^* for some $x \in X$. Hence $(X^*)^* = X$. [1]

Proof. Assume that ϕ^* is a weak*-continuous linear functional on X^*. Then $|\phi^*(\phi)| < 1$ for all ϕ in some set $V(x_1, x_2, \ldots, x_n, \epsilon)$. This means that there is a constant C such that $|\phi^*(\phi)| \leq C \max_j |x_j^*(\phi)|$ for all $\phi \in X^*$.

Let N be the set on which $x_j^* = 0$, $j = 1, 2, \ldots, n$. Then ϕ^* is zero on N, so we can factor X^* by N so that we are in the finite dimensional situation. We can identify X^*/N with $\text{Span}(x_1, x_2, \ldots, x_n)^*$. In X^*/N, $\phi^* = \sum_j \alpha_j x_j^*$, and so this must be the case in X as well. Hence

$$\phi^* = (\sum_j \alpha_j x_j)^*,$$

and we are done. □

[1] It is important that on X^* we have the weak* topology, if we put a different topology on it, $(X^*)^*$ might not equal X.

We should point out that the weak* toplogy is the *coarsest* topology in which all functionals of the form x^* are continuous. Indeed, if we require that x^* is continuous, then the sets $V(x,\epsilon)$ are open for every ϵ. Certainly intersections of such sets must also be open, thus the sets of the form $V(x_1,x_2,...,x_n,\epsilon)$ are open. And once we consider the topology generated by these sets, the functionals of the form x^* are continuous.

There is another way to look at this topology. We can view functionals on X^* simply as functions, and functions as elements in the catesian product $\mathbb{R}^{X^*}$. If we endow the latter with the product topology, then the weak* topology is the induced topology.

Let us recall some facts about the product topology. If A_α, $\alpha \in I$, is a family of sets, then the cartesian product $A = \prod_\alpha A_\alpha$ together with the projection maps $\pi_\alpha : A \to A_\alpha$ is caracterized by the following property: for every set X and family of functions $f_\alpha : X \to A_\alpha$ there exists a unique function $f : X \to A$ such that $\pi_\alpha \circ f = f_\alpha$.

If we require A_α to be topological spaces, then A itself has a unique topology that makes every π_α continuous, and moreover, for every topological space X and continous maps $f_\alpha : X \to A_\alpha$, there exists a unique continuous function $f : X \to A$ such that $\pi_\alpha \circ f = f_\alpha$. This topology on $A = \prod_\alpha A_\alpha$ is called the *product topology*. It is the coarsest topology for which all the maps π_α are continuous. The fact that it is defined by this categorical construction makes it the most natural topology.

The product topology is generated by sets of the form $U_\alpha \times \prod_{\beta\neq\alpha} A_\beta$, where U_α is an open set in A_α. Its open sets are arbitrary unions of sets of the form $U_{\alpha_1} \times \cdots \times U_{\alpha_n} \times \prod_{\beta\neq\alpha_j} A_\beta$. An important result in topology is Tychonoff's Theorem, which states that a product of compact sets is compact. We will use this theorem below.

Theorem 2.5.2. (Banach-Alaoglu) Let X be a topological vector space, V a neighborhood of 0, and

$$K = \{\phi \in X^* \,|\, |\phi(x)| \leq 1, \text{ for all } x \in V\}.$$

Then K is compact in the weak* topology.

Proof. For every $x \in V$ define

$$K_x = \{\lambda \in \mathbb{C} \,|\, |\lambda| \leq 1\}.$$

The set

$$\prod_{x\in V} K_x$$

is compact in the product topology, by Tychonoff's Theorem. Define

$$\Phi : K \to \prod_{x\in V} K_x, \quad \Phi(\phi) = (\phi(x))_{x\in V}.$$

We will show

(1) $\Phi(K)$ is closed.

(2) $\Phi : K \to \Phi(K)$ is a homeomorphism.

For (1) assume that $(a_x)_{x\in V}$ is in $\overline{\Phi(K)}$. Define $\phi(x) = \frac{1}{t}a_{tx}$ for t such that $tx \in V$, Approximating $(a_x)_{x\in V}$ with linear functionals $\phi_n \in \Phi(K)$, $n = 1, 2, \ldots, n$. we have

$$\phi_n(\alpha x + \beta y) = \alpha\phi_n(x) + \beta\phi_n(y).$$

For n large enough, $\phi_n(\alpha x+\beta y)$ approximates well $\phi(x)$, while $\phi_n(x)$ and $\phi_n(y)$ approximate $\phi(x)$ and $\phi(y)$. By passing to the limit $n \to \infty$ we obtain that ϕ is linear.

Also for $x \in V$, $|\phi_n(x)| \leq 1$, and again by passing to the limit, $|\phi(x)| \leq 1$. This implies the continuity of ϕ, as well as the fact that it lies in $\Phi(K)$. This proves (1).

For (2), note that Φ is one-to-one, hence it is an inclusion. Moreover, as explained above, the weak* topology was chosen so that it coincides with the topology induced by the product topology. Hence the conclusion. □

2.5.2 The dual of a normed vector space

If X is a normed vector space, then X^* is also a normed space with the norm

$$\|\phi\| = \sup\{|\phi(x)| \mid \|x\| \leq 1\}.$$

Proposition 2.5.2. The dual of a normed space is a Banach space.

Proof. The only difficult part is to show that X^* is complete. Let ϕ_n, $n \geq 1$, be a Cauchy sequence in X^*. Then $\phi_n(x)$ is Cauchy for every x, hence convergent. So we can define $\phi(x) = \lim_{n\to\infty}\phi_n(x)$. It is not hard to check that ϕ is linear. On the other hand, because ϕ_n is Cauchy, $\|\phi_n\|$ is Cauchy as well, by the triangle inequality $(|\|\phi_n\| - \|\phi_m\|| < \|\phi_n - \phi_m\|)$. For a given x, if we choose n large enough then

$$|\phi(x) - \phi_n(x)| < \|x\|,$$

so

$$|\phi(x)| < (\|\phi_n\| + 1)\|x\|.$$

Because $\|\phi_n\|$, $n \geq 1$, is a bounded sequence (being Cauchy), it follows that ϕ is a bounded linear functional, and we are done. □

So X^* has two topologies the one induced by the norm, and the weak* topology. It is not hard to check that the second is coarser than the first. The two topologies coincide only in the finite dimensional case. Here is an example of the dual of a Banach space.

Theorem 2.5.3. Let $p \in [1, \infty)$. Then $(L^p([0, 1]))^* = L^q([0, 1])$, where q satisfies

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

A function $g \in L^q([0, 1])$ defines a functional by

$$\phi_g(f) = \int_0^1 f(x)g(x)dx.$$

Proof. Note that every $g \in L^q([0,1])$ defines a continuous linear functional by the above formula because of Hölder's inequality:

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Moreover, it is not hard to see that if g_1 and g_2 define the same functional then $g_1 = g_2$ almost everywhere. This follows from the fact that if

$$\int_0^1 f(x)[g_1(x) - g_2(x)]dx = 0$$

for all f then $g_1 - g_2 = 0$ almost everywhere.

Let us show that every continuous linear functional ϕ is of this form. The map

$$\mu_\phi : A \mapsto \phi(\chi_A)$$

is a measure on the Lebesgue measurable sets in $[0,1]$. Note that μ_ϕ is absolutely continuous with respect to the Lebesgue measure, since if the Lebesgue measure of A is zero, then χ_A is the zero vector in $L^p([0,1])$, and so $\mu_\phi(A) = \phi(\chi_A) = 0$.

Using the Radon-Nikodym Theorem, we deduce that there is a function $g \in L^1([0,1])$ such that

$$\phi(\chi_A) = \int_0^1 \chi_A(x) g(x) dx. \tag{2.5.1}$$

Case 1. $p = 1$. We have

$$\left| \int_A g(x)dx \right| = |\phi(\chi_A)| \leq \|\phi\| \|\chi_A\|_1 = \|\phi\| m(A),$$

where $m(A)$ is the Lebesgue measure of A. So $|g| \leq \|\phi\|$ almost everywhere, showing that $g \in L^\infty([0,1])$.

Case 2. $p > 1$. Looking at (2.5.1) and approximating the functions in $L^p([0,1])$ by step functions, and using the continuity of both the left-hand side on $L^p([0,1])$ and of the right-hand side on $L^\infty([0,1]) \subset L^p([0,1])$, we deduce that

$$\phi(f) = \int_0^1 f(x)g(x)dx \text{ for all } f \in L^p([0,1]).$$

We want to show that if $\int_0^1 f(x)g(x)dx$ is finite for all $f \in L^p([0,1])$, then $g \in L^q([0,1])$. By multiplying f by $|g|/g$, we can make g be positive, so let us consider just this case.

Let h be a step function that approximates g^q from below, $h \geq 0$. Consider

$$\phi(h^{1/p}) = \int_0^1 g(x)h^{1/p}(x)dx \geq \int_0^1 h^{1/q}(x)h^{1/p}(x)dx = \int_0^1 h(x)dx.$$

By the continuity of ϕ, this inequality forces

$$\int_0^1 h(x)dx \leq \|\phi\| \|h^{1/p}\|_p. \tag{2.5.2}$$

We also have

$$\|h^{1/p}\|_p = \left[\int_0^1 h(x)dx\right]^{1/p},$$

so from (2.5.2) we get

$$\int_0^1 h(x)dx \leq \|\phi\| \left(\int_0^1 h(x)dx\right)^{1/p}.$$

Dividing through we get

$$\left(\int_0^1 h(x)dx\right)^{1/q} \leq \|\phi\|.$$

Passing to the limit with $h \to g^q$, we obtain $\|g\|_q \leq \|\phi\|$, as desired. □

Here is this result in full generality.

Theorem 2.5.4. Let (X, μ) be a measure space, and let $p \in [1, \infty)$ and q such that $1/p + 1/q = 1$. Then $(L^p(X))^* = L^q(X)$, where $g \in L^q(X)$ defines the functional

$$\phi(f) = \int_X fg d\mu.$$

Moreover $\|\phi\| = \|g\|_q$.

Theorem 2.5.5. $(C([0,1]))^*$ is the set of finite complex valued measures on $[0,1]$.

Proof. Each finite measure μ defines a continuous linear functional by

$$\phi(f) = \int_0^1 f(t)d\mu.$$

Let us prove conversely, that every linear functional is of this form. For every complex continuous linear functional ϕ, we have $\phi = \mathrm{Re}\phi + i\mathrm{Im}\phi$ where the real and the imaginary part are themselves continuous. So we reduce the problem to real functionals. We show that each such functional is the difference between two positive functionals, and then apply the Riesz Representation Theorem.

For $f \geq 0$, set

$$\phi^+(f) = \sup\{\phi(g) \,|\, g \in C([0,1]), 0 \leq g \leq f\}.$$

Because ϕ is continuous, hence bounded, ϕ^+ takes finite values. Since $g = 0 \leq f$, and $\phi(0) = 0$, we have that ϕ^+ is positive.

It is clear that $\phi^+(cf) = c\phi^+(f)$, for $c \geq 0, f \geq 0$.

Also, for $f_1, f_2 \geq 0$, $\phi^+(f_1 + f_2) \geq \phi^+(f_1) + \phi^+(f_2)$ because we can use for $f_1 + f_2$ the function $g_1 + g_2$ with $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$. On the other hand, if $f = f_1 + f_2$, and

$g \leq f$, set $g_1 = \max(g - f_2, 0)$ and $g_2 = \min(g, f_2)$. Then $g_1 + g_2 = g$, and $0 \leq g_j \leq f_j$, $j = 1, 2$. Hence

$$\phi(g) = \phi(g_1) + \phi(g_2) \leq \phi^+(f_1) + \phi^+(f_2).$$

Consequently $\phi^+(f) = \phi^+(f_1 + f_2) \leq \phi^+(f_1) + \phi^+(f_2)$. Therefore we must have equality.

For arbitrary f, write $f = f_1 - f_2$, where $f_1, f_2 \geq 0$, and define $\phi^+(f) = \phi^+(f_1) - \phi^+(f_2)$. It is not hard to see that ϕ^+ is well defined, linear, and positive. Also $\phi^+ - \phi$ is a linear positive functional. We have

$$\phi = \phi^+ - (\phi^+ - \phi),$$

and the claim is proved. We can therefore write every continuous complex linear functional as

$$\phi = \phi_1 - \phi_2 + i(\phi_3 - \phi_4),$$

where ϕ_j, $j = 1, 2, 3, 4$ are positive. Each of these is given by a positive measure μ_j, by the Riesz representation Theorem, so ϕ is given by the complex measure

$$\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4).$$

□

Remark 2.5.1. Using the general form of the Riesz Representation Theorem, we see that $[0, 1]$ can be replaced by any compact space.

Theorem 2.5.6. (Banach-Alaoglu) Let X be a normed vector space. Then the closed unit ball in X^* is weak*-compact.

Here are some applications.

Proposition 2.5.3. Place a number from the interval $[0, 1]$ at each node of the lattice $\mathbb{Z}^2$ such that the number at each node is the average of the four numbers at the closest nodes. Then all numbers are equal.

Proof. Consider the Banach space $L^\infty(\mathbb{Z}^2)$. Let K be the set of elements in $L^\infty(\mathbb{Z}^2)$ satisfying the condition from the statement. Then K is a weak*-closed subset of the unit ball; by applying the Banach-Alaoglu theorem we deduce that it is weak*-compact. It is also convex. By the Krein-Milman Theorem, $K = \overline{\text{co}(\text{Ext}(K))}$. Let $f : \mathbb{Z}^2 \to [0, 1]$ be an extremal point in K. Let L, U be the operators that shift up and left. Then $Uf, U^{-1}f, Lf, L^{-1}f$ are functions with the same property, and

$$f = \frac{1}{4}(Uf + U^{-1}f + Lf + L^{-1}f).$$

Because f is extremal, $f = Uf = U^{-1}f = Lf = L^{-1}f$, meaning that f is constant. In fact $f = 0$ or $f = 1$. The convex hull of the two extremal constant functions is the set of all constant functions with values in $[0, 1]$, this set is closed, so K consists only of constant functions. Done. □

Theorem 2.5.7. The space $L^1(\mathbb{R})$ is not the dual of any normed space.

Proof. If $L^1(\mathbb{R})$ were the dual of a normed space, then the Banach-Alaoglu Theorem implies that the closed unit ball in $L^1(\mathbb{R})$ is weak*-compact. By the Krein-Milman Theorem it has extreme points. But this is not true, since every function in the closed unit ball of L^1 can be written as the convex combination of two functions in the unit ball. □

Remark 2.5.2. Here we should compare with the case of L^p spaces. Just focus on positive functions. The convex combination of two norm 1 such functions in L^1 has also norm 1. But this is not true for L^p spaces.

Consider $\mathbb{C}([0,1])$, the Banach space of complex valued continuous functions on $[0,1]$, with the norm $\|f\| = \sup_{x\in[0,1]} |f(x)|$.

Theorem 2.5.8. (Stone-Weierstrass) Let $A \subset C([0,1])$ be a subalgebra with the following properties

(1) if $f \in A$ then $\bar{f} \in A$,

(2) the function identically equal to 1 is in A,

(3) A separates the points of $[0,1]$.

Then A is dense in $\mathbb{C}([0,1])$.

Proof. (de Brange) We argue by contradiction, and assume that $\bar{A}$ is not dense, that is $\bar{A} \neq C([0,1])$. By replacing A with $\bar{A}$, we can assume that A is closed. Let

$$K = \{\phi \in C([0,1])^* \mid \|\phi\| \leq 1, \phi|A = 0\}.$$

By the Banach-Alaoglu Theorem, it is compact in the weak* topology. K is also convex, so by the Krein-Milman Theorem it has extremal points. Moreover, the Hahn-Banach Theorem implies that $K \neq \{0\}$ (because there is a functional that separates A from a point that does not belong to it), so then Krein-Milman implies that moreover there exist at least two extremal points. This means that there is an extremal functional $\phi \in K$ that is not identically equal to zero.

Because $C([0,1])^*$ is the space of finite complex measures (Theorem 2.5.5), ϕ is given by a measure μ. We claim that every function in A is constant on the support of μ. If this is so then because the functions in A separate points, the support of μ consists of just one point, so $\mu = c\delta_{x_0}$ for $x_0 \in [0,1]$ and $c \in \mathbb{C}$. Because $\mu|A = 0$, and $1 \in A$, we get that $c = 0$, a contradiction. Hence the conclusion.

Let us prove the claim. Suppose there is $f \in A$ not constant on the support of μ. We have $f = f_1 + if_2$, so $f_1 = (f + \overline{f})/2$ and $f_2 = (f - \overline{f})/2i$, and because $\overline{f} \in A$, $f_1, f_2 \in A$ as well. One of these is nonconstant, so we may assume that f is real valued. Replacing f by $(f + A)/B$ we may assume $0 \leq f \leq 1$. Define the measures μ_1 and μ_2 by

$$d\mu_1 = fd\mu, \quad d\mu_2 = (1-f)d\mu.$$

Then $\mu = \mu_1 + \mu_2$. Note that μ_1, μ_2 are both zero on A.

We have $\|\mu_1\| = \int_0^1 f d|\mu|$, $\|\mu_2\| = \int_0^1 (1-f) d|\mu|$. And also $\|\mu\| = \int d|\mu| = \|\mu_1\| + \|\mu_2\|$. Then

$$\mu = \frac{\|\mu_1\|}{\|\mu\|}\left(\frac{\|\mu\|}{\|\mu_1\|}\mu_1\right) + \frac{\|\mu_2\|}{\|\mu\|}\left(\frac{\|\mu\|}{\|\mu_2\|}\mu_2\right).$$

Note that

$$\frac{\|\mu\|}{\|\mu_1\|}\mu_1, \frac{\|\mu\|}{\|\mu_2\|}\mu_2 \in K$$

Because μ is an extreme point, either μ_1 or μ_2 is zero. So f must be identically equal to 1, a contradiction. The claim is proved, and so is the theorem. □

Remark 2.5.3. Using the general form of the Riesz Representation Theorem, we see that $[0, 1]$ can be replaced by any compact space.

Chapter 3

Bounded Linear Operators

3.1 Continuous linear operators

3.1.1 The case of general topological vector spaces

We now start looking at continuous linear operators between topological vector spaces:

$$T : X \to Y.$$

Proposition 3.1.1. Let $T : X \to Y$ be a linear operator between topological vector spaces that is continuous at 0. Then T is continuous everywhere, moreover, for every open neighborhood V of 0 there is an open neighborhood U of 0 such that if $x - y \in U$ then $Tx - Ty \in V$.

Definition. A linear operator is called bounded if it maps bounded sets to bounded sets.

Proposition 3.1.2. A continuous linear operator is bounded.

Proof. Let $T : X \to Y$ be a continuous linear operator. Consider a bounded set $E \subset X$. Let also V be a neighborhood of 0 in Y. Because T is continous, there is a neighborhood U of 0 in X such that $T(U) \subset V$. Choose $\lambda \in \mathbb{C}\backslash\{0\}$ such that $\lambda E \subset U$. Then $T(\lambda E) = \lambda T(E) \subset V$. It follows that $T(E)$ is bounded. □

Definition. Let $T : X \to Y$ be a linear operator. The *kernel* of T is

$$\ker(T) = \{x \in X \,|\, Tx = 0\}.$$

The *range* or *image* of T is

$$\mathrm{im}(T) = \{y \in Y \,|\, \text{ there is } x \in X \text{ with } Tx = y\}.$$

Both $\ker(T)$ and $\mathrm{im}(T)$ are vector spaces. If T is a continuous linear operator between topological vector spaces, then $\ker(T)$ is closed. This is not necessarily true about $\mathrm{im}(T)$.

3.2 The three fundamental theorems

3.2.1 Baire category

A subset of a topological space is called nowhere dense if its closure has empty interior. Another way to say this is that its complement contains a dense open set.

Definition. A topological space is said to be of the *first category* if it is a countable union of nowhere dense subsets. Otherwise it is said to be of the *second category*.

Theorem 3.2.1. (Baire Category Theorem) A complete metric space is of the second category.

Proof. Assume by contradiction that X is a complete metric space of the first category. Write $X = \cup_{n=1}^{\infty} X_n$, with $X_n = X \backslash V_n$ where V_n is a dense open set. Define inductively the set of balls B_n such that $\overline{B_n} \subset V_n$ and $\overline{B_n} \subset B_{n-1}$, and the radius of B_n is less than half of the radius of B_{n-1}. The centers of the balls form a Cauchy sequence that converges to a point $x \in X$. This point x belongs to all B_n and hence it is in the complement of every X_n. But this is impossible because X is the union of all X_n. □

Corollary 3.2.1. If X is of second category and $X = \cup_{n=1}^{\infty} X_n$, then there is n such that $\overline{X_n}$ contains an open subset.

3.2.2 Bounded linear operators on Banach spaces

From now on we will focus just on continuous linear operators between Banach spaces.

Theorem 3.2.2. Let $T : X \to Y$ be a linear operator. Then T is continuous if and only if it is bounded.

Proof. We have seen that if T is continuous then T is bounded. Let us show the converse. Assume that T is bounded by is not continuous. Then there is a neighborhood $V \subset Y$ of 0 such that $T^{-1}(V)$ is not a neighborhood of 0. This means that there is a sequence $x_n \in X \backslash T^{-1}(V)$ such that $x_n \to 0$. So there is a sequence $x_n \to 0$ such that $Tx_n \notin V$, $n \geq 1$. We know that T is bounded, so $\{Tx_n\}_n$ is bounded. But now we can write $x_n = \alpha_n y_n$, where $\alpha_n \to 0$ and $y_n \to 0$. Then $\{Ty_n\}_n$ is still bounded, which implies that $Tx_n = \alpha_n T y_n \to 0$. This is a contradiction. Hence T is bounded.

Here is another way to prove this. Let V be a neighborhood of 0 in Y. We want to show that $T^{-1}(V)$ is a neighborhood of 0 in X. Consider the unit ball $X_1 \subset X$. Then $T(X_1)$ is bounded, so there is $t > 0$ such that $T(X_1) \subset tV$. But then $X_1 \subset T^{-1}(tV) = tT^{-1}(V)$, so $t^{-1} X_1 \subset T^{-1}(V)$. Thus $T^{-1}(V)$ is a neighborhood of 0, as desired. □

Definition. Let T be a bounded linear operator (which is the same as a continuous operator). The *norm* of T is

$$\|T\| = \sup\{\|Tx\| \,|\, \|x\| \leq 1\}.$$

Proposition 3.2.1. The set of continuous linear operators $T : X \to Y$ endowed with the operator norm is a Banach space.

Proof. Indeed, if T_n is Cauchy, then $T_n x$ is also Cauchy, because

$$\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|.$$

This latter sequence is convergent, and we define $Tx = \lim_{n\to\infty} T_n x$. Note that

$$\begin{aligned}
&\|T(\alpha x + \beta y) - \alpha Tx - \beta Ty\| \\
&\leq \|T(\alpha x + \beta y) - \alpha Tx - \beta Ty - [T_n(\alpha x + \beta y) - \alpha T_n x - \beta T_n y] + T_n(\alpha x + \beta y) - \alpha T_n x - \beta T_n y\| \\
&= \|(T - T_n)(\alpha x + \beta y) - \alpha(T - T_n)x - \beta(T - T_n)y\| \\
&\leq \|(T - T_n)(\alpha x + \beta y)\| + |\alpha|\|(T - T_n)x\| + |\beta|\|(T - T_n)y\|,
\end{aligned}$$

and the right-hand side goes to 0 when n goes to infinity. This implies that T is linear.

For a fixed $x \neq 0$, we can choose n such that $\|T - T_n(x)\| \leq \|x\|$ (as we can make this as small as possible).

$$\|Tx\| = \|(T - T_n)x + T_n x\| \leq \|(T - T_n)x\| + \|T_n x\| \leq \|x\| + \|T_n\|\|x\| \leq (\|T_n\| + 1)\|x\|,$$

so T is bounded. □

Proposition 3.2.2. If $T : X \to Y$ and $S : Y \to Z$ are bounded linear operators between Banach spaces, then $\|ST\| \leq \|S\|\|T\|$.

Proof. We have

$$\|STx\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|,$$

hence the conclusion. □

Here is an example of a bounded linear operator from P.D. Lax, *Functional Analysis*:

Example. The Laplace transform

$$L : L^2([0,\infty)) \to L^2([0,\infty)), \quad (Lf)(s) = \int_0^\infty f(t)e^{-st}dt$$

is a bounded linear operator.

We prove that it is bounded and compute its norm. We have

$$\begin{aligned}
|(Lf)(s)|^2 &= \left(\int_0^\infty f(t)e^{-st}dt\right)^2 = \left(\int_0^\infty (f(t)e^{-st/2}t^{1/4})(e^{-2st/2}t^{-1/4})dt\right)^2 \\
&\leq \int_0^\infty |f(t)|^2 e^{-st}t^{1/2}dt \int_0^\infty e^{-st}t^{-1/2}dt,
\end{aligned}$$

where for the last step we have applied the Cauchy-Schwarz inequality. By changing variables we can compute the second integral as

$$\begin{aligned}
&\int_0^\infty e^{-st}t^{-1/2}dt = s^{-1/2}\int_0^\infty e^{-u}u^{-1/2}du = s^{-1/2}\int_0^\infty e^{-x^2}x^{-1}2xdx \\
&= 2s^{-1/2}\int_0^\infty e^{-x^2}dx = s^{-1/2}\sqrt{\pi}.
\end{aligned}$$

We conclude that

$$|(Lf)(s)|^2 \leq s^{-1/2}\sqrt{\pi}\int_0^\infty |f(t)|^2 e^{-st} t^{1/2} dt.$$

Integrating with respect to s we obtain

$$\begin{aligned}
\|Lf\|^2 &= \int_0^\infty |(Lf)(s)|^2 ds \leq \sqrt{\pi}\int_0^\infty\int_0^\infty |f(t)|^2 e^{-st} t^{1/2} s^{-1/2} dt ds \\
&= \sqrt{\pi}\int_0^\infty\int_0^\infty |f(t)|^2 e^{-st} t^{1/2} s^{-1/2} ds dt = \sqrt{\pi}\int_0^\infty |f(t)|^2 t^{1/2}\int_0^\infty e^{-st} s^{-1/2} ds dt \\
&= \sqrt{\pi}\int_0^\infty |f(t)|^2 t^{1/2} t^{-1/2} dt = \sqrt{\pi}(\sqrt{\pi}\|f\|^2),
\end{aligned}$$

where for the last step we have used the integral computed above. Hence $\|Lf\| \leq \sqrt{\pi}\|f\|$. Thus $\|L\| \leq \sqrt{\pi}$.

In the above computation, the Cauchy-Schwarz inequality is the only place where an inequality occurred. We can get close to the equality case by choosing $f = 1/\sqrt{t}$ on an interval $[a, b]$ with a small and b large and zero outside of this interval (which ensures that f is in $L^2([0,\infty))$). Thus we can make $\|Lf\| \geq (\sqrt{\pi} - \epsilon)\|f\|$, for all ϵ, which then implies $\|L\| \geq \sqrt{\pi} - \epsilon$ for all ϵ. We conclude that $\|L\| = \sqrt{\pi}$.

Theorem 3.2.3. (Banach-Steinhaus) Let X be a Banach space, let Y be a normed space, and let $\mathcal{F}$ be a family of continuous operators from X to Y. Suppose that for all $x \in X$, $\sup_{T\in\mathcal{F}} \|Tx\| < \infty$. Then $\sup_{T\in\mathcal{F}} \|T\| < \infty$.

Proof. Let

$$X_n = \{x \in X \mid \|Tx\| \leq n \text{ for all } T \in \mathcal{F}\}.$$

These sets are convex and balanced. They are also closed, so by the Baire Category Theorem there is n such that the interior of X_n is nonempty. Because X_n is convex and balanced, its interior contains the origin. Hence there is a ball $B_{0,r}$ centered at origin such that $\|Tx\| \leq n$ for all $T \in \mathcal{F}$ and x with $\|x\| \leq r$. We have $\|T\| \leq n/r$ for all $T \in \mathcal{F}$, and the theorem is proved. □

Here is an application that I have learned from Hari Bercovici. We have

$$\frac{1}{x+1} = \sum_{n=1}^\infty (-1)^{n-1} x^{n-1}.$$

The left-hand side takes the value $1/2$ when $x = 1$, so it is natural to impose that the right-hand side converges to $1/2$. A way to do this is to consider the sequence $s_n = \sum_{k=1}^n (-1)^{k-1} x^{k-1}$ and then notice that

$$\frac{1}{n}(s_1 + s_2 + \cdots + s_n) \tag{3.2.1}$$

converges to the same limit as s_n when the latter converges (Cesàro), but moreover for $x = 1$ (3.2.1) converges to $1/2$.

Definition. A summation method associates to each convergent sequence s_n, $n \geq 1$ another convergent sequence σ_n, $n \geq 1$ such that

(1) $\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} s_n$;

(2) $\sigma_n = \sum_{k=1}^{\infty} \alpha_{nk} s_k$ for $n = 1, 2, \ldots$, where α_{nk} is an array of complex numbers that does not depend on s_n and defines the summation method.

An example of a summation method, introduced by Cesàro, is $\alpha_{nk} = 1/n$, $n = 1, 2, \ldots$, $1 \leq k \leq n$, and $\alpha_{nk} = 0$ otherwise.

Theorem 3.2.4. (Toeplitz) The array α_{nk}, $n, k \geq 1$, defines a summation method if and only if it satisfies the following three conditions

(1) $\lim_{n\to\infty} \alpha_{nk} = 0$, for all $k = 1, 2, \ldots$;

(2) $\lim_{n\to\infty} \sum_{k=1}^{\infty} \alpha_{nk} = 1$;

(3) $\sup_n \sum_{k=1}^{\infty} |\alpha_{nk}| < \infty$.

Proof. Let us prove that the three conditions are necessary. If $s_n = \delta_{nk}$ for some k, then $\sigma_n = \alpha_{nk}$. The fact that $s_n \to 0$ implies $\lim_{n\to\infty} \alpha_{nk} = 0$, hence (1).

If $s_n = 1$, $n \geq 1$, then $\sigma_n = \sum_{k=1}^{\infty} \alpha_{nk}$. Because $s_n \to 1$, it follows that $\lim_{n\to\infty} \sum_k \alpha_{nk} = 1$, hence (2).

For (3) we apply the Banach-Steinhaus Theorem. Denote by C_0 the Banach space of convergent sequences with the sup norm (i.e. continuous functions on $\mathbb{N} \cup \{\infty\}$ with the sup norm, where $\mathbb{N} \cup \{\infty\}$ is given the topology such that the map $f(x) = 1/x$ from it to $\mathbb{R}$ is a homeomorphism onto the image). Let α_k, $k \geq 1$, be a sequence such that $\sum_{k=1}^{\infty} \alpha_k x_k$ converges for every convergent sequence x_k, $k \geq 1$. We claim that $\sum_{k=1}^{\infty} |\alpha_k| < \infty$.

Indeed, if this is not the case, then choose $r_k > 0$ such that that $r_k \to 0$ and $\sum |\alpha_k| r_k = \infty$. The sequence $x_k = r_k \overline{\alpha_k}/|\alpha_k|$ converges to 0, but $\sum_k \alpha_k x_k = \sum |\alpha_k| r_k = \infty$, which is impossible. This proves our claim.

Additionally,

$$\sup_{(x_k)_k \in C_0, \|(x_k)_k\| \leq 1} \left| \sum_k \alpha_k x_k \right| = \sum_{k=1}^{\infty} |\alpha_k|.$$

The fact that the left-hand side does not exceed the right-hand side follows from the triangle inequality. On the other hand, if $x_k = \overline{\alpha_k}/|\alpha_k|$ for $1 \leq k \leq N$ and zero otherwise makes $\sum \alpha_k x_k = \sum_{k=1}^{N} |\alpha_k|$. Taking $N \to \infty$ we obtain that the right-hand side is less than or equal to the left-hand side. Hence the two are equal.

Define

$$\phi_n : C_0 \to \mathbb{C}, \quad \phi_n((s_k)_k) = \sum_{k=1}^{\infty} \alpha_{nk} s_k.$$

Then the above argument shows that $\phi_n \in (C_0)^*$ and

$$\|\phi_n\| = \sum_{k=1}^{\infty} |\alpha_{nk}|.$$

The sequence $\phi_n((s_k)_k)$, $n \geq 1$ is bounded for every convergent sequence $(s_k)_k$ (because we are in the presence of a summation method, and convergent sequences are bounded). Hence by the Banach-Steinhaus Theorem, $\|\phi_n\|$, $n \geq 1$, is bounded, which is (3).

Now let us check that the conditions are sufficient. Let $M = \sup_n \sum_{k=1}^{\infty} |\alpha_{nk}|$. Consider a sequence s_n converging to s. We want to show that σ_n converges to s as well. We compute

$$\begin{aligned} |\sigma_n - s| &\leq \left| \sum_{k=1}^{\infty} \alpha_{nk} s_k - \sum_{k=1}^{\infty} \alpha_{nk} s \right| + \left| \left(\sum_{k=1}^{\infty} \alpha_{nk} - 1 \right) s \right| \\ &\leq \sum_{k=1}^{\infty} |\alpha_{nk}||s_k - s| + \left| \sum_{k=1}^{\infty} \alpha_{nk} - 1 \right| |s| \\ &\leq \sum_{k=1}^{N} |\alpha_{nk}||s_k - s| + M \sup_{k \geq N+1} |s_k - s| + \left| \sum_{k=1}^{\infty} \alpha_{nk} - 1 \right| |s|. \end{aligned}$$

We obtain $\lim_{n\to\infty} |\sigma_n - s| = 0$, since each of the three terms converges to zero as $n \to \infty$. □

Theorem 3.2.5. (Open Mapping Theorem) Let $T : X \to Y$ be a surjective bounded linear operator between Banach spaces. Then T maps open sets to open sets.

Proof. It is enough to show that the set

$$A = \{Tx \mid \|x\| < 1\}$$

is a neighborhood of 0 in Y. We have

$$Y = \cup_{n=1}^{\infty} nA.$$

Because Y is of the second category (by the Baire Category Theorem), it follows that there is n such that $\overline{nA}$ has nonempty interior. Consequently $\overline{A}$ has nonempty interior.

But A is convex and balanced, because it is the image through a linear map of a convex and balanced set. Hence so it $\overline{A}$, and consequently $\overline{A}$ contains a neighborhood of 0. Let $\epsilon > 0$ be such that

$$\{y \mid \|y\| < \epsilon\} \subset \overline{A} = \overline{\{Tx \mid \|x\| < 1\}}.$$

We want to show that

$$\{y \mid \|y\| < \epsilon\} \subset \{Tx \mid \|x\| < 2\}.$$

Fix $y \in Y$, $\|y\| < \epsilon$ and fix $0 < \delta < 1$. Choose x_1 in the unit ball of X such that $\|y - Tx_1\| < \delta$. There is $x_2 \in X$, $\|x_2\| < \delta/\epsilon$ with $\|y - Tx_1 - Tx_2\| < \delta^2$, ..., there is x_n with $\|x_n\| < \delta^{n-1}/\epsilon$ and $\|y - Tx_1 - Tx_2 - \cdots - Tx_n\| < \delta^n$. Because X is Banach, there is a point $x \in X$ such that $x = \sum_{n=1}^{\infty} x_n$. Choosing δ small enough, we can ensure that $\|x\| < 2$. We have

$$\|y - Tx\| = \lim_{n\to\infty} \|y - Tx_1 - Tx_2 - \cdots - Tx_n\| = 0,$$

so $y = Tx$. The theorem is proved. □

Corollary 3.2.2. Let $T : X \to Y$ be a bounded linear operator between Banach spaces that is onto. Then there is a constant $C > 0$ such that for every $y \in Y$, there is $x \in X$ such that $Tx = y$ and $\|x\| \leq C\|y\|$.

Proof. The image of the unit ball of X is open in Y. Let $\delta > 0$ such that $\|y\| < \delta$ implies $y = Tx$ with $\|x\| < 1$. Then $C = 1/\delta$ does the job. □

Theorem 3.2.6. (Inverse Mapping Theorem) Let $T : X \to Y$ be an invertible bounded linear operator between Banach spaces. Then T^{-1} is also a bounded linear operator.

Proof. Because T maps open sets to open sets, the preimage of an open set through T^{-1} is open, showing that T^{-1} is continuous. □

Definition. Let $f : A \to B$ be a function. The graph of f is the set

$$\{(x, f(x)) \,|\, x \in A\} \subset A \times B.$$

We denote the graph of f by G_f.

Theorem 3.2.7. (Closed Graph Theorem) Let X and Y be Banach spaces and let $T : X \to Y$ be a linear operator such that the graph of T is closed in $X \times Y$ with the product topology. Then T is continuous.

Proof. The product space $X \times Y$ is a Banach space. The graph G_T is a linear subspace. By hypothesis it is closed, so it is a Banach subspace. Define

$$\pi_1 : G_T \to X, \quad \pi_1(x, Tx) = x$$

and

$$\pi_2 : G_T \to Y, \quad \pi_2(x, Tx) = Tx.$$

Both these operators are linear and continuous. The operator π_1 is invertible and bijective. By the Inverse Mapping Theorem (Theorem 3.2.6) its inverse is also continuous. We have $T = \pi_2 \circ \pi_1^{-1}$, and hence T is continuous. □

Here is an application found online in a note by Jesús Gil de Lamadrid:

Example. Let $T : L^2([0, 1]) \to L^2([0, 1])$ be a bounded linear operator so that if $f \in C([0, 1])$ then $Tf \in C([0, 1])$. Then the restriction of T to $C([0, 1])$ is a bounded operator.

Indeed, we have $\|f\|_2 < \|f\|_\infty$, so the topology induced on $C([0, 1])$ by the sup norm is finer than the one induced by the L^2 norm. Because T is continuous, its graph is closed in the product topology induced by the L^2 norm on each factor, and consequently it is closed in the product topology induced by the sup norm on each factor. Hence T is bounded on $C([0, 1])$ with the sup norm.

Definition. A linear operator $P : X \to X$ is called a projection if $P^2 = P$.

Proposition 3.2.3. Let X be a Banach space and let $P : X \to X$ be a projection. Then P is continuous if and only if both the kernel and the image of P are closed.

Proof. Assume that the kernel and the image are closed. Because $x = Px + (x - Px)$, and $P(x - Px) = Px - P^2x = 0$ every element in X is the sum of an element in $kerP$ and an element in imP. Moverover, if $x \in kerP \cap imP$, then $x = Py$, for some y, so $P^2y = Px = 0$. But $P^2y = Py = x$, so $x = 0$. It follows that

$$X = kerP \oplus imP.$$

Let us show that the graph of P is closed. Consider a sequence (x_n, Px_n), $n \geq 1$, that converges to (x, y); we want to show that $y = Px$. Because imP is closed, $y \in imP$. The sequence $x_n - Px_n$ converges to $x - y$. Because $x_n - Px_n \in kerP$, there is $z \in kerP$ such that $x_n - Px_n \to z$. So $x - y = z$. It follows that $x - y \in kerP$. We thus have $P(x-y) = Pz = 0$. But $Py = y$, so $P(x-y) = Px - y$. It follows that $Px = y$. From the Closed Graph Theorem it follows that P is continuous.

Conversely, if P is continuous, then $kerP = P^{-1}(0)$ is closed. Also, $imP = ker(1 - P)$, and $1 - P$ is also continuous. Hence imP is closed. □

Corollary 3.2.3. If P is a continuous projection then $X = kerP \oplus imP$ is a decomposition of X as a direct sum of two closed subspaces.

Example. Let A be a closed subset of $[0, 1]$, and let $C_A([0, 1])$ be the set of continuous functions that are zero on A. Then there is a closed subspace Y of $C([0, 1])$ such that

$$C([0, 1]) = C_A([0, 1]) \oplus Y.$$

Indeed, there is a bounded linear operator $T : C(A) \to C([0, 1])$ such that $Tg|A = g$ (the complement of A is a disjoint union of open intervals, and on such an interval (a, b) we can define $Tg(ta + (1 - t)b) = tg(a) + (1 - t)g(b)$). If $R : C([0, 1]) \to C(A)$ is the restriction operator, then $P = T \circ R$ is a projection. It is also continuous because T and R are continuous. Hence

$$C([0, 1]) = kerP \oplus imP = C_A([0, 1]) \oplus imP.$$

Set $Y = imP$.

The operator T defined in this example is called a simultaneous extension. It has been proved that such operators exist in more general situations (e.g. for compact spaces). The existence of such an operator is a stronger version of the Tietze Extension Theorem.

3.3 The adjoint of an operator between Banach spaces

Definition. Let $T : X \to Y$ be a bounded linear operator between Banach spaces. The adjoint of T, denoted by T^*, is the operator $T^* : Y^* \to X^*$ given by $T^* = \phi \circ T$.

Theorem 3.3.1. The operator T^* is linear and bounded, and $\|T^*\| = \|T\|$.

Proof. We have

$$(\alpha\phi_1 + \beta\phi_2) \circ T = \alpha\phi_1 \circ T + \beta\phi_2 \circ T$$

which shows that T^* is linear.

Lemma 3.3.1. If X is a Banach space and $x \in X$, then

$$\|x\| = \sup\{|\phi(x)| \mid \phi \in X^*, \|\phi\| \leq 1\}$$

Proof. We have $|\phi(x)| \leq \|\phi\|\|x\|$, so the left-hand side is greater than or equal to the right-hand side. For the converse inequality, define $\phi_0 : \mathbb{R}x \to C$, $\phi_0(tx) = t\|x\|$. Then $\|\phi_0\| = 1$. By the Hahn-Banach Theorem, there is a continuous linear functional $\phi : X \to \mathbb{C}$ such that $\|\phi\| = 1$, and $\phi(x_0) = \|x_0\|$. □

Returning to the theorem and using the lemma, we have

$$\begin{aligned}\|T\| &= \sup\{\|Tx\| \mid \|x\| \leq 1\} = \sup\{\|\phi(Tx) \mid \|x\| \leq 1, \|\phi\| \leq 1\} \\ &= \sup\{\|(T^*\phi)(x)\| \mid \|x\| \leq 1, \|\phi\| \leq 1\} = \sup\{\|T^*\phi\| \mid \|\phi\| \leq 1\} = \|T^*\|.\end{aligned}$$

□

Example. Let $X = \mathbb{C}^m, Y = \mathbb{C}^n$ and let $T : X \to Y$ be a linear operator. If A is the matrix of T in the standard basis, then the matrix of T^* is the transpose of A.

Proposition 3.3.1. Let $T : X \to Y$ be a bounded linear operator between Banach spaces and let T^* be its adjoint. Then $\phi \in \ker(T^*)$ if and only if $\phi|\mathrm{im}(T) = 0$ and $x \in \ker(T)$ if and only if $\phi(x) = 0$ for all $\phi \in \mathrm{im}(T^*)$.

Proof. We have

$$\phi \in \ker(T^*) \Leftrightarrow T^*\phi = 0 \Leftrightarrow (T^*\phi)(x) = \phi(Tx) = 0, \forall x \Leftrightarrow \phi|\mathrm{im}(T) = 0.$$

and

$$x \in \ker(T) \Leftrightarrow Tx = 0 \Leftrightarrow \phi(Tx) = (T^*\phi)(x) = 0, \forall \phi \Leftrightarrow \phi(x) = 0, \forall \phi \in \mathrm{im}(T^*).$$

□

Corollary 3.3.1. $\ker(T^*)$ is weak* closed, $\mathrm{im}(T)$ is dense if and only if T^* is injective, and T is injective if and only if $\mathrm{im}(T^*)$ is weak* dense.

Theorem 3.3.2. Let $T : X \to Y$ be a bounded linear operator between Banach spaces. The following conditions are equivalent:
(a) $\mathrm{im}(T)$ is closed in Y;
(b) $\mathrm{im}(T^*)$ is weak* closed in X^*;
(c) $\mathrm{im}(T^*)$ is norm closed in X^*.

Proof. Suppose (a) holds. Then by Proposition 3.3.1, $\phi(x) = 0$ for all $\phi \in \mathrm{im}(T^*)$ if and only if $x \in \ker(T)$. We claim that the functionals that are zero on $\ker(T)$ are the weak* closure of $\mathrm{im}(T^*)$. Indeed, this set is weak* closed and contains $\mathrm{im}(T^*)$. To prove the converse inclusion, recall that the dual of X^* with the weak* topology is X. Assume that there is ϕ_0 that is zero on $\ker(T)$ but ϕ is not in the weak*-closure of $\mathrm{im}(T^*)$. Then by the Hahn-Banach Theorem, there is $x \in X$ such that $\phi_0(x) \neq 0$ and $\phi(x) = 0$ for all $\phi \in \mathrm{im}(T^*)$.

But $\phi(Tx) = 0$ for all $\phi \in Y^*$ means that $Tx = 0$, so $x \in \ker(T)$. Then $\phi_0(x) = 0$, a contradiction. This proves our claim.

We are left to show that any functional that is zero on $\ker(T)$ is in the image of T^*. Let ϕ be such a functional. Define a linear functional ψ on $\mathrm{im}(T)$ by

$$\psi(Tx) = \phi(x).$$

It is not hard to see that ψ is well defined. Apply the Open Mapping Theorem to

$$T : X \to \mathrm{im}(T)$$

to conclude that there is $C > 0$ such that for every $y \in \mathrm{im}(T)$ there is $x \in X$ such that $Tx = y$ and $\|x\| \leq C\|y\|$. Hence

$$|\psi(y)| = |\psi(Tx)| = |\phi(x)| \leq \|\phi\|\|x\| \leq C\|\phi\|\|y\|.$$

Hence ψ is continuous. Extend ψ to the entire space using Hahn-Banach. Because

$$\phi(x) = \psi(Tx) = (T^*\psi)(x),$$

it follows that $\phi = T^*\psi$. Hence $\phi \in \mathrm{im}(T^*)$, as desired. We thus proved that (a) implies (b).

(b)⇒(c) is straightforward.

Now let us suppose that (c) holds. Let Z be the closure of $\mathrm{im}(T)$ in Y. Define $S : X \to Z$, $Sx = Tx$. As a corollary to Proposition 3.3.1, $S^* : Z^* \to X^*$ is one-to-one.

If $\phi \in Z^*$, the Hahn-Banach Theorem provides an extension $\psi \in Y^*$ of ϕ. For every $x \in X$, we have

$$(T^*\psi)(x) = \psi(Tx) = \phi(Sx) = (S^*\phi)(x).$$

Hence $S^*\phi = T^*\psi$. It follows that S^* and T^* have identical images, in particular the image of S^* is closed. Apply the Inverse Mapping Theorem to $S^* : Z^* \to \mathrm{im}(S^*)$ to conclude that it is invertible. The conclusion follows from the following result.

Lemma 3.3.2. Suppose $S : X \to Z$ is a bounded linear operator such that $S^* : Z^* \to X^*$ is invertible. Then S is onto.

Proof. Because S^* is invertible, there is $C > 0$ such that $\|\phi\| \leq C\|S^*\phi\|$ for all $\phi \in Z^*$.

Let B_X and B_Z be the unit balls in X and Z. We will show that $B_Z \subset C\overline{S(B_X)}$, namely that $\delta B_Z \subset \overline{S(B_X)}$, where $\delta = 1/C$.

Choose $z_0 \notin \overline{S(B_X)}$. Because $\overline{S(B_X)}$ is convex, closed, and balanced, an application of the Hahn-Banach Theorem shows that we can separate it from z_0, so there is $\phi \in Z^*$ such that $\|\phi(z)\| \leq 1$ for $z \in \overline{S(B_X)}$ but $|\phi(z_0)| > 1$. If $x \in B_X$, then

$$|S^*\phi(x)| = |\phi(Sx)| \leq 1.$$

Hence $\|S^*\phi\| \leq 1$. We have

$$\delta < \delta|\phi(z_0)| \leq \delta\|\phi\|\|z_0\| \leq \|z_0\|\|S^*\phi\| \leq \|z_0\|.$$

We deduce that if $\|z\| \leq \delta$ then necessarily $z \in \overline{S(B_X)}$.

Now let us show that moreover $z \in S(B_X)$. Rescaling S we may assume $\delta = 1$. Then $\overline{B_Z} \subset \overline{T(B_X)}$, and hence for every $z \in Z$ and every $\epsilon > 0$ there is $x \in X$ such that $\|x\| \leq \|y\|$ and $\|y - Tx\| < \epsilon$. Choose $z_1 \in B_Z$. Let $\epsilon_n = \frac{1}{3^n}(1 - \|z_1\|)$. Define the sequences x_n and z_n inductively as follows. Assume z_n is already picked, and let x_n be such that $\|x_n\| \leq \|z_n\|$ and $\|z_n - Tx_n\| < \epsilon_n$. Set $z_{n+1} = z_n - Tx_n$.

If $x = \sum x_n$, then $Tx = \sum Tx_n = \sum(y_n - y_{n+1}) = z_1$. Hence $z_1 \in T(B_X)$. This proves our claim. The conclusion follows. □

Using the lemma we conclude that $\mathrm{im}(S) = \overline{\mathrm{im}(S)}$, and so the image of S is closed. But $\mathrm{im}(S) = \mathrm{im}(T)$, and so the theorem is proved. □

As a corollary, we obtain the following result.

Theorem 3.3.3. Let $T : X \to Y$ be a bounded linear operator between Banach spaces. Then $\mathrm{im}(T) = Y$ if and only if T^* is one-to-one and $\mathrm{im}(T^*)$ is norm closed.

3.4 The adjoint of an operator on a Hilbert space

Let H be a Hilbert space over $\mathbb{C}$ and let $T : H \to H$ be a bounded linear operator. There is a different construction of T^* based on the Riesz representation theorem. Recall that there is an antilinear isometry between H^* and H which associates to each functional ϕ the element $z \in H$ such that $\phi(x) = \langle x, z \rangle$.

The linear operator $\phi \mapsto T^*\phi$ induces a linear operator $z \mapsto T^*z$. Moreover, the two operators have the same norm. We will use the notation T^* for the second. A direct way to define this operator is by the equality

$$\langle Tx, y \rangle = \langle x, T^*y \rangle . \tag{3.4.1}$$

Because the adjoint is defined using the inner product, we will use the following lemma several times. This lemma is only true for Hilbert spaces over $\mathbb{C}$!

Lemma 3.4.1. Two linear operators S and T on a Hilbert space H are equal if and only if

$$\langle Sx, x \rangle = \langle Tx, x \rangle \text{ for all } x \in H.$$

Proof. Recall the polarization formula for the inner product:

$$\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2).$$

We can adapt it to write

$$\langle Tx, y \rangle = \frac{1}{4}(\langle T(x + y), x + y \rangle - \langle T(x - y), x - y \rangle + i \langle T(x + iy), x + iy \rangle - i \langle T(x - iy), x - iy \rangle).$$

So if $\langle Tx, x \rangle = \langle Sx, x \rangle$ for all $x \in H$, then

$$\langle Tx, y \rangle = \langle Sx, y \rangle \text{ for all } x, y \in H.$$

This condition implies $Tx = Sx$ for all $x \in H$, i.e. $T = S$. □

By Theorem 3.3.1 $\|T^*\| = \|T\|$. Note that (3.4.1) implies that

$$(T^*)^* = T.$$

Also, it is easy to check that

$$\begin{aligned}(T+S)^* &= T^* + S^* \\ (\alpha T)^* &= \bar{\alpha} T^* \\ (ST)^* &= T^* S^*.\end{aligned}$$

Example. If $H = \mathbb{C}^n$, and $T : H \to H$ is linear, then the matrix of T is the transpose conjugate of the matrix of T.

Proposition 3.4.1. If $T : H \to H$ is a bounded linear operator on a Hilbert space, then

$$\|T^*T\| = \|T\|^2.$$

Proof. We have

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle \leq \|T^*Tx\|\|x\| \leq \|T^*T\|\|x\|^2,$$

where for the inequality we used Cauchy-Schwarz. So $\|T\|^2 \leq \|T^*T\|$. On the other hand,

$$\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2.$$

Hence the equality. □

As a corollary of Proposition 3.3.1, we obtain the following result.

Proposition 3.4.2. Let $T : H \to H$ be a bounded linear operator on a Hilbert space. Then

$$\ker(T^*) = \operatorname{im}(T)^{\perp} \text{ and } \ker(T) = \operatorname{im}(T^*)^{\perp}.$$

Proof. This can be proved directly as follows: $T^*y = 0$ if and only if $\langle x, T^*y \rangle = 0$ for all x. This is further equivalent to $\langle Tx, y \rangle = 0$ for all x, meaning that $y \in \operatorname{im}(T)^{\perp}$. □

Definition. A bounded linear operator T on a Hilbert space is said to be

- *normal* if $TT^* = T^*T$
- *self-adjoint* if $T = T^*$
- *unitary* if $TT^* = T^*T = I$
- an *isometry* if $T^*T = I$

It is standard to denote unitaries by U and isometries by V. An alternative way to say that V is an isometry is to say that $\|Vx\| = \|x\|$. It is also important to note that isometries preserve the inner product, meaning that

$$\langle Vx, Vy \rangle = \langle x, y \rangle .$$

U is unitary if it is an invertible isometry. Isometries and in particular unitaries have norm 1. Note also that self-adjoint operators are normal.

Example. If we consider the real vector space $L^2([0,\infty))$ of square integrable real valued functions on $[0,\infty)$ with values in $\mathbb{C}$, then the Laplace transform

$$L : L^2([0,\infty)) \to L^2([0,\infty)), \quad (Lf)(s) = \int_0^\infty f(t)e^{-st}dt.$$

is self-adjoint. Indeed,

$$\begin{aligned}\langle Lf, g\rangle &= \int_0^\infty (Lf)(s)\overline{g(s)}ds = \int_0^\infty \left(\int_0^\infty f(t)e^{-st}dt\right)\overline{g(s)}ds \\ &= \int_0^\infty f(t)\left(\int_0^\infty \overline{g(s)}e^{-st}ds\right)dt = \langle f, Lg\rangle .\end{aligned}$$

By Proposition 3.4.1, $L^2 = L^*L$ has norm equal to the square of the norm of the Laplace transform. Thus

$$\|L^2\| = \pi.$$

We compute

$$\begin{aligned}(L^2f)(u) &= \int_0^\infty (Lf)(s)e^{-us}ds = \int_0^\infty \int_0^\infty f(t)e^{-st}dte^{-us}ds \\ &= \int_0^\infty f(t)\int_0^\infty e^{-(t+u)s}dsdt = \int_0^\infty \frac{f(t)}{t+u}dt.\end{aligned}$$

The later is called the Hilbert-Hankel operator, and we have shown that it is a bounded (self-adjoint) operator with norm equal to π.

Example. Let ℓ^2 be the Hilbert space of complex valued square integrable sequences. The operator $S : \ell^2 \to \ell^2$, $S(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$ is an isometry that is not onto. It is called a shift.

Theorem 3.4.1. (H. Wold) Every isometry of a Hilbert space into itself can be decomposed as an orthogonal sum of operators that are unitary equivalent to the shift and a unitary operator.

Proof. Let H be the Hilbert space and let V be the isometry. Consider the inclusions

$$H \supset V(H) \supset V^2(H) \supset V^3(H) \supset \cdots \supset \cap_{n=1}^\infty V^n(H).$$

Let $H_\beta = \cap_{n=1}^\infty V^n(H)$ and $H_\alpha = H \ominus H_\beta$. Then $V|H_\beta$ is onto so it is unitary.

Let us examine $V|H_\alpha$. Define $H_k = V^k(H) \ominus V^{k-1}(H)$, $k \geq 1$. Then $V : H_k \to H_{k+1}$ is an isometric isomorphism. Decompose $H_1 = \oplus_i \mathbb{C}e_i$. Then $V|\oplus_n \mathbb{C}V^n(e_i)$, is a shift for every i, so we obtain the decomposition of $V|H_\alpha$ as an orthogonal sum of shifts. □

Proposition 3.4.3. T is normal if and only if

$$\|Tx\| = \|T^*x\|, \text{ for all } x \in H.$$

Consequently $\ker(T) = \ker(T^*)$.

Proof. Note that the equality from the statement yields

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 = \|T^*x\|^2 = \langle T^*x, T^*x\rangle = \langle TT^*x, x\rangle \,.$$

By Lemma 3.4.1 $T^*T = TT^*$, meaning that T is normal.

Conversely

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle = \langle TT^*x, x\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|^2,$$

and the proposition is proved. □

Proposition 3.4.4. If T is normal then the following properties hold:

- $\mathrm{im}(T)$ is dense if and only if T is one-to-one.
- T is invertible if and only if there is $\delta > 0$ such that $\|Tx\| \geq \delta\|x\|$ for all x.

Proof. The first property is a consequence of Proposition 3.4.3 and Proposition 3.4.2.

Assume that T is invertible. Then by the Inverse Mapping Theorem the inverse of T is continuous, so we can choose $\delta = \|T^{-1}\|$.

For the converse, the existence of such a δ implies that $\ker(T) = \{0\}$. Moreover, $\mathrm{im}(T)$ is closed, because if $(Tx_n)_n$ is Cauchy, then so is $(x_n)_n$, and if the limit of the latter is x, then $Tx = \lim Tx_n$. Finally, by the first property $\mathrm{im}(T)$ is dense. So T is one-to-one and onto, hence invertible. □

Proposition 3.4.5. An operator A is self-adjoint if and only if $\langle Ax, x\rangle$ is real for all $x \in H$.

Proof. If A is self-adjoint, then $\langle Ax, x\rangle = \langle x, Ax\rangle$. But by the properties of the inner product, $\langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$. Hence the quantity must be real. Conversely, if the quantity is real then

$$\langle x, A^*x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle \,.$$

So $A = A^*$ by Lemma 3.4.1. □

Of course the concept of a self-adjoint operator can be defined for Hilbert spaces over $\mathbb{C}$, but neither this proposition, nor Lemma 3.4.1 hold in that case.

It is important to point out that if T is an arbitrary operator, then T^*T is not only self-adjoint, but $\langle T^*Tx, x\rangle$ is nonnegative for all x. We will see later that the converse of this is also true: if $\langle Ax, x\rangle \geq 0$ for all x, there is an operator T such that $A = T^*T$.

A projection P is called orthogonal if $\mathrm{im}(P) = \ker(P)^\perp$.

Proposition 3.4.6. A projection P is orthogonal if and only if P is self-adjoint.

Proof. Assume that P is orthogonal. Then every $x \in H$ is of the form $x = y + z$ with $y \in \ker(P)$ and $z \in \mathrm{im}(P)$. Then

$$\langle Px, x\rangle = \langle z, y + z\rangle = \|z\|^2$$

and

$$\langle x, Px\rangle = \langle y + z, z\rangle = \|z\|^2.$$

Hence $P = P^*$.

For the converse, note that $P = P^*$ implies P normal, so $\ker(P) = \mathrm{im}(P^*)^\perp = \mathrm{im}(P)^\perp$. But P is a projection, so $\mathrm{im}(P)$ is closed. The conclusion follows. □

As a corollary, a property that characterizes orthogonal projections is $\langle Px, x\rangle = \|Px\|^2$.

Proposition 3.4.7. Let N be a normal operator. Then there are self-adjoint operators A_1 and A_2 that commute such that $N = A_1 + iA_2$.

Proof. $A_1 = (N + N^*)/2$, $A_2 = (N - N^*)/2i$. □

For a bounded linear operator A on a Banach space we can define

$$\exp(A) = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

This operator can be defined by

$$\exp(A)x = x + \frac{A}{1!}x + \frac{A^2}{2!}x + \frac{A^3}{3!}x + \cdots$$

and because

$$\left\|\frac{A^m}{m!}x + \frac{A^{m+1}}{(m+1)!}x + \cdots + \frac{A^n}{n!}x\right\| \le \left\|\frac{A^m}{m!}x\right\| + \left\|\frac{A^{m+1}}{(m+1)!}x\right\| + \cdots + \left\|\frac{A^n}{n!}x\right\|$$
$$\le \frac{\|A\|^m}{m!}\|x\| + \frac{\|A\|^{m+1}}{(m+1)!}\|x\| + \cdots + \frac{\|A\|^n}{n!}\|x\|$$

we see that the truncations of the series form a Cauchy sequence in the Banach space, which converges. Clearly the limit is a linear operator. Moreover, setting $m = 0$ in the above inequality we obtain $\|\exp(A)x\| \le e^{\|A\|}\|x\|$, showing that the limit is a bounded operator. Thus $\exp(A)$ is a well defined bounded operator on a Banach space.

Note that in the same manner for every bounded operator A and every holomorphic function f on the whlie plane we can define $f(A)$. Later we will extend this definition to functions that are not defined on the whole plane, and in the case of normal and self-adjoint operators, to L^∞ functions.

Proposition 3.4.8. Let A be a self-adjoint operator. Then $\exp(iA)$ is unitary.

Proof. First, note that

$$\exp(iA) = I + \frac{iA}{1!} - \frac{A^2}{2!} - \frac{iA^3}{3!} + \cdots .$$

Taking the adjoint term-by-term we see that

$$\exp(iA)^* = \exp(-iA),$$

and because iA and $-iA$ commute,

$$\exp(iA)\exp(-iA) = \exp(-iA)\exp(iA) = \exp(i(A - A)) = I.$$

It follows that $\exp(iA)$ is unitary. □

Corollary 3.4.1. If T is a bounded operator, then $\exp[i(T+T^*)]$ and $\exp(T-T^*)$ are unitary.

Proof. We have $(T+T^*)^* = T+T^*$, and $[(T-T^*)/i]^* = (T-T^*)/i$. □

Theorem 3.4.2. (Fuglede-Putnam-Rosenblum) Assume that M, N, T are bounded linear operators on a Hilbert space such that M and N are normal and

$$MT = TN.$$

Then

$$M^*T = TN^*.$$

Proof. From the statement we obtain by induction that $M^nT = TN^n$ for all n, so

$$\exp(M)T = T\exp(N).$$

It follows that

$$T = \exp(-M)T\exp(N).$$

Multiply to the right by $\exp(M^*)$ and to the left by $\exp(-N^*)$ to obtain

$$\exp(M^*)T\exp(-N^*) = \exp(M^*)\exp(-M)T\exp(N)\exp(-N^*),$$

and because $MM^* = M^*M$ and $NN^* = N^*N$, we obtain

$$\exp(M^*)T\exp(-N^*) = \exp(M^*-M)T\exp(N-N^*).$$

Set $U_1 = \exp(M^*-M)$, $U_2 = \exp(N-N^*)$. In view of the above corollary, these are unitary, in particular $\|U_1\| = \|U_2\| = 1$. We then obtain

$$\|\exp(M^*)T\exp(-N^*)\| \leq \|\exp(M^*-M)\|\|T\|\|\exp(N-N^*)\| = \|T\|.$$

Now replace M and N by $\overline{\lambda}M$ and $\overline{\lambda}N$ and repeat the same argument to conclude that

$$\|\exp(\lambda M^*)T\exp(-\lambda N^*)\| \leq \|T\| \text{ for all } \lambda \in \mathbb{C}.$$

Define the operator valued function

$$f(\lambda) = \exp(\lambda M^*)T\exp(-\lambda N^*).$$

Then for every pair of vectors $x, y \in H$, the function

$$f_{x,y} : \mathbb{C} \to \mathbb{C}, \quad f_{x,y}(\lambda) = \langle f(\lambda)x, y\rangle$$

is holomorphic. Using the Cauchy-Schwarz inequality, we conclude that

$$|f_{x,y}(\lambda)| \leq \|f(\lambda)x\|\|y\| \leq \|f(\lambda)\|\|x\|\|y\| \leq \|T\|\|x\|\|y\|,$$

namely that $f_{x,y}$ is bounded. By Liouville's theorem $f_{x,y}$ is constant. It follows that f itself is constant, so $f(\lambda) = f(0) = T$ for all λ. Hence

$$\exp(\lambda M^*)T\exp(\lambda N^*) = f(\lambda) = T.$$

Write this as

$$\exp(\lambda M^*)T = T\exp(\lambda N^*).$$

This gives for every $x, y \in H$, the equality of two power series

$$\langle \exp(\lambda M^*)Tx, y\rangle = \langle T\exp(\lambda N^*)x, y\rangle ,$$

which must be equal term-by-term. Considering the λ-term we obtain that for all x, y,

$$\langle M^*Tx, y\rangle = \langle TN^*x, y\rangle .$$

Hence $M^*T = TN^*$, as desired. □

Corollary 3.4.2. If N is normal and T commutes with N, then T commutes with N^* and N commutes with T^*.

Show that the hypothesis of the theorem does not necessarily imply $MT^* = T^*N$.

3.5 The heat equation

This section is taken from P.D. Lax, *Functional Analysis*.

Let us consider the solutions $u(x,t)$ to the heat equation

$$u_t = u_{xx},$$

that are defined for all x and $t \geq 0$ and which tend to zero sufficiently rapidly as $|x| \to \infty$.

Lemma 3.5.1. Let $u(x,t)$ be a solution as above. Then for $T > 0$,
(1) $\|u(\cdot,T)\|_\infty \leq \|u(\cdot,0)\|_\infty$; (2) $\|u(\cdot,T)\|_1 \leq \|u(\cdot,0)\|_1$; (3) $\|u(\cdot,T)\|_2 \leq \|u(\cdot,0)\|_2$.

Proof. (1) Let $k > 0$. Define $v(x,t) = ue^{-kt}$. Then v satisfies the equation

$$v_t + kv = v_{xx}.$$

Since u was assume to tend to zero rapidly as $|x| \to \infty$, the same is true for v. So in the strip $\mathbb{R} \times [0,T]$, $|v(x,t)|$ has a max, say at (x_0, t_0). We claim $t_0 = 0$. Arguing by contradiction, assume that $t_0 \in (0,T]$. If $v(x,t_0) > 0$, then (x_0,t_0) is a maximum for v, so $v_t(x_0,t_0) > 0$ and $v(\cdot,t_0)$ has a maximum at x_0, so $v_{xx}(x_0,t_0) < 0$. This is impossible. If v is negative at the max, then the max of $|v|$ is a min for v, and we get another contradiction. Now let $k \to 0$ to obtain the conclusion.

(2) Consider the space of solutions $w(x,t)$ to the backward heat equation $w_t = -w_{xx}$ defined for $0 \leq t \leq T$ and that tend rapidly to zero at infinity. Multiply this equation by u, the heat equation by w then add to obtain

$$(uw)_t = wu_{xx} - uw_{xx}.$$

Integrate (by parts) this with respect to x and use the condition at ∞ to write

$$0 = \int (uw)_t dx = \frac{d}{dt}\int uwdx,$$

so $\int uwdx = (u, w)$ is independent of time. We have

$$\int u(x,0)w(x,0)dx = \int u(x,T)w(x,T)dx.$$

In particular, if we let $u(\cdot,T) = S(T)u(\cdot,0)$ and $w(\cdot,0) = S'(T)w(\cdot,T)$, then we have $(u, S'(T)w) = (S(T)u, w)$. It is not hard to check that

$$\|u\|_1 = \sup_{\|w\|_\infty=1} |(u,w)|.$$

By part (1), $\|S'(T)w(\cdot,T)\|_\infty \le \|w(\cdot,T)\|_\infty$, and using the equality $(u, S'(T)w) = (S(T)u, w)$ we obtain the desired conclusion.

(3) Multiply the heat equation by $2u$ and integrate with respect to x. Integrate by parts the right-hand side. Then

$$\frac{d}{dt}\int u^2 dx = -\int u_x^2 dx.$$

This shows that $\int u^2(x,t)dx$ is a decreasing function of t. The lemma is proved. □

For every initial condition $u(x,0)$ we can solve the equation explicitly:

$$u(x,t) = \frac{1}{2\sqrt{\pi t}}\int u(y,0)e^{-(x-y)^2/4t}dy.$$

If we check that this gives, indeed, a solution for every initial condition in L^p, $p = 1, 2, \infty$, then the operator $S(t) : L^p \to L^p$, $p = 1, 2, \infty$, $Su(x,0) = u(x,t)$ has the property that $|S(t)| \le 1$. We notice that the solution is an integral operator $\mathbf{K}$ of the form

$$f \mapsto \int K(x,y)f(y)dy.$$

And we have the following theorem:

Theorem 3.5.1. (1) If $\sup_x \int |K(x,y)|dy < \infty$ then $\mathbf{K} : L^\infty \to L^\infty$ is bounded.
(2) If $\sup_y \int |K(x,y)|dx < \infty$ then $\mathbf{K} : L^1 \to L^1$ is bounded.
(3) If both quantities defined above are bounded then $\mathbf{K} : L^2 \to L^2$ is bounded.

Proof. For (1) we have

$$\|(\mathbf{K}f)(x)\| \le \int |K(x,y)|dy\|f\|_\infty.$$

For (2) we have

$$\|(\mathbf{K}f)(x) \leq \iint |K(x,y)||f(y)|dydx = \int \left(\int |K(x,y)|dx \right) |f(y)|dx$$
$$\leq \sup_y \int |K(x,y)|dx \|f\|_1.$$

For (3) we start with the observation that the Cauchy-Schwarz inequality implies $\|g\|_2 = \max_{\|h\|_2=1} \langle g, h \rangle$. We have

$$\langle \mathbf{K}f, h \rangle = \iint K(x,y)f(y)h(x)dydx.$$

Using the fact that if $a, b, c > 0$ then $ab \leq ca^2/2 + b^2/2c$, we see that for every $c > 0$ the right-hand side of the above is less than or equal to

$$\iint |K(x,y)| \left(\frac{c}{2}|f(y)|^2 + \frac{1}{2c}|h(x)|^2 \right) dxdy.$$

Integrate in the first term first with respect to x then with respect to y, and the other way around in the second to obtain that this is further less than or equal to

$$\frac{c}{2} \sup_y \int |K(x,y)|dx\|f\|_2^2 + \frac{1}{2c} \sup_x \int |K(x,y)dy\|h\|_2^2.$$

Next take $\|f\|_2 = \|h\|_2 = 1$, and vary c in this expression. Note that its min is

$$(\sup_y \int |K(x,y)|dx)^{1/2} (\sup_y \int |K(x,y)|ds)^{1/2}.$$

So this is an upper bound for the norm of $\mathbf{K}$. The theorem is proved. □

It is easy to see that the solution to the heat equation satisfies all three hypotheses of the theorem, because we are integrating a Gaussian.

Chapter 4

Banach Algebra Techniques in Operator Theory

4.1 Banach algebras

This section and the next follow closely R.G. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press 1972 with some input from Rudin's *Functional Analysis*.

Definition. A Banach algebra is an associative algebra with unit 1 over the complex (or real) numbers that is a Banach space and its norm satisfies

$$\|ab\| \leq \|a\|\|b\|, \text{ and } \|1\| = 1.$$

Example. The Banach algebra $\mathcal{B}(X)$ of bounded linear operators on a Banach space X.

Example. The Banach algebra of continuous functions $C([0,1])$.

We will almost always be concerned with Banach algebras over the complex numbers.

Definition. A series

$$\sum_{n=0}^{\infty} c_n a_n$$

with $c_n \in \mathbb{C}$ and $a_n \in \mathcal{B}$ is called absolutely convergent if

$$\sum_{n=0}^{\infty} |c_n|\|a_n\| < \infty$$

Proposition 4.1.1. An absolutely convergent series is convergent.

Theorem 4.1.1. Let $\mathcal{B}$ be a Banach algebra and let $a \in \mathcal{B}$ be an element such that $\|1-a\| < 1$. Then a is invertible and

$$\|a^{-1}\| \leq \frac{1}{1 - \|1-a\|}.$$

Proof. Set $b = 1 - a$. Then $\|b\| \le 1$. Then the series

$$1 + b + b^2 + b^3 + \cdots$$

is absolutely convergent so it is convergent. We have

$$(1-b)(1+b+b^2+b^3+\cdots) = \lim_{n\to\infty}(1-b)(1+b+b^2+\cdots+b^n)$$
$$= 1 - \lim_{n\to\infty} b^n = 1.$$

Hence

$$a^{-1} = (1-b)^{-1} = 1 + b + b^2 + b^3 + \cdots .$$

By the triangle inequality

$$\|a^{-1}\| \le 1 + \|b\| + \|b\|^2 + \|b\|^3 + \cdots = \frac{1}{1-\|b\|} = \frac{1}{1-\|1-a\|}.$$

□

Definition. For a Banach algebra $\mathcal{B}$, let $\mathcal{G}$, $\mathcal{G}_r$, and $\mathcal{G}_l$ be respectively the sets of invertible elements, right invertible elements that are not invertible, and left invertible elements that are not invertible.

Proposition 4.1.2. If $\mathcal{B}$ is a Banach algebra, then each of the sets $\mathcal{G}$, $\mathcal{G}_r$, and $\mathcal{G}_l$ is open.

Proof. If a is invertible, and

$$\|a-b\| < \frac{1}{\|a^{-1}\|},$$

then

$$\|1 - a^{-1}b\| \le \|a^{-1}\|\|a-b\| < 1.$$

Hence $1 - a^{-1}b$ is invertible, and so is $a(1-a^{-1}b) = a - b$. This proves that for every $a \in \mathcal{G}$ there is a ball of radius $1/\|a^{-1}\|$ centered at a and contained in $\mathcal{G}$. Hence $\mathcal{G}$ is open.

By the same argument, if $a \in \mathcal{G}_l$ and $b \in \mathcal{B}$ is such that $ba = 1$, then if c is such that $\|c-a\| < 1/\|b\|$ then bc is invertible. We have $((bc)^{-1}b)c = 1$, showing that c is left invertible. Note that c itself cannot be invertible, or else bc and c invertible implies b invertible, so a is invertible, too. This proves $\mathcal{G}_l$ open. The proof that $\mathcal{G}_r$ is open is similar. □

Proposition 4.1.3. If $\mathcal{B}$ is a Banach algebra and $\mathcal{G}$ is the subgroup of invertible elements, then the map

$$\mathcal{G} \to \mathcal{G}, \quad a \mapsto a^{-1}$$

is continuous.

Proof. Fix $a \in \mathcal{G}$. We want to show that for every $\epsilon > 0$, there is $\delta > 0$ such that if $\|b - a\| < \delta$ then $\|b^{-1} - a^{-1}\| < \epsilon$. We have

$$\|a^{-1} - b^{-1}\| = \|a^{-1}(a-b)b^{-1}\| \leq \|a^{-1}\|\|a-b\|\|b^{-1}\|.$$

If $\|b - a\| < 1/(2\|a^{-1}\|)$, then $\|1 - a^{-1}b\| < 1/2$ and so by Theorem 4.1.1

$$\|b^{-1}\| \leq \|b^{-1}a\|\|a^{-1}\| = \|(a^{-1}b)^{-1}\|\|a^{-1}\| \leq \frac{1}{1-\frac{1}{2}}\|a^{-1}\| = 2\|a^{-1}\|.$$

Hence it suffices to choose

$$\delta = \min\left(\frac{1}{2\|a^{-1}\|}, \frac{\epsilon}{2\|a^{-1}\|^2}\right).$$

□

We conclude that $\mathcal{G}$ is a topological group.

Proposition 4.1.4. Let $\mathcal{B}$ be a Banach algebra whose group of invertible elements is $\mathcal{G}$. Let $\mathcal{G}_0$ be the connected component of $\mathcal{G}$ that contains the identity element. Then $\mathcal{G}_0$ is an open and closed normal subgroup of $\mathcal{G}$. Consequently $\mathcal{G}/\mathcal{G}_0$ is a group whose induced topology is discrete.

Proof. $\mathcal{B}$ is a locally path connected space, so connected is equivalent to path connected. It is a standard fact in topology that $\mathcal{G}_0$ is open and closed. If a and b are in $\mathcal{G}_0$ and γ_a and γ_b are paths connecting them to the identity, then $\gamma_a\gamma_b$ and $(\gamma_a)^{-1}$ are paths connecting 1 to ab respectively 1 to a^{-1}. Hence $\mathcal{G}_0$ is a group. Moreover, for every $a \in \mathcal{G}_0$ and $c \in \mathcal{G}$, $c\gamma_a c^{-1}$ connects 1 to cac^{-1}, hence $cac^{-1} \in \mathcal{G}_0$. This shows that $\mathcal{G}_0$ is normal. □

Definition. The group $\Lambda_{\mathcal{B}} = \mathcal{G}/\mathcal{G}_0$ is called the *abstract index group* for $\mathcal{B}$. The *abstract index* is the natural homomorphism $\mathcal{G} \to \Lambda_{\mathcal{B}}$.

4.2 Spectral theory for Banach algebras

Let $\mathcal{B}$ be a Banach algebra.

Definition. Let a be an element of $\mathcal{B}$. The *spectrum* of a is the set

$$\sigma_{\mathcal{B}}(a) = \{\lambda \in \mathbb{C} \,|\, a - \lambda \notin \mathcal{G}\}.$$

The resolvent is the set

$$\rho_{\mathcal{B}}(a) = \{\lambda \in \mathbb{C} \,|\, a - \lambda \in \mathcal{G}\}.$$

So the spectrum consists of those λ for which $a - \lambda$ is not invertible, and the resolvent is the complement in $\mathbb{C}$ of the spectrum. When there is no risk of confusion, we ignore the subscript, but be careful, the spectrum depends on the algebra in which your element lies (in case the given element can be put inside several Banach algebras).

Example. If $\mathcal{B} = M_n(\mathbb{C})$, the algebra of $n \times n$ matrices, and $A \in M_n(\mathbb{C})$, then $\sigma(A)$ is the set of eigenvalues.

Theorem 4.2.1. The spectrum of an element $a \in \mathcal{B}$ is nonempty and compact. Moreover, the spectrum lies inside the closed disk of radius $\|a\|$ centered at the origin.

Proof. First, note that if $|\lambda| > \|a\|$, then by Theorem 4.1.1, $1 - a/\lambda$ is invertible. Hence $\lambda(1 - a/\lambda) = \lambda - a$ is invertible. This shows that the spectrum is included in the closed disk of radius $\|a\|$ centered at the origin.

Let us show that the spectrum is nonempty. Assume to the contrary that for some element a the spectrum is empty. Let ϕ be a continuous linear functional on $\mathcal{B}$. Consider the function

$$f_\phi : \mathbb{C} \to \mathbb{C}, \quad f_\phi(\lambda) = \phi((a - \lambda)^{-1}).$$

We claim that f_ϕ is holomorphic. Indeed,

$$\begin{aligned}\lim_{\lambda \to \lambda_0} \frac{f_\phi(\lambda) - f_\phi(\lambda_0)}{\lambda - \lambda_0} &= \phi\left(\lim_{\lambda \to \lambda_0} \frac{(a - \lambda_0)^{-1}[(a - \lambda_0) - (a - \lambda)](a - \lambda)^{-1}}{\lambda - \lambda_0}\right) \\ &= \phi(\lim_{\lambda \to \lambda_0} (a - \lambda_0)^{-1}(a - \lambda)^{-1}) = \phi((a - \lambda_0)^{-2}).\end{aligned}$$

For $|\lambda| > \|a\|$, we have by Theorem 4.1.1 that $1 - a/\lambda$ is invertible and

$$\left\|(1 - a/\lambda)^{-1}\right\| < \frac{1}{1 - \|a/\lambda\|}.$$

Hence

$$\begin{aligned}\limsup_{|\lambda| \to \infty} |f_\phi(\lambda)| &= \limsup_{|\lambda| \to \infty} |\phi\left(\frac{1}{\lambda}(a/\lambda - 1)^{-1}\right) \\ &\leq \limsup \frac{1}{|\lambda|}\|\phi\|\|(a/\lambda - 1)^{-1}\| \leq \limsup \frac{1}{|\lambda|}\|\phi\|\frac{1}{1 - \|a/\lambda\|}\end{aligned}$$

where for the last step we used Theorem 4.1.1. This last limit is zero. Hence f_ϕ is a bounded holomorphic function. By Liouville's Theorem it is constant.

Using the Hahn-Banach Theorem we deduce that $\lambda \mapsto (a - \lambda)^{-1}$ is constant, and since the inverse is unique, it follows that $\lambda \mapsto a - \lambda$ is constant. But this is clearly not true. Hence our assumption was false, and the spectrum is nonempty.

Since the map $\lambda \to a - \lambda$ is continuous, and $\mathcal{G}$ (the set of invertible elements) is open, the inverse image of $\mathcal{G}$ through this map is open. But the inverse image of $\mathcal{G}$ is the resolvent. Hence the resolvent is open, and therefore the spectrum is closed. Being bounded (as it lies inside the disk of radius $\|a\|$), it is compact. □

Example. Let

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then $\sigma(A) = \{0\}$. Note that $\|A\| = 1$.

Example. Let $S : \ell^2 \to \ell^2$, $S(x_1, x_2, \ldots, x_n, \ldots) = (0, x_1, \ldots, x_{n-1}, \ldots)$ be the shift. Then by Proposition 3.4.1, $\|S\|^2 = \|S^*S\| = \|I\| = 1$, since S is an isometry. Thus by Theorem 4.2.1, $\sigma(S) \subset \{\lambda \in \mathbb{C} \,\|\, |\lambda| \leq 1\}$.

Note that $0 \in \sigma(S)$ because $S = S - 0$ is not onto. Also, if $|\lambda| < 1$, then the sequence $(\lambda^{n-1})_{n\geq 1}$ is in ℓ^2. However, if we try to solve $(\lambda - S)((x_n)_{n\geq 1}) = (\lambda^{n-1})_{n\geq 1}$, we notice that

$$x_n = \lambda^{-n}\left(1 + \frac{1 - \lambda^{2n}}{\lambda^{-1} - \lambda}\right),$$

and it is not hard to see that $\lim_{n\to\infty} x_n \neq 0$, so $\lambda - S$ is not onto. Thus the spectrum contains the closure of the open unit disk, and so $\sigma(S) = \{\lambda \in \mathbb{C} \,|\, |\lambda| \leq 1\}$.

Example. Let $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ be the translation operator $(Tf)(x) = f(x+1)$. It is unitary, so its spectrum is a priori a closed subset of the closed unit disk. If we consider the Fourier transform

$$(\mathcal{F})f(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ixy} f(x)dx,$$

then $\mathcal{F}T\mathcal{F}^{-1}$ is the operator of multiplication by the function $f(y) = e^{-iy}$. This operator has the spectrum equal to the unit circle, so the same is true for T.

In view of Theorem 4.2.1 we define the spectral radius to be

$$r_{\mathcal{B}}(a) = \sup\{|\lambda| \,|\, \lambda \in \sigma_{\mathcal{B}}(a)\}.$$

Proposition 4.2.1. (Beurling-Gelfand) The spectral radius is given by the formula

$$r_{\mathcal{B}}(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \inf\{\|a^n\|^{1/n} \,|\, n \geq 1\}$$

Proof. Fix an element $a \in \mathcal{B}$ and let $|\lambda| > \|a\|$. Then using Theorem 4.1.1 we can write

$$(\lambda - a)^{-1} = \lambda^{-1} + \lambda^{-2}a + \lambda^{-3}a^2 + \cdots .$$

The series converges absolutely on every circle $C(0, r)$ centered at the origin and radius $r > \|a\|$. We can therefore multiply by λ^n, then integrate term by term and write

$$a^n = \frac{1}{2\pi i} \int_{C(0,r)} \lambda^n (\lambda - a)^{-1} d\lambda, \quad n = 1, 2, 3, \ldots \tag{4.2.1}$$

Here we used the fact that λ^k has an antiderivative in the plane for all $k \neq -1$, so its integral is zero, while the integral of λ^{-1} on the circle is $2\pi i$.

Let ϕ be a continuous linear functional. Then as we saw before $\phi((\lambda - a)^{-1})$ is holomorphic. From (4.2.1) we deduce

$$\phi(a^n) = \frac{1}{2\pi i} \int_{C(0,r)} \lambda^n \phi((\lambda - a)^{-1}) d\lambda.$$

The right-hand side is an integral of a holomorphic function, and so by Cauchy's theorem the equality also holds true for all circles for which $\phi((\lambda - a)^{-1})$ is defined. Thus the equality holds for $r > r_{\mathcal{B}}(a)$. Because of the Hahn-Banach theorem we can conclude that

$$a^n = \frac{1}{2\pi i} \int_{C(0,r)} \lambda^n (\lambda - a)^{-1} d\lambda, \text{ for } n \geq 0, r > r_{\mathcal{B}}(a).$$

Let $M(r)$ be the maximum of $\|(\lambda-a)^{-1}\|$ on $C(0,r)$, (which is finite because $\lambda \mapsto (\lambda-a)^{-1}$ is continuous). Then

$$\|a^n\| \leq r^{n+1}M(r), \quad n \geq 0, r > r_{\mathcal{B}}(a).$$

But $M(r)$ is bounded when $r \to \infty$, so

$$\limsup_{n\to\infty} \|a^n\|^{1/n} \leq r, \quad r > r_{\mathcal{B}}(a).$$

Hence

$$r_{\mathcal{B}}(a) \geq \limsup_{n\to\infty} \|a^n\|^{1/n}.$$

On the other hand, if $\lambda \in \sigma_{\mathcal{B}}(a)$, then $\lambda^n \in \sigma_{\mathcal{B}}(a^n)$, because $\lambda^n - a^n = (\lambda - a)(\lambda^{n-1} + \cdots + a^{n-1})$, which is therefore not invertible. Hence $|\lambda^n| \leq \|a^n\|$. We thus have

$$r_{\mathcal{B}}(a) \leq \inf_{n\geq 1} \|a^n\|^{1/n}.$$

Combining the two inequalities we deduce

$$r_{\mathcal{B}}(a) = \lim_{n\to\infty} \|a^n\|^{1/n} = \inf\{\|a^n\|^{1/n} \mid n \geq 1\}$$

and we are done. □

Here is a first application of the notion of spectrum.

Theorem 4.2.2. (Gelfand-Mazur) Let $\mathcal{B}$ be a Banach algebra which is a division algebra (i.e. every nonzero element has an inverse). Then there is a unique isometric isomorphism of $\mathcal{B}$ onto $\mathbb{C}$.

Proof. If $a \in \mathcal{B}$, then $\sigma(a) \neq \emptyset$. If $\lambda \in \sigma(a)$, then $a - \lambda$ is not invertible. Hence $a - \lambda = 0$, that is $a = \lambda$. Moreover, if $\lambda' \neq \lambda$, then $\lambda' - a = \lambda' - \lambda$, which is invertible. Hence the spectrum of each element consists of only one point. The map that associates to each element the unique point in its spectrum is an isometric isomorphism of $\mathcal{B}$ onto $\mathbb{C}$ (it is isometric because $\|\lambda\| = |\lambda|\|1\| = 1$ is a requirement in the definition of a Banach algebra). Moreover, if ψ were an arbitrary isometric isomorphism, and if a is an element in $\mathcal{B}$ with spectrum $\{\lambda\}$, then we saw that $a = \lambda$. So $\psi(a) = \psi(\lambda 1) = \lambda\psi(1) = \lambda$, showing that ψ is the above constructed homomorphism. Hence the conclusion. □

4.3 Functional calculus with holomorphic functions

Let $a \in \mathcal{B}$. Then $\sigma(a)$ is a compact subset of the plane. Consider a domain D that contains $\sigma(a)$, and let Γ be a smooth oriented contour (maybe made out of several curves) that does not cross itself such that $\sigma(a)$ is surrounded by Γ in D and such that Γ travels around $\sigma(a)$ in the counterclockwise direction.

For a holomorphic function in D, we have the Cauchy formula

$$f(z_0) = \frac{1}{2\pi i}\int_\Gamma f(z)(z - z_0)^{-1} dz.$$

Now let us replace z_0 by a. Then on Γ, the element $(z - a)^{-1}$ is defined. With Cauchy's formula in mind, we can define

$$f(a) = \frac{1}{2\pi i}\int_\Gamma f(z)(z - a)^{-1} dz. \tag{4.3.1}$$

Lemma 4.3.1. The operator $f(a)$ is well defined and does not depend on the contour Γ.

Proof. Because on $z \mapsto \|(z-a)^{-1}\|$ is continuous on $\rho(a)$ and Γ is a compact subset of $\rho(a)$, it follows that $\sup_\Gamma \|(z-a)^{-1}\| < \infty$. So the integral can be defined using limits of Riemann sums, which converge by Proposition 4.1.1. Hence the definition makes sense.

Let ϕ be a continuous linear functional. The function $z \mapsto f(z)\phi((z-a)^{-1})$ is holomorphic. By Cauchy's theorem, the integral

$$\frac{1}{2\pi i}\int_\Gamma f(z)\phi((z-a)^{-1})dz = \phi\left(\frac{1}{2\pi i}\int_\Gamma f(z)(z-a)^{-1}dz\right)$$

does not depend on Γ. So $f(a)$ itself does not depend on Γ. □

However, if $f(z) = \sum_n c_n z^n$ is an entire function, then we can define the element

$$f(a) = \sum_n c_n a^n,$$

since again the series converges. The integral formula (4.3.1) would be meaningful only if in this particular situation the two versions coincide. And indeed, we have the following result.

Proposition 4.3.1. If $f(z) = \sum_n c_n z^n$ is a series that converges absolutely in a disk centered at the origin that contains $\sigma(a)$, then

$$\sum_n c_n a^n = \frac{1}{2\pi i}\int_\Gamma f(z)(z-a)^{-1}dz$$

for every oriented contour Γ that surrounds $\sigma(a)$ counterclockwise.

Proof. Choose N large enough so that $\sum_{n>N} |c_n|\|a\|^n$ and $\sup_\Gamma \sum_{n>N} |c_n z^n|$ are as small as we wish. Then we can ignore these sums and consider just the case where $f(z) = \sum_{n=0}^N c_n z^n$. To prove the result in this case, it suffices to check it for f a power of z. Thus let us show that

$$a^n = \frac{1}{2\pi i}\int_\Gamma z^n (z-a)^{-1}dz.$$

Now we can rely on Cauchy's theorem about the integral of a holomorphic function, to to make Γ a circle of radius greater than $\|a\|$. Because on Γ $|z| > \|a\|$, we can expand

$$(z-a)^{-1} = \sum_{k\geq 0} a^k/z^{k+1}.$$

The series on the right is absolutely convergent, so we can integrate term-by-term to write

$$\frac{1}{2\pi i}\int_\Gamma z^n(z-a)^{-1}dz = \sum_{k=0}^{\infty}\frac{1}{2\pi i}\left(\int_\Gamma z^{n-k-1}dz\right)a^k.$$

All of the integrals are zero, except for the one where $k = n$, which is equal to $2\pi i$. Hence the result is a^n, as desired. $\square$

A slight modification of the proof yields the following more general result.

Proposition 4.3.2. Suppose $R(z) = P(z)/Q(z)$ is a rational function with poles outside of the spectrum of a. Then $R(a)$ is well defined and $Q(a)$ is invertible, and

$$R(a) = P(a)Q(a)^{-1}.$$

Theorem 4.3.1. (The Spectral Mapping Theorem for Polynomials) Let $P(z)$ be a polynomial and a an element in $\mathcal{B}$. Then

$$\sigma(P(a)) = P(\sigma(a)).$$

Proof. Let $\lambda \in \sigma(a)$. Then

$$P(a) - P(\lambda) = (a-\lambda)Q(a).$$

Because $a-\lambda$ is not invertible, neither is $P(a)-P(\lambda)$. Hence $P(\lambda) \in \sigma(P(a))$. Consequently $P(\sigma(a)) \subset \sigma(P(a))$.

Let $\lambda \in \sigma(P(a))$, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the roots of $P(z) = \lambda$. Then

$$P(a) - \lambda = (a-\lambda_1)(a-\lambda_2)\cdots(a-\lambda_n).$$

Because $P(a) - \lambda$ is not invertible, there is k such that $a - \lambda_k$ is not invertible. Then $\lambda_k \in \sigma(a)$, and $\lambda = P(\lambda_k) \in P(\sigma(a))$. This proves $\sigma(P(a)) \subset P(\sigma(a))$. The double inclusion yields the desired equality. $\square$

Theorem 4.3.2. Let D be a domain in $\mathbb{C}$ that contains $\sigma(a)$. Endow the space of holomorphic functions on D, $Hol(D)$, with the topology of uniform convergence on compact subsets. Then the map $Hol(D) \to \mathcal{B}$, $f \mapsto f(a)$ is a continuous algebra homomorphism.

Proof. The only difficult step is multiplicativity. But we have multiplicativity for polynomials, and hence for rational functions. By Runge's theorem, every function in $Hol(D)$ is the limit of rational functions. By passing to the limit in $f_n(a)g_n(a) = (f_ng_n)(a)$, we conclude that multiplicativity holds in general. $\square$

Theorem 4.3.3. (The Spectral Mapping Theorem for Holomorphic Functions) Let f be a holomorphic function in a neighborhood of the spectrum of a. Then

$$\sigma(f(a)) = f(\sigma(a)).$$

Proof. Let $\lambda \in \sigma(a)$. Then as before $f(z) - f(\lambda) = (z-\lambda)g(z)$ with g a holomorphic function with the same domain as f. By the previous theorem

$$f(a) - f(\lambda) = (a - \lambda)g(a),$$

so $f(a) - f(\lambda)$ is not invertible. Hence $f(\sigma(a)) \subset \sigma(f(a))$.

For the opposite inclusion, let $\lambda \in \sigma(f(a))$. If $f(z) - \lambda$ is nowhere zero on the spectrum of a, then $g(z) = (f(z) - \lambda)^{-1}$ is defined on the spectrum of a, and then

$$(f(a) - \lambda)(f - \lambda)^{-1}(a) = 1$$

which cannot happen. So $f(z) - \lambda$ is zero for some $z \in \sigma(a)$, that is $\lambda \in f(\sigma(a))$. □

4.4 Compact operators, Fredholm operators

In this section we will construct a Banach algebra which is not the algebra of bounded linear operators on a Banach space. For this we introduce the notion of a compact operator.

Definition. Let X be a Banach space. An operator $K \in \mathcal{B}(X)$ is called compact if the closure of the image of the unit ball is compact.

Example. If R is such that $\mathrm{im}(R)$ is finite dimensional, then R is compact. Such an operator is said to be of finite rank.

Theorem 4.4.1. The set $\mathcal{K}(X)$ of compact linear operators on X is a closed two-sided ideal of $\mathcal{B}(X)$.

Proof. Let K_1 and K_2 be compact operators. Then $\overline{K_1(B_{0,1})}$ and $\overline{K_2(B_{0,1})}$ are compact. Then

$$\overline{(K_1 + K_2)(B_{0,1})} \subset \overline{K_1(B_{0,1})} + \overline{K_2(B_{0,1})}$$

and the latter is compact because is the image through the continuous map $(x, y) \mapsto x + y$ of the compact set $\overline{K_1(B_{0,1}} \times \overline{K_2(B_{0,1})} \subset X \times X$. This proves that $K_1 + K_2$ is compact.

Also for every $\lambda \in \mathbb{C}$, if K is compact then λK is compact, because the image of the set $\overline{K_1(B_{0,1})}$ through the continuous map $x \mapsto \lambda x$ is compact.

Finally, if $T \in \mathcal{B}(X)$ and $K \in \mathcal{K}(X)$ then $T(\overline{K(B_{0,1})})$ is the image of a compact set through a continuous map, so it is compact. It follows that $TK(B_{0,1})$ lies inside a compact set, so its closure is compact. So TK is compact.

On the other hand, $T(B_{0,1})$ is a subset of $B_{0,n}$ for some n, so $\overline{KT(B_{0,1})}$ is a closed subset of the compact set $\overline{K(B_{0,n})}$, hence is compact. This proves that KT is compact.

We thus showed that $\mathcal{K}(X)$ is an ideal. Let us prove that it is closed. Let K_n, $n \geq 1$, be a sequence of compact operators that is norm convergent to an operator T. We want to prove that T is compact. For this we use the characterization of compactness in metric spaces: "Every sequence contains a convergent subsequence."

Let $x_k, k \geq 1$ be a sequence of points in the unit ball of X. Let us examine the sequence $Tx_k, n \geq 1$. For every $\epsilon > 0$, there is $n(\epsilon)$ such that for $n \geq n(\epsilon)$, $\|K_n x_k - Tx_k\| \leq \|K_n - T\| \leq \epsilon$. For $n \geq n(\epsilon)$,

$$\|Tx_k - Tx_l\| \leq \|Tx_k - K_n x_k\| + \|K_n x_k - K_n x_l\| + \|K_n x_l - Tx_l\| \leq 2\epsilon + \|K_n x_k - K_n x_l\|.$$

The sequence $K_n x_k$ has a convergent subsequence, and so we can find a subsequence $T_n x_{k_m}$ such that $\|Tx_{k_m} - Tx_{k_r}\| < 3\epsilon$ for all m, r. Do this for $\epsilon = 1$, then choose the first term of a sequence y_k to be x_{k_1}. Inductively let $\epsilon = 1/k$, and choose from the previous sequence x_{k_m} a subseqence such that $\|Tx_{k_m} - Tx_{k_r}\| < 3\epsilon$ and let y_k be the first term of this subsequence. The result is a Cauchy sequence Ty_k, which therefore converges. We conclude that T is compact. □

Theorem 4.4.2. Let $K \in \mathcal{B}(X)$ be a compact operator. Then

(a) If $\text{im}(K)$ is closed, then $\text{dimim}(K) < \infty$.

(b) If $\lambda \neq 0$, then $\text{dimker}(K - \lambda I) < \infty$.

(c) If $\text{dim}X = \infty$, then $0 \in \sigma(K)$.

Proof. (a) If $\text{im}(K)$ is closed, then it is a Banach space. The Open Mapping Theorem implies that the image of the unit ball is a neighborhood of the origin. This neighborhood is compact, and this only happens if $\text{im}(K)$ is finite dimensional.

(b) The operator $K|\text{ker}(K - \lambda I)$ is a multiple of the identity operator. This operator is also compact. By (a) this can only happen if we are in a finite dimensional situation.

(c) The operator K cannot be onto.

□

Theorem 4.4.3. Let $\mathcal{B}$ be a Banach algebra and let $\mathcal{M}$ be a two-sided closed ideal. Then $\mathcal{B}/\mathcal{M}$ is a Banach algebra with the norm

$$\|[a]\| = \inf\{\|a + m\| \mid m \in \mathcal{M}\}.$$

Here we denote by $[a]$ the image of a under the quotient map.

Proof. Let us show first that $\|\cdot\|$ is a norm. Clearly if $[a] = 0$ then $a \in \mathcal{M}$ so $\|[a]\| \leq \|a - a\| = 0$. Now assume that $\|[a]\| = 0$. Then there is a sequence $m_n \in \mathcal{M}$ such that $\lim_{n\to\infty} \|a + m_n\| = 0$. Since $\mathcal{M}$ is closed, it follows that $a \in \mathcal{M}$, so $[a] = 0$. Thus $\|[a]\| = 0$ if and only if $[a] = 0$.

If $a \in \mathcal{B}$ and $\alpha \in \mathbb{C}$, then

$$\|\alpha[a]\| = \|[\alpha a]\| = \inf\{\|a + \alpha m\| \mid m \in \mathcal{M}\} = |\alpha| \inf\{\|a + m'\| \mid m' \in \mathcal{M}\} = |\alpha|\|[a]\|.$$

Also

$$\begin{aligned}\|[a] + [b]\| = \|[a + b]\| &= \inf\{\|a + b + m\| \mid m \in \mathcal{M}\} = \inf\{\|a + m + b + m'\| \mid m, m' \in \mathcal{M}\} \\ &\leq \inf\{\|a + m\| \mid m \in \mathcal{M}\} + \inf\{\|a + m\| \mid m \in \mathcal{M}\} = \|[a]\| + \|[b]\|.\end{aligned}$$

Thus $\|\cdot\|$ is a norm.

Next, let us show that the norm satisfies the requirements from the definition of a Banach algebra. First,

$$\|[1]\| = \inf\{\|1+m\| \,|\, m \in \mathcal{M}\} = 1,$$

where the equality is attained for $m = 0$, and one cannot have $\|1+m\| < 1$ for in that case m must be invertible and hence cannot be an element of an ideal.

Secondly, for $a, b \in \mathcal{B}$, we have

$$\begin{aligned}\|[a][b]\| = \|[ab]\| = \inf\{\|ab+m\| \,|\, m \in \mathcal{M}\} &\leq \inf\{\|(a+m_1)(b+m_2) \,|\, m_1, m_2 \in \mathcal{M}\} \\ &\leq \inf\{\|a+m_1\| \,|\, m_1 \in \mathcal{M}\} \inf\{\|b+m_2\| \,|\, m_2 \in \mathcal{M}\} = \|[a]\|\|[b]\|.\end{aligned}$$

Finally, let us show that $\mathcal{B}/\mathcal{M}$ is complete. Showing that every Cauchy sequence is convergent is equivalent to showing that every absolutely convergent series is convergent. It is clear that the fact that every Cauchy sequence is convergent implies that every absolutely convergent series is convergent. For the converse, let $x_n, n \geq 1$, be a Cauchy sequence. By choosing a subsequence, we may assume that $|y_n - y_m| \leq 1/2^k$ whenever $n, m \geq k$. Set $x_k = y_{k+1} - y_k$. Then $\sum x_k$ is absolutely convergent, and its sum is the limit of y_k.

So let $\sum_n [a_n]$ be a series such that $\sum_n \|[a_n]\| = M < \infty$. Then for each n there is m_n sucht that $\|a_n + m_n\| \leq \|a_n\| + 1/2^n$. Hence $\sum_n (a_n + m_n)$ is absolutely convergent, and therefore convergent in $\mathcal{B}$. If a is its sum, then $a + \mathcal{M}$ is the sum of the original series in $\mathcal{B}/\mathcal{M}$. This concludes the proof that $\mathcal{B}/\mathcal{M}$ is a Banach algebra. □

Corollary 4.4.1. The algebra $\mathcal{B}(X)/\mathcal{K}(X)$ is a Banach algebra.

Definition. The algebra $\mathcal{B}(X)/\mathcal{K}(X)$ is called the Calkin algebra.

Definition. An operator with finite dimensional kernel and with closed image of finite codimension is called *Fredholm*.

Example. A very standard example is the shift.

Theorem 4.4.4. Let H be a Hilbert space. Then an operator is compact if and only if it is the limit of a sequence of finite rank operators.

Theorem 4.4.5. (Atkinson) Let H be a Hilbert space. Then the Fredholm operators form the preimage through the quotient map of the invertible elements of $\mathcal{B}(H)/\mathcal{K}(H)$.

Corollary 4.4.2. The Fredholm operators form an open set.

Definition. If T is Fredholm, then the index of T is

$$\operatorname{ind}(T) = \dim\ker(T) - \operatorname{codim}\operatorname{im}(T) = \operatorname{codim}\operatorname{im}(T^*) - \dim\ker(T^*).$$

Theorem 4.4.6. The index is continuous and invariant under compact perturbations.

4.5 The Gelfand transform

Definition. Let $\mathcal{B}$ be a Banach algebra. A complex linear functional ϕ on $\mathcal{B}$ is said to be multiplicative if
(a) $\phi(ab) = \phi(a)\phi(b)$ for all a, b,
(b) $\phi(1) = 1$.

We denote the set of all multiplicative functionals by $M_{\mathcal{B}}$.

Proposition 4.5.1. If $\mathcal{B}$ is a Banach algebra and $\phi \in M_{\mathcal{B}}$, then $\|\phi\| = 1$.

Proof. Since $\phi(a - \phi(a)) = 0$ it follows that every element in $\mathcal{B}$ is of the form $\lambda + a$, for some $\lambda \in \mathbb{C}$ and $a \in \ker(\phi)$. Note that if $\lambda \neq 0$ and $\|\lambda + a\| < |\lambda| = |\phi(\lambda + a)|$, then a is invertible. This cannot happen, because $\phi(a) = 0$ implies $1 = \phi(aa^{-1}) = \phi(a)\phi(a^{-1}) = 0$. Hence $|\phi(b)| \leq \|b\|$ for all b. Because $\phi(1) = 1$, the equality is attained, so $\|\phi\| = 1$. □

Proposition 4.5.2. $M_{\mathcal{B}}$ is a compact subspace of X^* endowed with the weak* topology.

Proof. As a corollary of the previous proposition, $M_{\mathcal{B}}$ is a subset of the unit ball in X^*. Because of the Banach-Alaoglu theorem, all we have to show is that $M_{\mathcal{B}}$ is weak*-closed. This amounts to showing that if a linear functional is in the weak*-closure of this set, then it is multiplicative.

Assume $\phi(1) \neq 1$, and let $\epsilon < |\phi(1) - 1|$. If $\psi \in M_{\mathcal{B}} \cap V(1, \epsilon)$, then

$$\epsilon < |\phi(1) - 1| = |\phi(1) - \psi(1)| < \epsilon.$$

This is impossible, so $\phi(1) = 1$.

Similarly, if $\phi(ab) \neq \phi(a)\phi(b)$ for some a, b (which we may assume to lie in the unit ball), choose $\epsilon = |\phi(ab) - \phi(a)\phi(b)|$, and $\psi \in V(a, b, ab, \epsilon/3)$. Then

$$\begin{aligned}
&\epsilon < |\phi(ab) - \phi(a)\phi(b)| = |\phi(ab) - \psi(ab) + \psi(ab) - \psi(a)\psi(b) + \psi(a)\psi(b) - \phi(a)\phi(b)| \\
&\leq |\phi(ab) - \psi(ab)| + |\phi(a)\phi(b) - \psi(a)\psi(b)| \leq \epsilon/3 + |\phi(a)\phi(b) - \phi(a)\psi(b) + \phi(a)\psi(b) - \psi(a)\psi(b)| \\
&\leq \epsilon/3 + |\phi(a)||\phi(b) - \psi(b)| + |\psi(b)||\phi(a) - \psi(a)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.
\end{aligned}$$

Again this is impossible, so ϕ is multiplicative. □

Proposition 4.5.3. If $\mathcal{B}$ is a commutative Banach algebra, then $M_{\mathcal{B}}$ is in one-to-one correspondence with the set of maximal two-sided ideals in $\mathcal{B}$.

Proof. The correspondence is $\phi \mapsto \ker(\phi)$.

So first, let us show that if ϕ is a multiplicative linear functional, then $\ker(\phi)$ is a maximal two-sided ideal. That it is an ideal follows from $\phi(a) = 0 \to \phi(ab) = \phi(a)\phi(b) = 0$. It is maximal because every element in $\mathcal{B}$ is of the form $\lambda + a$ where $a \in \ker(\phi)$. If a were in an ideal and $\lambda + a$ were in an ideal, then λ and hence 1 would be in an ideal, which is impossible. Hence the kernel is maximal.

Conversely, let $\mathcal{M}$ be a maximal two-sided ideal. We will prove that there is $\phi \in M_{\mathcal{B}}$ such that $\ker(\phi) = \mathcal{M}$. Because if $a \in \mathcal{M}$, then a is not invertible, then $\|1 - a\| \geq 1$, so 1 is not in the closure of $\mathcal{M}$. Thus the closure of $\mathcal{M}$ is an ideal, and because of maximality, this ideal must be $\mathcal{M}$. So $\mathcal{M}$ is closed.

The quotient algebra $\mathcal{B}/\mathcal{M}$ is a division algebra, because $\mathcal{M}$ is maximal. So by the Gelfand-Mazur Theorem it is $\mathbb{C}$. The quotient map is the desired multiplicative functional. □

Recall that for every $a \in \mathcal{B}$, the function $\hat{a} : (\mathcal{B}^*)_1 \to \mathcal{C}$ given by $\hat{a}(\phi) = \phi(a)$ is continuous, where $(\mathcal{B}^*)_1$ is the unit ball in $\mathcal{B}^*$, endowed with the weak* topology.

Definition. The *Gelfand transform* of the Banach algebra $\mathcal{B}$ is the function $\Gamma : \mathcal{B} \to C(M_\mathcal{B})$ given by $\Gamma(a) = \hat{a}|M_\mathcal{B}$.

Theorem 4.5.1. The Gelfand transform is an algebra homomorphism and $\|\Gamma(a)\| \leq \|a\|$ for all $a \in \mathcal{B}$.

Proof. Γ is clearly linear and $\Gamma(1) = 1$. Let us check that Γ is multiplicative. We have

$$[\Gamma(ab)](\phi) = \phi(ab) = \phi(a)\phi(b) = [\Gamma(a)](\phi)[\Gamma(b)](\phi) = [\Gamma(a)\Gamma(b)](\phi).$$

Next, let us check that Γ is contractive. We have

$$\|\Gamma(a)\| = \sup\{|\phi(a)| \mid \phi \in M_\mathcal{B}\} \leq \sup\{\|\phi\|\|a\| \mid \phi \in M_\mathcal{B}\} = \|a\|.$$

□

If $\mathcal{B}$ is not commutative, the Gelfand transform has large kernel which is generated by the elements of the form $ab - ba$. For this reason it is not so interesting.

Proposition 4.5.4. If $\mathcal{B}$ is a commutative Banach algebra and $a \in \mathcal{B}$, then a is invertible in $\mathcal{B}$ if and only if $\Gamma(a)$ is invertible in $C(M_\mathcal{B})$.

Proof. If a is invertible, then $\Gamma(a^{-1}) = (\Gamma(a))^{-1}$. If a is not invertible, then $\mathcal{M}_0 = \{ab \mid b \in \mathcal{B}\}$ is a proper ideal. It is contained in a maximal ideal, whose associated functional is zero on a. Hence $\Gamma(a)$ is not invertible. □

Remark 4.5.1. The fact that a invertible implies $\Gamma(a)$ invertible does not use the fact that the Banach algebra is commutative. Because $\Gamma(ab-ba) = \Gamma(a)\Gamma(b) - \Gamma(b)\Gamma(a) = 0$, it follows that $ab - ba$ is not invertible. This means that the canonical commutation relations for the position and momentum operators in quantum mechanics

$$PQ - QP = \frac{\hbar}{i} I$$

cannot be modeled with bounded linear operators.

Proposition 4.5.5. If $\mathcal{B}$ is a commutative Banach algebra and $a \in \mathcal{B}$, then $\sigma_\mathcal{B}(a) = \mathrm{im}(\Gamma(a))$ and $r_\mathcal{B}(a) = \|\Gamma(a)\|$.

Proof. If λ is not in $\sigma(a)$, then $a - \lambda$ is invertible. This is equivalent to $\Gamma(a) - \lambda$ is invertible. And this is further equivalent to the fact that λ is not in the image of $\Gamma(a)$. □

Chapter 5

C^* algebras

5.1 The definition of C^*-algebras

Again, most of this chapter is from the book of Ronald Douglas.

Definition. A C^*-algebra is a Banach algebra over the complex numbers with an involution $*$ that satisfies

- $(a+b)^* = a^* + b^*$
- $(\lambda a)^* = \bar{\lambda} a^*$
- $(ab)^* = b^* a^*$
- $(a^*)^* = a$.

Additionally, the involution should satisfy

$$\|a^* a\| = \|a\| \|a^*\|. \tag{5.1.1}$$

Alternatively, the involution should satisfy

$$\|a^* a\| = \|a\|^2. \tag{5.1.2}$$

The two conditions (5.1.1) and (5.1.2) are equivalent, though it is hard to show that (5.1.1) implies (5.1.2). Thus our working definition will be the one with (5.1.2), what is usually called a B^*-algebra. This condition implies (5.1.1) as follows:

$$\|x\|^2 = \|x^* x\| \leq \|x\| \|x^*\|.$$

Hence $\|x\| \leq \|x^*\|$ and $\|x^*\| \leq \|(x^*)^*\| = \|x\|$. So $\|x\| = \|x^*\|$. Then $\|x^* x\| = \|x\|^2 = \|x\| \|x^*\|$. From these calculations we conclude that in a C^* algebra the involution is an isometry.

Example. The algebra $\mathcal{B}(H)$ of bounded linear operators on a Hilbert space with the involution defined by taking the adjoint.

Example. Let X be a compact Hausdorff space. The algebra $C(X)$ of complex valued continuous linear functions on X with the sup norm and the involution given by $(f^*)(x) = \overline{f(x)}$ is a C^*-algebra.

Example. The algebra $\mathcal{K}(H)$ of compact operators on a Hilbert space H is a C^*-algebra. We know that it is a subalgebra of $\mathcal{B}(H)$, so all we have to check is that it is closed under taking the adjoint. Thus we have to show that the adjoint of a compact operator is compact.

Let T be compact and consider a sequence $y_n, n \geq 1$ in the unit ball $B_{0,1}$ centered at the origin of H. Let us prove that T^*y_n has a convergent subsequence. Define the functions $f_n : \overline{T(B_{0,1})} \to \mathbb{C}$,

$$f_n(x) = \langle x, y_n \rangle$$

where $\langle \cdot, \cdot \rangle$ is the inner product. Note that since T is compact, the domains of these functions are compact. Then

$$|f_n(x)| \leq \|x\|\|y_n\| \leq M.$$

So $f_n, n \geq 1$ is a bounded sequence. Also,

$$|f_n(x) - f_n(x')| \leq \|y_n\|\|x - x'\| \leq \|x - x'\|.$$

Thus for every $\epsilon > 0$, if $\|x - x'\| < \delta = \epsilon$, then $|f_n(x) - f_n(x')\| < \epsilon$ for all n, so the sequence f_n is also equicontinuous. By the Arzela-Ascoli theorem, f_n has a convergent subsequence in $C(\overline{T(B_{0,1})})$. Note also that

$$\|f_n\| = \sup\{|f_n(x)| \,|\, x \in \overline{T(B_{0,1})}\} = \sup\{|\langle Tx, y_n \rangle| \,|\, x \in B_{0,1}\} = \sup\{|\langle x, T^*y_n \rangle| \,|\, x \in B_{0,1}\} = \|T^*y_n\|.$$

So we have a sequence of linear functionals that converges in norm, and the limit is also a linear functional. Thus T^*y_n has a norm convergent subsequence, showing that T^* is compact.

We obtain that compact operators form a C^*-algebra. Moreover, $\mathcal{B}(H)/\mathcal{K}(X)$, the quotient of all operators module compact operators is a C^*-algebra.

Definition. If $\mathcal{B}$ and $\mathcal{B}'$ are C^*-algebras then $f : \mathcal{B} \to \mathcal{B}'$ is called a homomorphism if it is an algebra homomorphism and $f(a^*) = f(a)^*$ for all a.

An element a is called self-adjoint if $a = a^*$, normal if $aa^* = a^*a$ and unitary if $aa^* = a^*a = 1$. A first observation is that

$$\sigma_{\mathcal{B}}(u^*) = \overline{\sigma_{\mathcal{B}}(u)}.$$

We also have the following result.

Theorem 5.1.1. In a C^*-algebra the spectrum of a unitary element is contained in the unit circle, and the spectrum of a self-adjoint element is contained in the real axis.

Proof. If u is unitary, then $1 = \|1\| = \|u^*u\| = \|u^2\|$, so $\|u\| = \|u^*\| = \|u^{-1}\| = 1$. Then if $|\lambda| > 1$, then $\lambda - u$ is invertible. So $\sigma_{\mathcal{B}}(u) \subset \{\lambda \,|\, |\lambda| \geq 1\}$. But then also $\sigma_{\mathcal{B}}(u^{-1}) \subset \{\lambda \,|\, |\lambda| \geq 1\}$, because u^{-1} is also unitary. Hence by the Spectral Mapping Theorem, $\sigma_{\mathcal{B}}(u) \subset \{\lambda \in \mathbb{C} \,|\, |\lambda| \geq 1\}$. Taking the intersection, we find that $\sigma_{\mathcal{B}}(u)$ is in the unit circle.

Let $\mathcal{B}$ be the C^*-algebra. If $a \in \mathcal{B}$ is self-adjoint, then $u = \exp(ia)$ is unitary. Indeed, $u^* = \exp(ia)^* = \exp(-ia)$, and

$$uu^* = \exp(ia)\exp(-ia) = \exp(ia - ia) = 1 = u^*u.$$

Because $\sigma_{\mathcal{B}}(u)$ is a subset of the unit disk, and, by the Spectral Mapping Theorem, $\sigma(u) = \exp(i\sigma(a))$, the spectrum of a must be real. □

5.2 Commutative C^*-algebras

Theorem 5.2.1. (Gelfand-Naimark) If $\mathcal{B}$ is a commutative C^*-algebra and $M_{\mathcal{B}}$ is the set of multiplicative functionals on $\mathcal{B}$, then the Gelfand transform is a $*$-isometrical isomorphism of $\mathcal{B}$ onto $C(M_{\mathcal{B}})$.

Proof. Let us show that Γ is a $*$-map. If $a \in \mathcal{B}$, then $b = \frac{1}{2}(a + a^*)$ and $c = \frac{1}{2i}(a - a^*)$ are self-adjoint operators such that $a = b + ic$ and $a^* = b - ic$. Recall that $\sigma_{\mathcal{B}}(b)$ and $\sigma_{\mathcal{B}}(c)$ are subsets of $\mathbb{R}$, by Theorem 5.1.1. By Proposition 4.5.5, the functions $\Gamma(b)$ and $\Gamma(c)$ are real valued. Hence

$$\overline{\Gamma(a)} = \overline{\Gamma(b) + i\Gamma(c)} = \Gamma(b) - i\Gamma(c) = \Gamma(a^*).$$

This shows that Γ is a homomorphism of C^*-algebras.

Let us show that it is an isometry. We have

$$\|a\|^2 = \|a^*a\| = \|(a^*a)^{2^n}\|^{1/2^n} = \lim_{n\to\infty} \|(a^*a)^{2^n}\|^{1/2^n} = r_{\mathcal{B}}(a^*a).$$

By Proposition 4.5.5, this is equal to the sup norm of $\Gamma(a^*a)$. We have

$$\|\Gamma(a^*a)\| = \|\Gamma(a^*)\Gamma(a)\| = \|\Gamma(a)^2\| = \|\Gamma(a)\|^2.$$

Hence $\|a\| = \|\Gamma(a)\|$, as desired.

Finally, if ϕ and ψ are multiplicative functionals, then $\Gamma(a)(\phi) = \Gamma(a)(\psi)$ for all a means that $\phi(a) = \psi(a)$ for all a, hence $\phi = \psi$. This shows that the functions in the image of Γ separate points. The image contains the identity function, and for each function it contains its complex conjugate. So by the Stone-Weierstrass theorem, they are all continuous functions on $M_{\mathcal{B}}$. □

Theorem 5.2.2. (The Spectral Theorem) If H is a Hilbert space and N is a normal operator on H, then the C^*-algebra $\mathcal{C}_N$ generated by N and N^* is commutative. Moreover, the maximal ideal space of $\mathcal{C}_N$ is homeomorphic to $\sigma(N)$ and hence the Gelfand map is a $*$-isometrical isomorphism of $\mathcal{C}_N$ onto $C(\sigma(N))$.

Proof. The algebra $\mathcal{C}_N$ is commutative because it is the closure of the algebra of all polynomials in N and N^*.

Let us show that the set of multiplicative functionals, $M_{\mathcal{C}_N}$, is homeomorphic to $\sigma(N)$. In view of Proposition 4.5.5, we can define the onto function

$$\Psi : M_{\mathcal{C}_N} \to \sigma(N), \quad \Psi(\phi) = \Gamma(N)(\phi).$$

This function is also one-to-one, because if $\Psi(\phi) = \Psi(\phi')$, then

$$\phi(N) = \Gamma(N)(\phi) = \Gamma(N)(\phi') = \phi'(N),$$

and also

$$\phi(N^*) = \Gamma(N^*)(\phi) = \overline{\Gamma(N)(\phi)} = \overline{\Gamma(N)(\phi')} = \Gamma(N^*)(\phi') = \phi'(N^*).$$

Hence ϕ and ϕ' coincide on $\mathcal{C}_N$, so they are equal.

Finally, let us show that Ψ is continuous. Let

$$B_{\lambda_0,r} = \{\lambda \in \sigma(N) \mid |\lambda - \lambda_0| < r\}.$$

Set $\phi_{\lambda_0} = \Psi^{-1}(\lambda_0)$. Then

$$\Psi^{-1}(B_{\lambda_0,r}) = \{\phi \in M_{\mathcal{C}_N} \mid |\phi(N) - \phi_{\lambda_0}(N)| < r\},$$

which is open in the weak* topology. Hence Ψ is continuous.

Because $M_{\mathcal{C}_N}$ and $\sigma(N)$ are compact Hausdorff spaces, Ψ is a homeomorphism. □

This theorem allows us to perform functional calculus with continuous functions on the spectrum of N. Note the particular case of self-adjoint operators.

5.3 C^*-algebras as algebras of operators

Definition. Given a C^*-algebra $\mathcal{B}$, a $*$-*representation* is a (continuous) C^*-homomorphism

$$\rho : \mathcal{B} \to \mathcal{B}(H),$$

for some Hilbert space H, that is non-degenerate in the sense that $\rho(a)x$ is dense when a ranges through a and x ranges through H. A vector x is called *cyclic* if the set $\{\rho(a)x \mid a \in \mathcal{B}\}$ is dense in H; in this case the representation is called cyclic.

Definition. A *state* on a C^*-algebra is a linear functional ϕ such that $\phi(a^*a) \geq 0$ for all a and $\phi(1) = 1$.

Proposition 5.3.1. If ϕ is a state, then
(a) (Cauchy-Schwarz) $|\phi(b^*a)|^2 \leq \phi(a^*a)\phi(b^*b)$;
(b) $\|\phi\| = 1$.

Proof. For (a), repeat the proof of Cauchy-Schwarz for inner products.

For (b), note that $\phi(1) = 1$ implies that $\|\phi\| \geq 1$. Thus it suffices to show that

$$|\phi(a)| \leq \|a\|, \text{ for all } a \in \mathcal{B}.$$

By Cauchy-Schwarz we have

$$\begin{aligned}|\phi(a)|^2 = |\phi(1^*a)| &\leq \phi(1^*1)\phi(a^*a) = \phi(1)\phi(a^*a)\\ &= \phi(a^*a) \leq \phi(\|a^*a\|) = \|a^*a\| = \|a\|^2\end{aligned}$$

because $\|a^*a\| \geq a^*a$. □

Theorem 5.3.1. (The Gelfand-Naimark-Segal Construction) Given a state ϕ of $\mathcal{B}$, there is a $*$-representation $\rho : \mathcal{B} \to \mathcal{B}(H)$ which is cyclic, and a cyclic vector x such that

$$\phi(a) = \langle \rho(a)x, x\rangle \quad \text{for all } a \in \mathcal{B}.$$

Proof. Let $a \in \mathcal{B}$ act on the left on $\mathcal{B}$ by

$$\rho_\phi(a)b = ab.$$

This is the left regular representation. We want this to be a representation on a Hilbert space, and for that reason we attempt to turn $\mathcal{B}$ into a Hilbert space. We define the inner product by

$$\langle a, b\rangle = \phi(b^*a).$$

This has all the nice properties of an inner product, except that it might be degenerate, in the sense that there might be a such that $\langle a, a\rangle = \phi(a^*a) = 0$. Adapting the Cauchy-Schwarz inequality, we deduce that the set N_ϕ of elements a such that $\langle a, a\rangle = 0$ form a subspace of $\mathcal{B}$.

Let us show that N_ϕ is also a left ideal of $\mathcal{B}$. This is because of the Cauchy-Schwarz inequality:

$$|\phi((a^*b^*ba)|^2 \leq \phi(a^*a)\phi((b^*ba)(a^*b^*b)) = 0.$$

Then $\mathcal{B}/N_\phi$ is an inner product space. Consider the completion H_ϕ of this space, which is therefore a Hilbert space. We have

$$\|a\|^2b^*b - b^*a^*ab = b^*(\|a\|^2 - a^*a)b = b^*c^*cb,$$

where

$$c = c^* = (\|a\|^2 - a^*a)^{1/2} = (\|a^*a\| - a^*a)^{1/2}.$$

The element c can by defined because the function $f(t) = (\|a^*a\| - a^*a)^{1/2}$ is continuous on $\sigma(a^*a)$, so we can use Theorem 5.2.2. So, because ϕ is positive,

$$\phi(b^*a^*ab) \leq \phi(\|a\|^2b^*b) = \|a\|^2\phi(b^*b).$$

It follows that

$$\|a(b + N_\phi)\|_{H_\phi} \leq \|a\|^2\|b + N_\phi\|_{H_\phi},$$

so $\rho_\phi(a)$ is continuous. This implies that $\rho_\phi(a)$ can be extended to the entire Hilbert space H_ϕ. This representation is cyclic, with cyclic vector $1 + N_\phi$. Also, $\langle \rho_\phi(a)1, 1\rangle = \phi(1^*a1) = \phi(a)$. □

The set of states is a weak* closed convex subset of the unit ball of $\mathcal{B}^*$. The extremal points are called pure states.

Theorem 5.3.2. (Gelfand-Naimark) Every C^*-algebra admits an isometric $*$-representation. If the C^*-algebra is separable, then the Hilbert space can be chosen to be separable as well.

Proof. Consider the set of pure states and define

$$\rho : \mathcal{B} \to \oplus_\phi \mathcal{B}(H_\phi), \quad \rho = \oplus \rho_\phi$$

where the sum is taken over all pure states. It suffices to show that ρ is faithful, namely one-to-one, because the fact that it is an isometric $*$-homomorphsm then follows from Theorem 5.4.3 proved in next section.

To prove that ρ is injective, let a be a nonzero element of $\mathcal{B}$. Then there is a state ϕ such that $\phi(a^*a) > 0$. Indeed, consider a real-valued linear functional $\psi : \mathbb{R}(a^*a) \to \mathbb{R}$ such that $\psi_0(a^*a) > 0$. Now $\mathcal{B}$ is also a real vector space, in which the positive elements form a cone (an element is positive if it is of the form a^*a). Use the theorem of M. Riesz about extension of positive functionals to extend ψ_0 to $\psi : \mathcal{B}_{sa} \to \mathbb{R}$, where $\mathcal{B}_{sa}$ consists of the self-adjoint elements. Now define

$$\phi_1(a) = c\left(\psi\left(\frac{a+a^*}{2}\right) + i\left(\psi(\frac{a-a^*}{2i}\right)\right),$$

where c is chosen so that $\phi_1(1) = 1$. Then ϕ_1 is positive and $\phi_1(a^*a) > 0$. By the Krein-Milman theorem, there is a pure state ϕ_1 that satisfies $\phi_1(a^*a) > 0$, because the closure of the convex hull of the pure states is the set of all states, so if all pure states are zero on a^*a, then all states are zero.

Consider the GNS representation associated to this pure state ϕ, and let x be its cyclic vector. Then

$$\|\rho_\phi(a)x\|^2 = \langle \rho_\phi(a)x, \rho_\phi(a)x\rangle = \langle \rho_\phi(a^*a)x, x\rangle = \phi(a^*a) > 0.$$

In this case $\rho_\phi \neq 0$, hence the representation is faithful. The theorem is proved. □

5.4 Functional calculus for normal operators

Throughout this section we assume that the Hilbert space H is separable.

Definition. Let H be a Hilbert space. The *weak operator topology* is the topology defined by the open sets

$$V(T_0; x_1, x_2, \ldots, x_k; y_1, y_2, \ldots, y_k; r) = \{T \in \mathcal{B}(H) \mid |\langle (T - T_0)x_j, y_j\rangle| < r,\ j = 1, 2, \ldots, k\}.$$

The *strong operator topology* is the topology defined by the open sets

$$V(T_0; x_1, x_2, \ldots, x_k; r) = \{T \in \mathcal{B}(H) \mid \|(T - T_0)x_j\| < r,\ j = 1, 2, \ldots, k\}.$$

Definition. A von Neumann algebra is a C^*-subalgebra of $\mathcal{B}(H)$ that is weakly closed.

Remark 5.4.1. If $\mathcal{C}$ is a self-adjoint subalgebra of $\mathcal{B}(H)$, then its weak closure is a von Neumann algebra. If $\mathcal{C}$ is commutative, then its closure is commutative.

Proposition 5.4.1. If N is a normal operator on H, then the von Neumann algebra $\mathcal{W}_N$ generated by N is commutative. If $M_{\mathcal{W}_N}$ is the set of multiplicative functionals on $\mathcal{W}_N$, then the Gelfand transform is a $*$-isometrical isomorphism of $\mathcal{W}_N$ onto $C(M_{\mathcal{W}_N})$.

We want to show that there is a unique $*$-isometrical isomorphism $\Gamma^* : \mathcal{W}_N \to L^\infty(\sigma(N))$ which extends the functional calculus with continuous functions defined by the Spectral Theorem (Theorem 5.2.2).

Assume we have a *finite positive regular Borel measure* on $\sigma(N)$. We can assume that the measure of the entire space is 1, so that we have a probability measure. For the moment, we work in this hypothesis.

The map $f \mapsto M_f$, where $M_f : L^2(\sigma(N)) \to L^2(\sigma(N))$ $M_f g = fg$ identifies $L^\infty(\sigma(N))$ with a maximal commutative von Neumann subalgebra of the algebra of operators on $L^2(\sigma(N))$.

Proposition 5.4.2. The weak operator topology and the weak* topology on L^∞ coincide.

Proof. $L^\infty = L_1^*$, and recall that every function in L^1 is the product of two L^2 functions. Thus an element of the form

$$\phi(f) = \int fg$$

with $f \in L^\infty$ and $g \in L^1$, can also be represented as

$$\int M_f g_1 \overline{g}_2 = \langle M_f g_1, g_2 \rangle$$

where $g = g_1 g_2$. Hence the conclusion. □

Proposition 5.4.3. The space $C(\sigma(N))$ is weak*-dense in $L^\infty(\sigma(N))$.

Proof. We will show that the unit ball in $C(\sigma(N))$ is weak*-dense in the unit ball in $L^\infty(\sigma(N))$. Consider a step function in the unit ball of L^∞, $f = \sum \alpha_j \chi_{E_j}$, $|\alpha_j| \leq 1$ with E_j disjoint and their union is $\sigma(N)$). For each j, choose $K_j \subset E_j$. Using Tietze's Extension Theorem we can find g in the unit ball of $C(\sigma(N))$ such that $g(x) = \alpha_j$ for $x \in K_j$. Then for $h \in L^1$,

$$\begin{aligned} \left| \int h(f-g) \right| &\leq \int |h||f-g| \\ &= \sum_{j=1}^n \int_{E_j \setminus K_j} |h||f-g| \leq \sum_{j=1}^n \int_{E_j \setminus K_j} |h| \end{aligned}$$

Because the measure is regular, we can choose K_j such that the integrals are as small as desired. □

Recall that a vector x is cyclic for an algebra $\mathcal{B} \subset \mathcal{B}(H)$ if $\mathcal{B}x$ is dense in H and separating if $Tx = 0$ implies $T = 0$. If $\mathcal{B}$ is commutative, then x cyclic implies x separating, because $Tx = 0$ implies $\mathcal{B}x \in \ker(T)$, hence $T = 0$.

Theorem 5.4.1. If N is a normal operator on H such that $\mathcal{C}_N$ has a cyclic vector, then there is a positive regular Borel measure ν supported on $\sigma(N) = M_{\mathcal{C}_N}$ and an isometrical isomorphism γ from H onto $L^2(\sigma(N), \nu)$ such that the map

$$\Gamma^* : \mathcal{W}_N \to \mathcal{B}(L^2(\sigma(N), \nu), \quad \Gamma^*(T) = \gamma T \gamma^{-1}$$

is a $*$-isometrical isomorphism from $\mathcal{W}_N$ onto $L^\infty(\sigma(N), \nu)$. Moreover, Γ^* is an extension of the Gelfand transform $\Gamma : \mathcal{C}_T \to C(\sigma(N))$. Lastly, if ν_1 is a positive regular Borel measure on $\sigma(N)$ and Γ_1^* us a $*$-isometrical isomorphism from $\mathcal{W}_N$ that extends Γ, then ν and ν_1 are mutually absolutely continuous, $L^\infty(\sigma(N), \nu) = L^\infty(\sigma(N), \nu_1)$ and $\Gamma_1^* = \Gamma^*$.

Proof. Let x be a cyclic vector for $\mathcal{C}_N$ with $\|x\| = 1$. Consider the linear functional on $C(\sigma(N))$ defined by $\phi(f) = \langle f(N)x, x\rangle$. Then ϕ is positive because if $f \geq 0$ then $f = g^2$ for some real valued function g, and then

$$f(N) = g(N)g(N) = \overline{g}(N)g(N) = (g(N))^* g(N),$$

and hence

$$\langle f(N)x, x\rangle = \langle g(N)x, g(N)x\rangle = \|g(N)x\|^2 \geq 0.$$

We also have

$$|\phi(f)| = |\,\langle f(N)x, x\rangle\,| \leq \|f(N)\|\|x^2\| = \|f\|,$$

thus ϕ is continuous. By the Reisz Representation Theorem (Theorem 2.1.2), there is a unique positive regular measure ν on $\sigma(N)$ such that

$$\int_{\sigma(N)} f d\nu = \langle f(N)x, x\rangle \text{ for } f \in C(\sigma(N)).$$

If the support of ν were not the entire spectrum, then, by Urysohn's lemma, we could find a continuous function f that is 1 somewhere on the spectrum and is zero on the support of ν. Then because f is not identically equal to zero, $f(N) \neq 0$ and because x is separating, we have

$$0 \neq \|f(N)x\|^2 = \langle f(N)x, f(N)x\rangle = \left\langle |f|^2(N)x, x\right\rangle = \int_{\sigma(N)} |f|^2 d\nu = 0,$$

impossible. So $\operatorname{supp}(\nu) = \sigma(N)$.

Define

$$\gamma_0 : \mathcal{C}_N x \to L^2(\sigma(N), \nu), \quad \gamma_0(f(N)x) = f.$$

The computation

$$\|f\|_2^2 = \int_{\sigma(N)} |f|^2 d\nu = \langle |f|^2(N)x, x\rangle = \|f(N)x\|^2$$

shows that γ_0 is a Hilbert space isometry. Because $\mathcal{C}_N$ is dense in H and $C(\sigma(N))$ is dense in $L^2(\sigma(N),\nu)$, γ_0 can be extended uniquely to an isometrical isomorphism

$$\gamma : H \to L^2(\sigma(N),\nu).$$

Moreover, if we define

$$\Gamma^* : \mathcal{W}_N \to \mathcal{B}(L^2(\sigma(N),\nu)), \quad \Gamma^*(T) = \gamma T \gamma^{-1}$$

then Γ^* is a $*$-isometrical isomorphism onto the image.

Let us show that Γ^* extends the Gelfand transform

$$\Gamma : \mathcal{C}_N \to \mathcal{B}(L^2(\sigma(N),\nu)).$$

Indeed, if $f \in C(\sigma(N))$, then for all $g \in C(\sigma(N))$,

$$[\Gamma^*(f(N))]g = \gamma f(N)\gamma^{-1}g = \gamma f(N)g(N)x = \gamma[(fg)(N)x] = fg = M_f g.$$

Since $C(\sigma(N))$ is dense in $L^2(\sigma(N),\nu)$, it follows that

$$\Gamma^*(f(N)) = M_f = \Gamma(f(N)).$$

Because the weak operator topology and the weak* topology coincide on L^∞ (Proposition 5.4.2), Γ^* is a continuous map from $\mathcal{W}_T$ with the weak operator topology to $L^\infty(\sigma(N),\nu)$ with the weak* topology. And because continuous functions are weak*-dense in L^∞, it follows that $\Gamma^*(\mathcal{W}_T) = L^\infty(\sigma(N),\nu)$. Thus Γ^* is a $*$-isometrical isomorphism mapping $\mathcal{W}_T$ onto $L^\infty(\sigma(N),\nu)$.

Finally, if (ν_1,Γ_1) are a different pair with the above properties, then $\Gamma^*\Gamma_1^{*-1}$ is a $*$-isometrical isomorphism from $L^\infty(\sigma(N),\nu_1)$ onto $L^\infty(\sigma(N),\nu)$ which is the identity on $C(\sigma(N))$. Then ν and ν_1 are mutually absolutely continuous, $L^\infty(\sigma(N),\nu) = L^\infty(\sigma(N),\nu_1)$ and $\Gamma^*\Gamma_1^{*-1}$ is the identity map. This completes the proof. □

However, not all operators have cyclic vectors. Instead we will use separating vectors and replace H by the smallest invariant subspace containing a separating vector. We proceed to show that every normal operator has a separating vector.

An easy application of Zorn's lemma yields the following result.

Proposition 5.4.4. Every commutative C^*-algebra is contained in a maximal commutative von Neumann algebra.

Definition. If $\mathcal{A} \subset \mathcal{L}(H)$, then the *commutant* of $\mathcal{A}$, denoted $\mathcal{A}'$, is the set of operators in $\mathcal{L}(H)$ which commute with every operator in $\mathcal{A}$.

Proposition 5.4.5. A C^*-algebra in $\mathcal{L}(H)$ is a maximal commutative von Neumann algebra if and only if it is equal to its own commutant.

Proof. Let $\mathcal{B}$ be the C^*-algebra. If $\mathcal{B}$ is commutative, then $\mathcal{B} \subset \mathcal{B}'$. If $\mathcal{B}$ is maximal commutative then necessarily we have equality.

Conversely, if equality holds, then $\mathcal{B}$ is a von Neumann algebra. It must be maximal commutative, for if A commutes with everything in $\mathcal{B}$, then $A \in \mathcal{B}' = \mathcal{B}$. □

Lemma 5.4.1. Let $T \in \mathcal{B}(H)$, V is a closed subspace of H, and P_V the orthogonal projection onto V. Then $P_V T = T P_V$ if and only if V is an invariant subspace for both T and T^*. Moreover, in this case both V and $V^\perp$ are invariant for T.

Proof. V is invariant for T if and only if $P_V T P_V = T P_V$. So V is invariant for both T and T^* if and only if $P_V T P_V = T P_V$ and $P_V T^* P_V = T^* P_V$. The latter is equivalent, by conjugating to $P_V T P_V = P_V T$. So V is invariant for both T and T^* if and only if $P_V T = T P_V$ (in which case the equality to $P_V T P_V$ is superfluous). Note also that V invariant for T^* implies $V^\perp$ invariant for T (by the equality $\langle Tx, y\rangle = \langle x, T^* y\rangle$). □

Definition. A subspace V of H is a reducing subspace for T if it satisfies any of the equivalent conditions from the statement of the above lemma.

Lemma 5.4.2. If $\mathcal{B}$ is a C^*-algebra contained in $\mathcal{B}(H)$ and $v \in H$, then the orthogonal projection onto $\overline{\mathcal{B}v}$ is in $\mathcal{B}'$.

Proof. By Lemma 5.4.1, it suffices to show that $\overline{\mathcal{B}v}$ is invariant for both T and T^* for every $T \in \mathcal{B}$. Note that $T^* \in \mathcal{B}$ so both T and T^* leave $\mathcal{B}v$ invariant, and so they do with $\overline{\mathcal{B}v}$. □

Theorem 5.4.2. If $\mathcal{B}$ is a maximal commutative von Neumann algebra on a separable Hilbert space H, then $\mathcal{B}$ has a cyclic vector.

Proof. Let $\mathcal{E}$ be the set of all collections of projections $\{E_\alpha\}_{\alpha\in A}$ in $\mathcal{B}$ such that

- For each $\alpha \in A$ there is $v_\alpha \in H\backslash\{0\}$ so that E_α is the projection onto $\overline{\mathcal{B}v_\alpha}$,
- $E_\alpha E_{\alpha'} = E_{\alpha'} E_\alpha = 0$ for $\alpha \neq \alpha'$.

Clearly $\mathcal{E}$ is not empty, since we can build an element in $\mathcal{E}$ starting with one vector, via Lemma 5.4.2. Order $\mathcal{E}$ by inclusion. The hypothesis of Zorn's Lemma is satisfied. Pick a maximal element $\{E_\alpha\}_{\alpha\in A}$.

Let $\mathcal{F}$ be the collection of all finite subsets of the index set A partially ordered by inclusion and let $\{P_F\}_{F\in\mathcal{F}}$ be the net of the orthgonal projections defined by

$$P_F = \sum_{\alpha\in\mathcal{F}} E_\alpha.$$

Then the net is increasing. If $F > F'$ then

$$\|(P_F - P_{F'})x\|^2 = \left\langle (P_F - P_{F'})^2 x, x\right\rangle = \langle (P_F - P_{F'})x, x\rangle = \langle P_F x, x\rangle - \langle P_{F'} x, x\rangle .$$

The net $\langle P_F x, x\rangle$ is increasing and bounded from above by $\|x\|^2$, so it is convergent. Hence it is Cauchy, and so is $\|P_F x\|$. Then $P_F x$ is norm convergent. Define $Px = \lim_F P_F x$. Then P is an orthogonal projection.

The range V of P has the property that both V and $V^\perp$ are invariant under $\mathcal{B}$. Moreover $P \in \mathcal{B}$ by Lemma 5.4.2. Note that if $v \in V^\perp$, then $\overline{\mathcal{B}v}$ is orthogonal to each E_α so we can add the projection onto this space to the family $\{E_\alpha\}_\alpha$, contradicting maximality. Hence $V^\perp = 0$, showing that P is the identity map.

Because H is separable, A is countable. Thus we can define $w = \sum_\alpha v_\alpha$. Then for each α, the range of E_α is contained in $\overline{\mathcal{B}w}$. So w is a cyclic vector for $\mathcal{B}$. □

Corollary 5.4.1. Every commutative C^*-algebra of operators on a separable Hilbert space has a separating vector.

Proof. Include the C^* algebra into a maximal one. The new algebra has a vector that is cyclic hence separable. This vector is also separable for the subalgebra. □

Theorem 5.4.3. If $\Phi : \mathcal{B}_1 \to \mathcal{B}_2$ is a $*$-homomorphism of C^*-algebras, then $\|\Phi\| \leq 1$ and Φ is an isometry if and only if it is one-to-one.

Proof. If $a \in \mathcal{B}_1$ and $a = a^*$ (i.e. a is self-adjoint), then $\mathcal{C}_a$ is a commutative C^*-algebra contained in $\mathcal{B}_1$ and $\overline{\Phi(\mathcal{C}_a)}$ is a commutative C^*-algebra contained in $\mathcal{B}_2$. If ϕ is a multiplicative linear functional on $\overline{\Phi(\mathcal{C}_a)}$, then $\phi \circ \Phi$ is a multiplicative linear functional on $\mathcal{C}_a$. Because of the Gelfand-Naimark Theorem, we can choose ϕ so that $|\phi(\Phi(a))| = \|\Phi(a)\|$. Then

$$\|a\| \geq |\phi(\Phi(a))| = \|\Phi(a)\|,$$

so Φ is a contraction on self-adjoint elements. For arbitrary $b \in \mathcal{B}_1$,

$$\|b\|^2 = \|b^*b\| \geq \|\Phi(b^*b)\| = \|\Phi(b)^*\Phi(b)\| = \|\Phi(b)\|^2.$$

Hence $\|\Phi\| \leq 1$.

For the second part, clearly if Φ is an isometry than it is one-to-one. Assume that Φ is not an isometry and choose b such that $\|b\| = 1$ but $\|\Phi(b)\| < 1$. Set $a = b^*b$; then $\|a\| = 1$ but $\|\Phi(a)\| = 1 - \epsilon$ with $\epsilon > 0$. Choose a function $f \in C([0, 1])$ such that $f(1) = 1$ and $f(x) = 0$ if $0 \leq x \leq 1 - \epsilon$. Using the functional calculus on $\mathcal{C}_a$, define $f(a)$. Since

$$\sigma(f(a)) = \mathrm{im}(\Gamma(f(a)) = f(\sigma(a)),$$

we conclude that $1 \in \sigma(f(a))$, so $f(a) \neq 0$. We have $\Phi(f(a)) = f(\Phi(a))$ (true on polynomials, then pass to the limit). But $\|\Phi(a)\| = 1 - \epsilon$, so $\sigma(\Phi(a)) \subset [0, 1 - \epsilon]$. But then $\sigma(f(\Phi(a)) = f(\sigma(\Phi(a)) = 0$ so $f(\Phi(a)) = 0$. Hence $\Phi(f(a)) = f(\Phi(a)) = 0$. Thus Φ is not one-to-one. □

Let H be a separable Hilbert space and N normal on H. By Corollary 5.4.1, the commutative von Neumann algebra $\mathcal{W}_N$ has a separating vector x. If we set $H_x = \overline{\mathcal{W}_N x}$, then both H_x and $H_x^\perp$ are invariant under $\mathcal{W}_N$. We can therefore define a map $\Phi : \mathcal{W}_N \to \mathcal{B}(H_x)$ by $\Phi(N) = N|H_x$.

Lemma 5.4.3. The map Φ defined above is a $*$-isometrical isomorphism. Moreover $\sigma_{\mathcal{B}(H)}(T) = \sigma_{\mathcal{B}(H_x)}(T|H_x)$ for all $T \in \mathcal{W}_N$.

Proof. In view of the previous theorem, let us show that Φ is one-to-one. And indeed, if $\Phi(T) = 0$ then $Tx = 0$, because $x = Ix \in \mathcal{W}_N x$. So $T = 0$, because x is separating of $\mathcal{W}_N$. The equality of spectra is proved as follows.

First,

$$\sigma_{\mathcal{B}(H)}(T) = \sigma_{\mathcal{W}_N}(T).$$

Indeed, $\sigma_{\mathcal{B}(H)}(T) \subset \sigma_{\mathcal{W}_N}(T)$ because the inverse of $\lambda - T$ might or might not be in $\mathcal{W}_N$. Moreover, because the resolvent is open both for $\mathcal{B}(H)$ and for $\mathcal{W}_N$, $\sigma_{\mathcal{W}_N}(T)$ is obtained from $\sigma_{\mathcal{B}(H)}(T)$ by adding to it some bounded components of its complement. So if $T - \lambda$ is invertible in $\mathcal{B}(H)$, then $(T - \lambda)(T^* - \bar{\lambda})$ is self-adjoint, so its spectrum is real and hence necessarily the same in $\mathcal{B}(H)$ and $\mathcal{W}_N$. So this operator must be invertible in $\mathcal{W}_N$, and hence so is $T - \lambda$. Next

$$\sigma_{\mathcal{W}_N}(T) = \sigma_{\Phi(\mathcal{W}_N)}(T|H_x)$$

because Φ is a $*$-isometrical isomorphism onto the image. Repeating the above argument we also have

$$\sigma_{\Phi(\mathcal{W}_N)}(T|H_x) = \sigma_{\mathcal{B}(H_x)}(T|H_x)$$

and we are done. □

Theorem 5.4.4. (Functional Calculus for Normal Operators - Version I) Let N be a normal operator on the separable Hilbert space H and let $\Gamma : \mathcal{C}_N \to C(\sigma(N))$ be the Gelfand transform. Then there is a positive regular Borel measure ν having support $\sigma(N)$ and a $*$-isometrical isomorphism Γ^* from $\mathcal{W}_N$ onto $L^\infty(\sigma(N), \nu)$ which extends Γ. Moreover ν is unique up to mutual absolute continuity while $L^\infty(\sigma(N), \nu)$ and Γ^* are unique.

Proof. Let x be a separating vector for $\mathcal{W}_N$, $H_x = \overline{\mathcal{W}_N x}$, and

$$\Phi_x : \mathcal{W}_N \to \mathcal{B}(H_x), \quad \Phi_x(T) = T|H_x.$$

Let $\mathcal{W}_x$ be the von Neumann algebra generated by $N|H_x$. The map Φ is continuous in the weak operator topology (because it is obtained by restricting the domain). Hence $\Phi(\mathcal{W}_N) \subset \mathcal{W}_x$. Moreover, if

$$\Gamma_0 : \mathcal{C}_{N|H_x} \to C(\sigma(N|H_x)) = C(\sigma(N))$$

is the Gelfand transform, then $\Gamma = \Gamma_0 \circ \Phi$.

Because $N|H_x$ is normal and has the cyclic vector x, by Theorem 5.4.1 there is a positive regular Borel measure ν with support $\sigma(N|H_x) = \sigma(N)$ (here we use the previous lemma), and a $*$-isometrical (onto) isomorphism

$$\Gamma_0^* : \mathcal{W}_x \to L^\infty(\sigma(N), \nu), \text{ such that } \Gamma_0^*|\mathcal{C}_{N|H_x} = \Gamma_0.$$

Moreover, Γ_0^* is continuous form the weak operator topology of $\mathcal{W}_x$ to the weak*-topology on $L^\infty(\sigma(N), \nu)$. Hence $\Gamma^* = \Gamma_0^* \circ \Phi$ is a $*$-isometrical isomorphism from $\mathcal{W}_N$ into $L^\infty(\sigma(N), \nu)$, continuous in the weak/weak* topologies, and which extends the Gelfand transform.

The only thing that remains to show is that Γ_* takes $\mathcal{W}_N$ onto $L^\infty(\sigma(N), \nu)$. For this we need the following result.

Lemma 5.4.4. Let H be a Hilbert space. Then the unit ball of $\mathcal{B}(H)$ is compact in the weak operator topology.

Proof. The proof is from the book of Kadison and Ringrose, *Fundamentals of the theory of operator algebras.* For two vectors $x, y \in H$, let $D_{x,y}$ be the closed disk of radius $\|x\| \cdot \|y\|$ in the complex plane. The mapping which assigns to each $T \in (\mathcal{B}(H))_1$ the point

$$\{\langle Tx, y\rangle \mid x, y \in H\} \subset \prod_{x,y} D_{x,y}$$

is a homeomorphism of $(\mathcal{B}(H))_1$ with the weak operator topology onto its image X in the topology induced on X by the product topology of $\prod_{x,y} D_{x,y}$. As the latter is a compact Hausdorff topology by Tychonoff's theorem, X is compact if it is closed. So let us prove that X is closed.

Let $b \in \overline{X}$. Choose $x_1, y_1, x_2, y_2 \in H$. Then for every $\epsilon > 0$ there is $T \in (\mathcal{B}(H))_1$ such that each of

$$|a \cdot b(x_j, y_k) - a\,\langle Tx_j, y_k\rangle\,|, \quad |b(x_j, y_k) - \langle Tx_j, y_k\rangle\,|,$$
$$|b(ax_1 + x_2, y_j) - \langle T(ax_1 + x_2), y_j\rangle\,|, \quad |b(x_j, ay_1 + y_2) - \langle Tx_j, ay_1 + y_2\rangle\,|$$

is less than ϵ. It follows that

$$|b(ax_1 + x_2, y_1) - a \cdot b(x_1, y_1) - b(x_2, y_1)| < 3\epsilon$$
$$|b(x_1, ay_1 + y_2) - \overline{a} \cdot b(x_1, y_1) - b(x_1, y_2)| < 3\epsilon.$$

Thus

$$b(ax_1 + x_2, y_1) = a \cdot b(x_1, y_1) + b(x_2, y_1) \quad b(x_1, ay_1 + y_2) = \overline{a}b(x_1, y_1) + b(x_1, y_2).$$

Additionally, $|b(x, y)| \leq \|x\| \cdot \|y\|$. Hence b is a conjugate-bilinear functional on H bounded by 1. Using the Riesz Representation Theorem, we conclude that there is an operator T_0 such that $b(x, y) = \langle T_0 x, y\rangle$. This operator has norm at most 1 and we are done. □

Using the lemma, we obtain that the unit ball in $\mathcal{W}_N$ is compact in the weak operator topology. It follows that its image is weak*-compact in $L^\infty(\sigma(N), \nu)$, and hence weak*-closed. Since this image contains the unit ball of $C(\sigma(N))$, it follows that it contains the unit ball in $L^\infty(\sigma(N), \nu)$. Hence Γ^* takes the unit ball in $\mathcal{W}_N$ onto the unit ball of $L^\infty(\sigma(N), \nu)$. So Γ^* is onto.

The uniqueness is as in Theorem 5.4.1. We are done. □

Definition. If N is a normal operator and $\Gamma^* : \mathcal{W}_N \to L^\infty(\sigma(N), \nu)$ is the map constructed in the above theorem, then for each $f \in L^\infty(\sigma(N), \nu)$ we can define

$$f(N) = \Gamma^{*-1}(f).$$

The spectral measure of a normal operator is defined as follows. For each Borel set $\Delta \in \sigma(N)$, let

$$E(\Delta) = \Gamma^{*-1}(\chi_\Delta).$$

Because

$$\chi_\Delta^2 = \chi_\Delta = \overline{\chi_\Delta}$$

$E(\Delta)$ is an orthogonal projection. Moreover, if $\Delta \cap \Delta' = \emptyset$, then $E(\Delta)E(\Delta') = E(\Delta')E(\Delta) = 0$. Hence

$$E\left(\sqcup_{k=1}^\infty \Delta_k\right) = \sum_{k=1}^\infty E(\Delta_k).$$

We conclude that E is a projection-valued measure.

For every $x, y \in H$, $\mu_{x,y}(\Delta) = \langle E(\Delta)x, y\rangle$ is a genuine positive regular Borel measure. Thus we can define for each function $f \in C(\sigma(N))$ an operator $f(N)$ by

$$\langle f(N)x, y\rangle = \int_{\sigma(N)} f d\mu_{x,y}.$$

It turns out that $f(N)$ is the functional calculus defined by the Gelfand transform. In fact more is true.

Let $f : \sigma(N) \to \mathbb{C}$ be a measurable function. There is a countable collection of open disks, D_i, $i \geq 1$, that form a basis for the topology on $\sigma(N)$. Let V be the union of those disks D_i for which $\nu(f^{-1}(D_i)) = 0$. Then $\nu(f^{-1}(V)) = 0$. The complement of V is the essential range of f. We say that f is essentially bounded if its essential range is bounded.

Theorem 5.4.5. (Functional Calculus for Normal Operators - Version II) There is a $*$-isometrical isomorphism $\Psi : L^\infty(\sigma(N), \nu) \to \mathcal{W}_N$ which is onto, defined by the formula

$$\langle \Psi(f)x, y\rangle = \int_{\sigma(N)} f d\mu_{x,y}.$$

Moreover, $\Psi = \Gamma^{*-1}$.

Proof. Check on step functions, then use density. □

This justifies the notation

$$f(N) = \int_{\sigma(N)} f dE.$$

In particular,

$$N = \int_{\sigma(N)} t dE.$$

Example. Say

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{pmatrix}.$$

Then the spectrum of A is $\sigma(A) = \{5, -1\}$, with eigenvectors for 5: $\mathbf{u}_1 = (-\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}})$, and for -1: $\mathbf{u}_2 = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, $\mathbf{u}_3 = (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. Then the spectral measure is

$$E(\{5\})(x) =< x, \mathbf{u}_1 > \mathbf{u}_1, \quad E(\{-1\})(x) =< x, \mathbf{u}_2 > \mathbf{u}_2 + < x, \mathbf{u}_3 > \mathbf{u}_3.$$

And for a function $f : \{5, -1\} \to \mathbb{C}$, we have

$$f(A) = f(5)E(\{5\}) + f(-1)E(\{-1\}).$$

In particular $A = 5E(\{5\}) - 1E(-1)$, which is actually the diagonalization of A.

In general if N is a normal operator on $\mathbb{R}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ then the spectral measure associates to each eigenvalue the projection onto its eigenspace.

Theorem 5.4.6. (The spectral mapping theorem) The spectrum of $f(N)$ is the essential range of f.

Proposition 5.4.6. If N is normal and has spectral measure E_N, and if $f \in L^\infty(E_N)$, then $f(N)$ is also normal and the spectral measure of $f(N)$ is defined by $E_{f(N)}(\Delta) = E(f^{-1}(\Delta))$.

Example. If A is self-adjoint, then the spectral measure of A is supported on a compact subset of $\mathbb{R}$. If U is unitary, then the spectral measure of U is supported on a compact subset of the unit circle.

Example. Let $A_1, A_2, ..., A_n$ be self-adjoint operators that commute pairwise. Then there exists a self-adjoint operator A and functions $f_1, f_2, ..., f_n$ such that $A_j = f_j(A)$ for all j.

By repeating some of the self-adjoints, we can assume that n is a power of 2, say $n = 2^m$. Now let $N_1 = A_1 + iA_2$. Then N_1 is a normal operator. Let $\sigma(N_1)$ be its spectrum and E_{N_1} be its spectral measure. Consider a square S that covers the spectrum and the continuous surjective map $\phi : C \to S$, where C is the Cantor set that defines the Peano curve. Define the spectral measure $E_{B_1}(\Delta) = E_{N_1}(\phi(\Delta))$, and let $B_1 = \int_0^1 t dE_{B_1}$. Then $N_1 = \phi(B_1)$ and $A_1 = (\phi(B_1) + \phi(B_1)^*)/2$, $A_2 = (\phi(B_1) - \phi(B_1)^*)/2i$.

Moreover, $B_1 = \phi^{-1}(N_1)$. We can define analogously $B_2, B_3, ...$, and they all commute. Thus we have reduced the number of self-adjoint operators to half that many. Now we can reason inductively to get the conclusion.

Chapter 6

Distributions

6.1 The motivation for using test functions

Let me point out that most of this material is taken from W. Rudin, *Functional Analysis*, McGraw Hill.

Assume that f is a differentiable one variable function and that ϕ is a compactly supported smooth function. Then integration by parts gives

$$\int f'\phi = -\int f\phi'.$$

But then using the right-hand side we can define the derivative of "any" function f to be a function g that satisfies

$$\int g\phi = -\int f\phi'$$

for all compactly supported ϕ.

Here is a practical application of this philosophy. Recall that a function f is harmonic if $\Delta f = 0$. Here is a "weak" characterization of harmonicity.

Theorem 6.1.1. (Weyl's Lemma) Let U be an open subset of $\mathbb{R}^n$ and let $f \in L^2(U)$. Then

$$\int_U f\,\Delta\phi = 0$$

for all smooth functions ϕ with compact support in U if and only if f is harmonic, meaning that $\Delta f = 0$.

Proof. Without loss of generality we may assume that $U = B_1^n$ is the unit ball centered at the origin. We can extend all functions on B_1^n to the entire $\mathbf{R}^n$ by setting them equal to zero outside B_1^n.

First note that if f is at least twice differentiable, then by Green's theorem

$$\int_{B_1^n} f\,\Delta\phi - \int_{B_1^n} \phi\,\Delta f = \int_{\partial B_1^n} f\frac{\partial\phi}{\partial n} - \phi\frac{\partial f}{\partial n} = 0.$$

Hence for all smooth functions ϕ with compact support in B_1^n,

$$\int_{B_1^n} \Delta f \, \phi = 0,$$

showing that $\Delta f = 0$, namely that f is harmonic.

But f is not necessarily a C^2-function. We resolve this issue by convoluting f with a mollifier that turns it into a smooth function. To this end, consider the bump function

$$\rho(\mathbf{x}) = \left\{ \begin{array}{ll} Ce^{-1/(1-\|\mathbf{x}\|^2)} & \text{if } \|\mathbf{x}\| \leq 1 \\ 0 & \text{if } \|\mathbf{x}\| \geq 1, \end{array} \right\}$$

where $C = (\int_U e^{-1/(1-\|\mathbf{x}\|^2)} d\mathbf{x})^{-1}$. Define the family of mollifiers

$$\rho_\epsilon(\mathbf{x}) = \epsilon^{-n} \rho\left(\frac{\mathbf{x}}{\epsilon}\right), \quad \epsilon > 0.$$

These functions have integrals equal to 1, are supported in the ball B_ϵ^n of radius ϵ centered at the origin and converge, in distributional sense, to Dirac's delta function as $\epsilon \to 0$. Let us convolute f with these functions to obtain

$$f_\epsilon(\mathbf{x}) = (f * \rho_\epsilon)(\mathbf{x}) = \int_{\mathbf{R}^n} f(\mathbf{y}) \rho_\epsilon(\mathbf{x} - \mathbf{y}) d\mathbf{y}.$$

Since f is an L^2 function, f_ϵ is smooth because we can differentiate under the integral sign (using the Dominated Convergence Theorem). By Young's inequality for convolutions,

$$\|f_\epsilon\|_2 \leq \|\rho_\epsilon\|_1 \|f\|_2 = \|f\|_2. \tag{6.1.1}$$

Next, we will show that

$$\|f - f_\epsilon\|_2 \to 0, \text{ as } \epsilon \to 0. \tag{6.1.2}$$

If g is another function, then

$$\begin{aligned} \|f - f_\epsilon\|_2 &\leq \|f - g\|_2 + \|g - g_\epsilon\|_2 + \|g_\epsilon - f_\epsilon\|_2 \\ &= \|f - f_\epsilon\|_2 \leq \|f - g\|_2 + \|g - g_\epsilon\|_2 + \|(g - f)_\epsilon\|_2. \end{aligned}$$

Combining this fact with (6.1.1), we see that if the convergence property (6.1.2) is true for a sequence of functions that approximates f in L^2, then it is true for f itself. We can choose the function g that approximates f to be continuous and compactly supported in B_1^n. We have

$$\begin{aligned} (g - g_\epsilon)(\mathbf{x}) &= \int_{\mathbb{R}^n} (g(\mathbf{x}) - g(\mathbf{x} - \mathbf{y})) \rho_\epsilon(\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^n} (g(\mathbf{x}) - g(\mathbf{x} - \epsilon\mathbf{z})) \rho(\mathbf{z}) d\mathbf{z}. \end{aligned}$$

The integrand is compactly supported, tends pointwise to zero almost everywhere and is bounded from above by a constant that depends on ρ only (and not on ϵ). Hence by the Dominated Convergence Theorem $g - g_\epsilon$ converges to zero in L^2. This proves (6.1.2).

For $\epsilon < 1$, and $\mathbf{y}$ in the ball $B^n_{1-\epsilon}$, the function $\mathbf{x} \mapsto \rho_\epsilon(\mathbf{x} - \mathbf{y})$ is smooth and compactly supported in B^n_1. Using the hypothesis, we deduce that

$$\Delta f_\epsilon = \int_{\mathbb{R}^n} f(\mathbf{y}) \Delta \rho_\epsilon(\mathbf{x} - \mathbf{y}) d\mathbf{y} = 0.$$

This shows that f_ϵ is harmonic in $B^n_{1-\epsilon}$.

Now let $0 < \epsilon < \delta$. Then f_ϵ and f_δ are both harmonic in $B^n_{1-\delta}$. Let us examine $(f_\epsilon * \rho_\delta)(\mathbf{x})$. We have

$$(f_\epsilon * \rho_\delta)(\mathbf{x}) = \int_{B^n_{\mathbf{x},1-\delta}} f_\epsilon(\mathbf{y}) \rho_\delta(\mathbf{x} - \mathbf{y}) d\mathbf{y}$$

where $B^n_{\mathbf{x},1-\delta}$ is the ball of radius $1 - \delta$ centered at $\mathbf{x}$. Switching to spherical coordinates this integral becomes

$$\int_0^{1-\delta} \rho_\delta(r) \int_{S^{n-1}_r} f_\epsilon(\mathbf{z}) d\mathbf{z} dr,$$

because ρ_δ is constant on spheres centered at the origin. The inner integral is $A(S^{n-1}_r) f_\epsilon(\mathbf{x})$, by the Mean Value Theorem for harmonic functions. Thus the integral is equal to

$$f(\mathbf{x}) \int_0^{1-\delta} \rho_\delta(r) A(S^{n-1}_r) dr = \int_{\mathbb{R}^n} \rho_\delta(\mathbf{x}) d\mathbf{x} = 1.$$

It follows that $f_\epsilon = f_\epsilon * \rho_\delta$ on $B^n_{1-\delta}$. For the same reason $f_\delta = f_\delta * \rho_\epsilon$ on $B^n_{1-\delta}$. We conclude that for $\epsilon < \delta$,

$$f_\delta = f_\delta * \rho_\epsilon = f * \rho_\delta * \rho_\epsilon = f * \rho_\epsilon * \rho_\delta = f_\epsilon * \rho_\delta = f_\epsilon \text{ on } B^n_{1-\delta}.$$

It follows that we can define a function f_h such that $f_h = f_\epsilon$ on $B^n_{1-\epsilon}$ for all $0 < \epsilon < 1$. This function is harmonic in the unit ball. Moreover

$$\|f_h - f_\epsilon\|_2 \to 0 \text{ when } \epsilon \to 0.$$

Combining this with (6.1.2), we deduce that $f = f_h$, so f is harmonic. The theorem is proved. □

6.2 Test functions

Let $\Omega \in \mathbb{R}^n$ be open, and let $\mathcal{D}(\Omega)$ be the set of compactly supported smooth functions. For a multiindex α let $|\alpha|$ be the sum of its entries.

So we want to be able to define derivatives of functions: $D^\alpha f = g$, by using the formula

$$\int g\phi = (-1)^{|\alpha|} \int f D^\alpha \phi, \quad \text{for all } \phi \in \mathcal{D}(\omega).$$

More generally, we want to be able to define the derivative of a linear functional Λ on $\mathcal{D}(\Omega)$ by the formula

$$D^\alpha \Lambda(\phi) = (-1)^{|\alpha|}\Lambda(D^\alpha \phi),$$

with the functionals of the form $\phi \mapsto \int f\phi$ being a particular example. Or course, we can use this as the definition, and that would be the end of the story. But arbitrary linear functionals are not that useful, we need a notion of continuity. This notion of continuity should be well behaved with respect to differentiation.

It is easy to put a topology on the smooth functions that have a common support. Thus let K be a compact subset of $\mathbb{R}$, and let $\mathcal{D}(K)$ be the set of smooth functions with support in K. It becomes a Fréchet space if endowed with the seminorms

$$\|\phi\|_{N,K} = \sup\{|D^\alpha \phi(x)\| \,|\, x \in K, \|\alpha\| \leq N\}.$$

Let τ_K be the Fréchet space topology, which we know is metrizable (being Fréchet, the space is complete in this metric).

$\mathcal{D}(\Omega)$ is the union of all $\mathcal{D}(K)$, $K \subset \Omega$, so we need to put all the topologies τ_K together in some nice topology τ.

Definition. The topology τ on $\mathcal{D}(\Omega)$ consists of the unions of the sets of the form $\phi + W$, where $\phi \in \mathcal{D}(\Omega)$ and W is a convex balanced set in $\mathcal{D}(\Omega)$ whose intersection with any $\mathcal{D}(K)$ lies in τ_K.

Theorem 6.2.1. The sets W from the definition form a system of neighborhoods of the origin, and τ makes $\mathcal{D}(\Omega)$ into a locally convex topological vector space.

Proposition 6.2.1. There is a sequence $K_n, \in \geq 1$ of compact subsets of Ω such that $\Omega = \cup_{n=1}^\infty K_n$ and $K_n \subset \mathrm{int} K_{n+1}$.

Proof. Set

$$K_n = \{x \in \Omega \,|\, \|x\| \leq n\} \cap \{z \in \Omega \,|\, \|x - y\| \geq \frac{1}{n}, \text{ for all } y \in \mathbb{R}^n \backslash \Omega\}. \quad \square$$

Just as an observation, we can define τ using only the sets K_n because every compact K lies in one of the K_n, as it lies in the open cover $\cup_n \mathrm{int} K_n$, and the inclusion $\mathcal{D}(K) \to \mathcal{D}(K_n)$ is a homeomorphism onto the image.

Theorem 6.2.2. (a) A convex balanced set is open in τ if and only if its intersection with every $\mathcal{D}(K)$ is in τ_K. Moreover the subspace topology induced by τ on $\mathcal{D}(K)$ is τ_K.
(b) If $E \subset \mathcal{D}(\Omega)$ is bounded then $E \subset \mathcal{D}(K)$ for some K and there is a sequence $M_N, N \geq 1$, such that

$$\|\phi\|_{N,K} \leq M_N, \quad N \geq 1.$$

Consequently, if a sequence $\phi_n, n \geq 1$, in $\mathcal{D}(\Omega)$ has the property that for every open set V there is N such that if $m, n \geq N$, then $\phi_m - \phi_n \in V$, then x_n lies in some $\mathcal{D}(K)$ and is it is a Cauchy sequence (hence convergent) in this subspace.
(c) Every closed and bounded subset of $\mathcal{D}(\Omega)$ is compact.

Proof. Part (c) is a consequence of

Theorem 6.2.3. (Arzela-Ascoli) A set $\mathcal{F} \subset C(G, Y)$ is normal (i.e. its closure is compact) if and only if the following two conditions are satisfied:
(a) for each $z \in G$, $\{f(z) \mid f \in \mathcal{F}\}$ has compact closure in Y.
(b) $\mathcal{F}$ is equicontinuous at every point in G.

□

Theorem 6.2.4. Suppose Λ is a mapping of $\mathcal{D}(\Omega)$ into some locally convex subspace Y. TFAE:
(a) Λ is continuous.
(b) Λ is bounded.
(c) If $\phi_i \to 0$ in $\mathcal{D}(\Omega)$ then $\Lambda\phi_i \to 0$ in Y.
(d) The restrictions of Λ to every $\mathcal{D}(K)$ are continuous.

Appendix A

Background results

A.1 Zorn's lemma

Theorem A.1.1. Suppose a partially ordered set M has the property that every totally ordered subset has an upper bound in M. Then the set M contains at least one maximal element.

Remark A.1.1. This result is proved using the Axiom of Choice.