Vitali Milman

An Introduction To Functional Analysis

WORLD 1999

dedications

Contents

Chapter 1

Linear spaces; normed spaces; first examples

1.1 Linear spaces

IN THIS course we study linear spaces E over the field of real or complex numbers $\mathbb{R}$ or $\mathbb{C}$. The simplest examples of linear spaces studied in a course of Linear Algebra are those of the n–dimensional vector spaces $\mathbb{R}^n$, $\mathbb{C}^n$ or the space of polynomials of degree, say, less than n.

An important example of linear space is the space $C[a, b]$ of continuous real (or complex) valued functions on the interval $[a, b]$.

A map $A : E_1 \mapsto E_2$ between two linear spaces E_1 and E_2 is called linear if and only if for every $x, y \in E_1$ and for every scalar a, b we have that

$$A(ax + by) = aA(x) + bA(y). \tag{1.1}$$

For such maps we usually write Ax instead of $A(x)$. Moreover, we define two important sets associated with a linear map A, its kernel $\ker A$ and its image $\mathrm{Im} A$ defined by:

$$\ker A = \{x \in E_1 \ : \ Ax = 0\} \tag{1.2}$$

and

$$\mathrm{Im} A = \{Ax \ : \ x \in E_1\}. \tag{1.3}$$

A linear map $A : E_1 \mapsto E_2$ between two linear spaces E_1 and E_2 is called *isomorphism* if $\ker A = 0$ and $\mathrm{Im} A = E_2$, that is A is an one to one and onto linear map and consequently it is invertible. We write A^{-1} for its inverse.

Examples of linear spaces.

1. s^* is the set of finite support sequences; that is, the sequences with all but finite zero elements. It is a linear space with respect to addition of sequences and obviously isomorphic to the space of all polynomials.

2. The set c_0 of sequences tending to zero.

3. The set c of all convergent sequences.

4. The set ℓ_∞ of all bounded sequences.

5. The set s of all sequences.

All of these form linear spaces and they relate in the following way:

$$s^* \subseteq c_0 \subseteq c \subseteq \ell_\infty \subseteq s. \tag{1.4}$$

Definition 1.1.1 *A linear space E_1 is called a subspace of the linear space E if and only if $E_1 \subseteq E$ and the linear structure of E restricted on E_1 gives the linear structure of E_1. We will write $E_1 \hookrightarrow E$.*

A set of vectors $x_1, x_2, \ldots, x_n$ is called linearly dependent set and the vectors linearly dependent vectors, if there exist numbers $(a_i)_{i=1}^n$ *not all of them zero*, so that

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0. \tag{1.5}$$

On the other hand $\{x_i\}_1^n$ are called linearly independent if they are *not* linearly dependent.

We define the linear span of a subset M of a linear space E to be the intersection of all subspaces of E containing M. That is,

$$\mathrm{span} M = \bigcap_\alpha \{E_\alpha \;:\; E_\alpha \hookrightarrow E \text{ and } M \subseteq E_\alpha\}. \tag{1.6}$$

An important theorem of linear algebra states:

Theorem 1.1.2 *Let $(x_i)_1^n$ be a maximal set of linearly independent vectors in E (meaning that there is no linear independent extension of this set). Then the number n is invariant and it is called the dimension of the space E. We write* $\dim E = n$ *and we say that the vectors $(x_i)_1^n$ form a basis of E.*

Next we introduce the notion of quotient spaces. For a subspace E_1 of E we define a new linear space called the *quotient space* of E with respect to E_1 in the following way. First we consider the collection of subsets

$$\{[x] = x + E_1 \ : \ x \in E\}. \tag{1.7}$$

The sets $[x]$ are called *cosets* of E. Note that two cosets $[x]$ and $[y]$ are either identical or they are disjoint sets. Indeed, if $z \in [x] \cap [y]$ then $z - x,\ z - y$ are both elements of E_1. Since E_1 is a linear space it follows that $y - x = (z - x) - (z - y) \in E_1$. Thus, if $a \in [x]$ we have $a - x \in E_1$ and again by the linear structure of E_1, $a - y = (a - x) - (y - x) \in E_1$ that is, $a \in [y]$. So, we showed that $[x] \subseteq [y]$. Similarly one shows that $[y] \subseteq [x]$.

Now we introduce a linear structure on E/E_1 by

$$\begin{aligned} [x] + [y] &= [x + y] \\ [ax] &= a[x] \end{aligned}$$

Note that $[0]$ is a zero of the new space E/E_1. The dimension dimE/E_1 is called the codimension of E_1 and we write $\mathrm{codim}_E E_1 = \mathrm{dim} E/E_1$ or simply $\mathrm{codim} E_1$ if it is obvious to which space E it refers to.

Example: The codimension of c_0 inside the space c of convergent sequences is equal to 1. Indeed, for every $x = (x_n)_1^\infty \in c$, $x + c_0 = a(\mathbf{1} + c_0)$ where $a = \lim x_n$.

Lemma 1.1.3 *If* $\mathrm{dim} E/E_1 = n$ *then there exist* $x_1, x_2, \ldots, x_n$ *such that for every* $x \in E$ *there are numbers* $a_1, a_2, \ldots, a_n$ *and* $y \in E_1$ *such that*

$$x = \sum a_i x_i + y. \tag{1.8}$$

Proof: Let $[x_1], [x_2], \ldots, [x_n]$ be a basis of E/E_1. Then the vectors $x_1, x_2, \ldots, x_n$ are linearly independent and moreover if $\sum_1^n a_i x_i \in E_1$ then $a_i = 0$ for all i [Indeed, $\sum_1^n a_i x_i \in E_1$ hence $\sum_1^n a_i([x_i]) = 0$ and now it follows from the linear independence of the $[x_i]$'s]. Now $\forall x \in E$, consider $x + E_1 = \sum a_i([x_i])$ that is $x \in \sum a_i x_i + E_1$. □

The reader should now try the exercises 1,2

1.2 Normed spaces; first examples

We now proceed to define the notion of "distance" in a linear space. This is necessary if one wants to do analysis and study convergence.

Definition 1.2.1 *A norm $p(x) = \|x\|$ for $x \in E$ is a function from E to $\mathbb{R}$ satisfying the following properties:*

1. $p(x) \geq 0$ and $p(x) = 0$ if and only if $x = 0$.

2. $p(\lambda x) = |\lambda| p(x)$.

3. $p(x + y) \leq p(x) + p(y)$ (triangle inequality)

for all $x, y \in E$ and for all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$ if the space is over the field $\mathbb{C}$)

With this definition the distance between two points x and y in E is the norm of the difference: $\|x - y\|$.

Examples. On the spaces c_0, c, ℓ_∞ we define the norm to be the supremum of the absolute value of the sequences: for $x = (a_i)_1^\infty$ we set $\|x\| = \sup |a_i|$ (exercise: check that this supremum defines a norm).

For the space $C[0, 1]$ we define the norm of a function f to be the $\max\{|f(x)| : x \in [0, 1]\}$. An other example is the space $\ell_1 = (\mathbb{R}^\infty, \|\cdot\|_1)$ which consists of all the sequences $x = (x_i)_1^\infty$ satisfying

$$\|x\|_1 = \sum_{i=1}^{\infty} |x_i| < \infty. \tag{1.9}$$

Similarly we define the spaces $\ell_p = (\mathbb{R}^\infty, \|\cdot\|_p)$ for $1 \leq p < \infty$ by requiring that

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} < \infty. \tag{1.10}$$

It is already not trivial that these sets (for $p > 1$) form linear spaces. In fact, we first prove that the function $\|\cdot\|_p$ is indeed a norm and the triangle inequality implies that if x and y are in ℓ_p then $x + y$ is also in ℓ_p. This follows from the following inequality of Hölder.

1.2.1 Hölder inequality.

Theorem 1.2.2 *For every sequence of scalars (a_i) and (b_i) we have:*

$$\left|\sum a_k b_k\right| \leq \left(\sum |a_k|^p\right)^{1/p} \left(\sum |b_k|^q\right)^{1/q}, \tag{1.11}$$

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

Let us first observe a few connections between the numbers p and q:

(1.12)
$$\frac{1}{p-1} = q-1 \text{ and } (p-1)q = p.$$

In order to prove the above inequality we set $c_i = \frac{|a_i|}{(\sum |a_j|^p)^{1/p}}$ and $d_i = \frac{|b_i|}{(\sum |b_j|^q)^{1/q}}$. Then $\sum c_i^p = 1$ and $\sum d_i^q = 1$. Now check that $c_i d_i \leq \frac{1}{p} c_i^p + \frac{1}{q} d_i^q$. Indeed, this is true because one considers the function $y = x^{p-1}$ and integrate this with respect to x from zero to c_i; and integrate with respect to y its inverse $x = \frac{1}{y^{p-1}} = y^{q-1}$ from zero to d_i. It is easy to see geometrically, that the sum of these two integrals always exceeds the product $c_i d_i$ and it equals $\frac{1}{p} c_i^p + \frac{1}{q} d_i^q$.

Adding up we get

(1.13)
$$\sum c_i d_i \leq \frac{1}{p} + \frac{1}{q} = 1.$$

□

The above inequality is called the *Cauchy inequality* if $p = q = 2$. From the inequality of Hölder follows the Minkowski inequality which is the triangle inequality for the spaces $\ell_p^n = (\mathbb{R}^n, \|\cdot\|_p)$.

1.2.2 Minkowski inequality

Theorem 1.2.3 *For every sequence of scalars $a = (a_i)$ and $b = (b_i)$ and for $1 \leq p \leq \infty$ we have:*

(1.14)
$$\|a+b\|_p \leq \|a\|_p + \|b\|_p.$$

Proof:

$$\begin{aligned}
\|a+b\|^p &= \sum |a_k + b_k|^p \\
&\leq \sum (|a_k| + |b_k|)^p \\
&= \sum (|a_k| + |b_k|)^{p-1} |a_k| + \sum (|a_k| + |b_k|)^{p-1} |b_k| \\
&\leq \left(\sum (|a_k| + |b_k|)^p\right)^{1/q} \left(\left(\sum |a_k|^p\right)^{1/p} + \left(\sum |b_k|^p\right)^{1/p} \right) \\
&= \left(\sum (|a_k| + |b_k|)^p\right)^{1/q} (\|a\|_p + \|b\|_p)
\end{aligned}$$

for q such that $\frac{1}{p} + \frac{1}{q} = 1$. □

A few topological remarks are due. We say that a sequence (x_n) converges to a point x in the space E if and only if $\|x_n - x\| \to 0$. An open ball of radius $r > 0$ centered at x_0 is defined to be the set

$$D_r(x_0) = \{x| \ \|x - x_0\| < r\} \tag{1.15}$$

and a set $\mathcal{O}$ is said to be open if and only if for every $x \in \mathcal{O}$ there exists $r > 0$ such that $D_r(x) \subseteq \mathcal{O}$. A set F is said to be closed if for every sequence $x_n \in F$ that converges to some $x \in E$ it follows that $x \in F$.

Lemma 1.2.4 *If $\mathcal{O}$ is an open set then the set $F = \mathcal{O}^c$ is closed. Conversely, if F is a closed set then the set $\mathcal{O} = F^c$ is open.*

Proof: Let $x_n \in F$ and $x_n \to x \in E$. If $x \in \mathcal{O}$ then for any $r > 0$ and n large enough, $\text{dist}(x_n, x) < r$ which implies that for n large enough, $x_n \in \mathcal{O}$ and not in F. For the converse now, for every $x \in F^c$ there exists $\varepsilon > 0$ such that $D_\varepsilon(x) \subseteq \mathcal{O}$. If not for every decreasing to zero sequence of $\varepsilon_n > 0$ there exist $x_n \in F$ and $\text{dist}(x_n, x) < \varepsilon_n$ which implies that $x_n \to x$ that is $x \in F$. □

We also note here that the union of open sets is open and the intersection of closed sets is closed.

We start now discussing some geometric ideas. If two points x and y are given then the set $\{\lambda x + (1 - \lambda)y\}$ for $0 \le \lambda \le 1$ is the line segment joining these two points. We also call this set an interval and we write $I[x, y]$.

Exercise. Check that if $z \in I[x, y]$ then $\|x - y\| = \|x - z\| + \|z - y\|$ that is, the triangle inequality becomes equality.

Definition 1.2.5 *A subset M of a linear space E is called convex if and only if for every two points $x, y \in M$ the interval $I[x, y]$ is contained in M.*

It is easy to see (an exercise) that if $(M_a)_a$ is a family of convex sets then the intersection $\cap_a M_a$ is also a convex set. We observe here that for E being a linear normed space the set

$$D(E) = \{x| \ \|x\| \le 1\}, \tag{1.16}$$

called the unit ball of the space E, is a convex (check!) and symmetric with respect to the origin (centrally symmetric) set.

Lemma 1.2.6 *If $E_0 \hookrightarrow E$, and E_0 is closed subspace then E/E_0 is a normed space and for $[x] \in E/E_0$ its norm is given by*

$$\|[x]\| = \inf_{y \in E_0} \|x - y\|. \tag{1.17}$$

Proof: If $\|[x]\| = 0$ then there exists a sequence $x_n \in E$ such that $x_n \to x$ for some $x \in E_0$ (E_0 is closed) hence $[x] = 0$. Homogeneity is easy since E_0 is a linear space and finally we prove the triangle inequality. Since $\|[x] + [y]\| = \inf_{z \in E_0} \|x + y + z\|$, for every $\varepsilon > 0$ we take $z_1, z_2 \in E_0$ so that

$$\|x + z_1\| \leq \|[x]\| + \varepsilon \tag{1.18}$$

and

$$\|y + z_2\| \leq \|[y]\| + \varepsilon. \tag{1.19}$$

So, for every $\varepsilon > 0$ we have that

$$\begin{aligned} \|[x] + [y]\| &\leq \|x + y + z_1 + z_2\| \\ &\leq \|x + z_1\| + \|y + z_2\| \\ &\leq \|[x]\| + \|[y]\| + 2\varepsilon, \end{aligned}$$

finishing the proof. □

A weaker notion than that of the norm is the notion of the seminorm.

$p(x)$ is a seminorm on a linear space E if it satisfies the properties of the norm except that it may be zero for non-zero vectors. So, a seminorm p satisfies

1. $p(0) = 0$
2. $p(\lambda x) = |\lambda| p(x)$
3. $p(x + y) \leq p(x) + p(y)$

for all $x, y \in E$ and all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$).

It is useful to note here that if p is a seminorm and we set E_0 to be its kernel, that is, $E_0 = \{x \in E \ : \ p(x) = 0\}$, then

1. E is a subspace

and

2. p can define a norm on the quotient E/E_0

Indeed the first is true from the triangle inequality and the second is true since $p(x + y)$ is independent of $y \in E_0$: $p(x + y_1) \leq p(x + y_2) + p(y_2 - y_1) = p(x + y_2)$ and similarly $p(x + y_2) \leq p(x + y_1) + p(y_1 - y_2) = p(x + y_1) \quad (y_1, y_2 \in E_0)$.

An analogue of the ℓ_p spaces is provided by the spaces of functions with finite p–norm: We define the space of continuous functions $C_{(p)}[a,b]$ so that if $f \in C_{(p)}[a,b]$ then

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{1/p} < \infty.$$

We note here that the quantity $\|f\|_p$ is a seminorm and hence we have to pass to a quotient space if we want to get a norm. Thus we pass to quotient as described above (quotient with respect to the set of zeroes of $\|\cdot\|_p$). In this space now we see the following "problem". It is easily seen that there exist sequences of continuous functions f_n and non continuous function f so that the quantity $\|f_n - f\|_p$ converges to zero. So f_n is inside the space but converges to a function "outside" the space of continuous functions. These spaces are called "incomplete". In the next section we look into the complete spaces.

The reader should now try the exercises 3 to 19

1.3 Completeness; completion

To approach the general picture we need the following definition.

Definition 1.3.1 *A normed space X is called complete if and only if every Cauchy sequence (x_n) in X converges to an element x of the space X.*

Examples.

1. It is well known from the standard calculus courses that if x_n is a Cauchy sequence in the space $C[a,b]$ equipped with the supremum norm (i.e., for every $\varepsilon > 0$ there is $N(\varepsilon) \in \mathbb{N}$ such that $\sup_{t\in[a,b]} |x_n(t) - x_m(t)| < \varepsilon$ for all n, m bigger than $N(\varepsilon)$ and for all $t \in [a,b]$ then there exists a *continuous* function $x(t) \in C[a,b]$ such that

$$\sup_{t\in[a,b]} |x_n(t) - x(t)| \to 0$$

as n goes to infinity. Thus the space $C[a,b]$ equipped with the norm $\|x\|_\infty = \sup_{t\in[a,b]} |x(t)|$ is a complete normed space.

2. The space ℓ_2 is a complete normed space, since if $x^n = (x^n_m)_m$ is a Cauchy sequence with the norm $\|\cdot\|_2$ then each sequence $(x^n_m)_n$ is a Cauchy sequence and by the completeness of $\mathbb{R}$ or $\mathbb{C}$ there exists x_m the limit of (x^n_m) as n tends to infinity. Let $x = (x_m)_m$ Then

$$\sum_1^k |x_m|^2 = \lim_{n\to\infty} \sum_1^k |x^n_m|^2 \le \sup \|x^n\|_2^2 < M < \infty$$

   for some $M > 0$. Thus $x \in \ell_2$ Finally $\|x^n - x^k\|_2 \le \varepsilon$ for big n, k so passing to the limit as k goes to infinity we get that $\|x^n - x\|_2 \le \varepsilon$, consequently $x^n \to x$.

   Note, that the same proof (with the obvious modifications) works for the ℓ_p spaces as well.

3. The space c_0 of sequences converging to zero, equipped with the supremum norm is complete (left as an exercise).

For non complete normed spaces there exists a procedure to “fill in the gaps” and make them complete.

Theorem 1.3.2 *Let E be a normed linear space. There exists a complete normed space $\hat{E}$ and a linear operator $T : E \to \hat{E}$ such that (i) $\|Tx\| = \|x\|$ (isometry into); (ii) $\operatorname{Im} T$ (= TE) is a dense set in $\hat{E}$ (i.e. $\overline{TE} = \hat{E}$).*

The reader should now try the exercise 20

[Also, in the sense which we don't explain now, such a $\hat{E}$ is unique. Explain this]

1.3.1 Construction of completion

Let $\mathcal{E}$ be the (linear) space of all Cauchy sequences

$$X = (x_i \in E)_{i=1}^{\infty} \tag{1.20}$$

in E. Introduce a seminorm in the space $\mathcal{E}$: $p(X) = \lim_{i\to\infty} \|x_i\|$, where $X = (x_i)$ is a Cauchy sequence. Note that the limit always exists. [Indeed, $|\|x_n\| - \|x_m\|| \le \|x_n - x_m\| \longrightarrow 0$ as $n > m \to \infty$, by the definition of Cauchy sequences; so $\{\|x_n\|\}$ is a Cauchy sequence of numbers and therefore converges.]

Define $N = \{X : p(X) = 0\}$, so that N is the subspace of all sequences which converge to 0. Then p defines a norm on the quotient space $\hat{E} = \mathcal{E}/N$ (as explained earlier) by the same formula $p(X) = \lim_{i\to\infty} \|x_i\|$ (for any representative X of an equivalent class $\chi = X + N \in \mathcal{E}/N$. The operator $T : E \to \hat{E}$ is defined by $Tx = \chi$ (= $X + N$) where X is the constant sequence $X = (x, x, \ldots, x, \ldots)$). (A constant sequence is, of course, a Cauchy sequence and $p(X) = \|x\|$.)

Now, to prove the theorem, we should prove (a) TE is dense in $\hat{E}$; and (b) $\hat{E}$ is a complete space.

Proof: (a) Forall $\varepsilon > 0$ and $X = (x_n)$, there exists $N \in \mathbb{N}$ such that $\|x_n - x_m\| < \varepsilon$ for $n > m \geq N$. Define $Y_n \in \mathcal{E}$, $Y_n = (x_n, x_n, \ldots, x_n, \ldots)$ a constant sequence; i.e. $Tx_n = Y_n$. Then the distance the distance from X to Y_n is $p(X - Y_n) \leq \varepsilon$. Thus, every X is approximable by elements of TE.

(b) Let $p(X^{(n)} - X^{(m)}) \longrightarrow 0$ as $n \geq m \to \infty$ (i.e. $X^{(n)}$ is a Cauchy sequence in $\mathcal{E}$ and represents a Cauchy sequence in $\hat{E} = \mathcal{E}/N$). Take $\varepsilon_n \searrow 0$ and $x_n \in E$ such that $p(X^{(n)} - Tx_n) < \varepsilon_n$. Then $\{x_n\}$ is a Cauchy sequence in E. Indeed, $\|x_n - x_m\| = p(Tx_n - Tx_m) \leq p(Tx_n - X^{(n)}) + p(X^{(n)} - X^{(m)}) + p(X^{(m)} - Tx_m) \longrightarrow 0$ as $n \geq m \to \infty$. Then $X^0 = (x_n)$ is a Cauchy sequence (so it belongs to $\mathcal{E}$) and $X^{(0)} = \lim X^{(n)}$. Indeed, $p(X^{(n)} - X^0) \leq p(X^{(n)} - Tx_n) + p(Tx_n - X^0) \to 0$. □

(Compare this with the construction of irrational numbers from rational ones.)

The completion of $C_{(p)}[a, b]$ is called $L_p[a, b]$. Hence, an element in $L_p[a, b]$ is a class of functions, but we will always choose a representative of this class and will treat it as an element of the space $L_p[a, b]$. The most important space for us is $L_2[a, b]$.

1.4 Exercises

1. Consider the linear space $\hat{c}$ of double sequences $x = (x_n)_{n=-\infty}^{\infty}$ such that the limits $b_1 = \lim_{n\to\infty} x_n$ and $\lim_{n\to-\infty} x_n$ exist. Consider moreover the subspace $\hat{c}_0$ of the sequences $y = (y_n)_{n=-\infty}^{\infty}$ such that $\lim_{n\to\pm\infty} y_n = 0$. Find the dimension and a basis of the space $\hat{c}/\hat{c}_0$.

2. Consider the linear space $c^{(3)}$ of all sequences $x = (x_n)_{n=1}^{\infty}$ such that $\{x_{3k+q}\}_{k=0}^{\infty}$ converges for $q = 0, 1, 2$. Find the dimension

and a basis for the space $c^{(3)}/c_0$.

3. Let V be an open set. Prove that the set $F = V^c$ is closed.

4. Let F be a closed set. Prove that the set $V = F^c$ is open.

5. Let E be a normed linear space and $E_0 \hookrightarrow E$ be a closed subspace. Then E/E_0 is a normed linear space with the norm $p([x]) = \inf_{y\in E_0} \| x + y \|$.

6. Prove that

   (a) the union of any family of open sets is open set.
   (b) the intersection of any family of closed sets is a closed set.

7. Prove that a ball is a convex set.

8. Show that there exists two vectors x and y in the space ℓ_∞ such that they are linearly independent, $\|x\| = \|y\| = 1$ and $\|x + y\| = 2$.

9. Prove that if the unit sphere of a normed linear space contains a line segment, then there exist vectors x and y such that $\|x + y\| = \|x\| + \|y\|$ and x, y are linearly independent (a line segment is a set of the form $\{\lambda u + (1 - \lambda)v \ : \ 0 \le \lambda \le 1\}$).

10. Prove that if a normed linear space X contains linearly independent vectors x and y such that $\|x\| = \|y\| = 1$, $\|x + y\| = \|x\| + \|y\|$, then there exists a line segment contained in the unit sphere of X.

11. Given the two spheres $\{x \ : \ \|x - y_0\| = \|y_0\|\}$ and $\{x \ : \ \|x + y_0\| = \|y_0\|\}$ in a normed linear space, how many points can these spheres have in common?

12. Find the intersection of the unit ball in $C[0, 1]$ with the following subspaces:

    (a) span$\{t\}$
    (b) span$\{1, t\}$
    (c) span$\{1 - t, t\}$.

13. Compute the norm of a vector in the factor space $\hat{c}/\hat{c}_0$ (see exercise 1).

14. Prove that if $p \geq q \geq 1$ then $l_q \subset l_p$.

15. Prove that if $p \geq q \geq 1$ then $L_p([a,\beta]) \subset L_q([a,\beta])$ for every finite segment $[a,\beta]$.

16. Prove that for $p \neq q$ no space $L_p(0,\infty)$ is a subspace of $L_q(0,\infty)$.

17. Let p,q,r be numbers such that $p,q,r > 1$ and $\frac{1}{p}+\frac{1}{q}+\frac{1}{r} = 1$. Let $f \in L_p(a,b),\ g \in L_q(a,b),\ h \in L_r(a,b)$. Prove that $fgh \in L_1(a,b)$ and $\| fgh \|_1 \leq \| f \|_p \| g \|_q \| h \|_r$.

18. Let $0 < \alpha \leq \beta < \infty$. For which p the function $f(x) = \frac{1}{x^\alpha + x^\beta}$ belongs to $L_p(0,\infty)$?

19. (a) Prove that the intersection of the two balls $\mathcal{D}_1 = \{z : \| \xi_1 - z \| \leq R_1\}$ and $\mathcal{D}_2 = \{z : \| \xi_2 - z \| \leq R_2\}$ in a normed lenear space is empty iff $\| \xi_1 - \xi_2 \| \leq R_1 + R_2$.

    (b) Is the intersection of the two balls $\mathcal{D}_1$ and $\mathcal{D}_2$ in the space l_2 empty if $R_1 = \frac{1}{2},\ \xi_1 = (1,0,0,\frac{1}{27},\frac{1}{81},\ldots),\ R_2 = \frac{3-\sqrt{2}}{2\sqrt{2}}$ and $\xi_2 = (0,\frac{1}{3},\frac{1}{9},0,0,\ldots)$?

20. Let $x^{(n)} \in c_0,\ x^{(n)} \xrightarrow{n} x = (x_k)_{k=1}^\infty$ in the supremum norm. Prove that $x \in c_0$.

Chapter 2

Hilbert spaces

2.1 Basic notions; first examples

Let H be a linear space over $\mathbb{C}$ with a given complex value function of two variables $\langle x, y\rangle : H \times H \to \mathbb{C}$, which has the following properties:

1. linearity with respect to the first argument:

   (2.1) $$\langle ax_1 + bx_2, y\rangle = a\,\langle x_1, y\rangle + b\,\langle x_2, y\rangle \;\; ;$$

2. complex conjugation: $\overline{\langle x, y\rangle} = \langle y, x\rangle$; this implies "semi-linearity" with respect the second argument: $\langle x, ay_1 + by_2\rangle = \overline{a}\,\langle x, y_1\rangle + \overline{b}\,\langle x, y_2\rangle$.

3. non-negativeness: $\langle x, x\rangle \geq 0$, and $\langle x, x\rangle = 0$ ifand only if $x = 0$.

Such a function is called "inner product". Consider also the function $p(x) = \langle x, x\rangle^{1/2}$. (We will see later that $p(x)$ is a norm and will write $p(x) = \|x\|$.)
Examples:

1. In $\mathbb{C}^n$ let $\langle x, y\rangle = \sum_1^n a_i\overline{b}_i$ where $x = (a_i)_1^n$, $y = (b_i)_1^n$.

2. In ℓ_2 let $\langle x, y\rangle = \sum_1^\infty a_i\overline{b}_i$ (by Hölder inequality $|\sum a_i\overline{b}_i| \leq \sqrt{\sum |a_i|^2} \cdot \sqrt{\sum |b_i|^2} < \infty$). So, $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$.

3. In $L_2[a, b]$ (re-write this parenthesis using integrals, think about functions you know how to

integrate; say, think about the Riemannian integral) let $\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}dt$. Again, $|\langle f, g\rangle| \le \|f\|_2 \cdot \|g\|_2$ by the "Cauchy-Schwartz" inequality, that is, Hölder inequality for $p = q = 2$.

2.1.1 Cauchy-Schwartz inequality

Theorem 2.1.1 (Cauchy-Schwartz inequality) *For all vectors x, y in a linear space H with inner product $\langle \cdot, \cdot\rangle$, the following inequality is true:*

$$|\langle x, y\rangle| \le \langle x, x\rangle^{1/2} \cdot \langle y, y\rangle^{1/2} \tag{2.2}$$

Proof: Recall our notation $p(x) = \langle x, x\rangle^{1/2}$. Then $0 \le \langle x - \lambda y, x - \lambda y\rangle = p(x)^2 - 2\mathrm{Re}(\lambda \langle y, x\rangle) + |\lambda|^2 p(y)^2$.

If $\langle x, y\rangle \ne 0$ take $\lambda = \frac{p(x)^2}{\langle y,x\rangle} \Rightarrow 0 \le -p(x)^2 + \frac{p(x)^4 p(y)^2}{|\langle y,x\rangle|^2}$ which implies the Cauchy-Schwartz inequality. Moreover $|\langle x, y\rangle| = p(x) \cdot p(y)$ if and only if $x = \lambda y$. $\square$

Exercise: Let $\langle x, y\rangle$ satisfy all three conditions of the inner product *except* that $\langle x, x\rangle$ may be zero for non-zero elements. Prove that the Cauchy-Schwartz inequality is still true.

Now we will prove that $p(x) = \|x\|$ is a norm. Indeed: $p(\lambda x) = |\lambda| p(x)$ and the triangle inequality holds:

$$\begin{aligned} p(x+y)^2 &= \langle x+y, x+y\rangle \\ &= p(x)^2 + 2\mathrm{Re}\,\langle x, y\rangle + p(y)^2 \le [p(x) + p(y)]^2, \end{aligned}$$

and because of the Cauchy-Schwartz inequality: $|\mathrm{Re}\,\langle x, y\rangle| \le |\langle x, y\rangle| \le p(x)p(y)$. We see that $p(x+y) \le p(x) + p(y)$ and we will use $\|x\|$ instead of $p(x)$.

So, H is a normed space with a norm $\|x\|$ defined by the inner product in H. We call H a Hilbert space if H is a complete normed space with this norm.

Moreover a general complete normed space X is called a **Banach space**.

Exercises:

1. Check that the inner product $\langle x, y\rangle$ is a continuous function with respect to both variables: $\langle x_n, y_n\rangle \to \langle x, y\rangle$ when $x_n \to x$

and $y_n \to y$. [Consider the expression

$$\begin{aligned}\langle x, y\rangle - \langle x_n, y_n\rangle &= \langle x, y\rangle - \langle x, y_n\rangle + \langle x, y_n\rangle - \langle x_n, y_n\rangle \\ &= \langle x, y - y_n\rangle + \langle x - x_n, y_n\rangle\end{aligned}$$

and use the Cauchy-Schwartz inequality.]

2. *Parallelogram Law:* $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$
3. Define the notion of *orthogonality:* $x \perp y$ if and only if $\langle x, y\rangle = 0$.
4. *Pythagorean Theorem:* if $x \perp y$ then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ (Proof: $(x+y, x+y) = (x,x) + (x,y) + (y,x) + (y,y) = \|x\|^2 + \|y\|^2$.)

Corollary 2.1.2 *If $\{e_i\}_1^n$ are pairwise orthogonal and normalized in H (i.e. $\|e_i\| = 1$) then $\|\sum_1^n \alpha_i e_i\| = (\sum_1^n |\alpha_i|^2)^{1/2}$; moreover,*

$$\lim_{n\to\infty} \|\sum_1^n \alpha_i e_i\| = \sqrt{\sum |\alpha_i|^2}$$

(under the condition $e_i, e_j = \delta_{ij}$). The completeness of the Hilbert space gives that if $\sum_1^\infty \alpha_i^2 < \infty$ then the series $\sum_i^\infty \alpha_i e_i$ converges.

The reader should now try the exercises from 1 to 8

Indeed $\|\sum_n^m \alpha_i e_i\| = \sqrt{\sum_n^m |\alpha_i|^2} \longrightarrow 0$ as $m > n \to \infty$.

2.1.2 Bessel's inequality

Theorem 2.1.3 (Bessel's inequality) *For any orthonormal system $\{e_i\}_{i\geq 1} \subset H$, and for every $x \in H$ we have: $\sum_{i\geq 1} |\langle x, e_i\rangle|^2 \leq \|x\|^2$.*

Proof: Consider $y_n = \sum_1^n \langle x, e_i\rangle e_i$. Then $|\langle y_n, x\rangle| \leq \|x\| \cdot \|y_n\|$ and $\|y_n\| = \sqrt{\sum_1^n |\langle x, e_i\rangle|^2}$. Hence $\sum_1^n |\langle x, e_i\rangle|^2 \leq \|x\| \cdot \sqrt{\sum_1^n |\langle x, e_i\rangle|^2}$ and the series converges. □

Corollary 2.1.4 *For any $x \in H$ and any orthonormal system $\{e_i\}_1^\infty$, there exists $y = \sum_{i=1}^\infty \langle x, e_i\rangle e_i$. [Indeed, the y_n in the previous proof converge to y as $n \to \infty$].*

Examples of orthonormal systems:

1. In ℓ_2 consider the vectors $\{e_n = (0, \ldots, 0, 1, 0, \ldots)\}_{n=1}^\infty$ where the 1 appears in the n-th position.

2. In $L_2[-\pi,\pi]$ consider the vectors $\{\frac{1}{\sqrt{2\pi}}e^{int}\}_{n=-\infty}^{\infty}$;

3. In $L_2[-\pi,\pi]$ consider the vectors $\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}$, for $n = 1, 2, \ldots$

We call a system $\{x_i\}_{i\geq 1}$ *a complete system* in H (or any other normed space X) if the linear span $\{ \sum_{i=1}^{n} \alpha_i x_i \mid \forall\, n \in \mathbb{N}, \text{forallscalars } \alpha_i\}$ is a dense set in H (or, correspondingly, in X).

Remarks. A few known theorems of Calculus state that some systems of functions are complete in some spaces. The Weierstrass approximation theorem, for example, states that the system $\{t^n\}_{n\geq 0}$ is complete in $C[0,1]$ (meaning that polynomials are dense in $C[0,1]$). Since $C[0,1]$ is dense in $L_2[0,1]$ (by the definition of $L_2[0,1]$) and the convergence $f_n \to f$ in $C[0,1]$ implies $f_n \to f$ in $L_2[0,1]$ (check it!), it follows that $\{t^n\}_{n\geq 0}$ is a complete system in $L_2[0,1]$ too.

Another version of the Weierstrass theorem states that the trigonometric polynomials are dense in the space of the continuous 2π-periodic functions on $[-\pi,\pi]$ (in the $C[-\pi,\pi]$-norm). As a consequence the system $\{1, \cos nx, \sin nx\}_{n=1}^{\infty}$ is complete in $L_2[-\pi,\pi]$. The same is true for $\{e^{int}\}_{n=-\infty}^{\infty}$.

Lemma 2.1.5 *If $\{f_i\}$ is a complete system and $x \perp f_i$, then $x = 0$.*

Proof: Indeed, $x \perp f_i$ implies $x \perp \mathrm{span}(f_i)$, implies that x is orthogonal to a dense set in H and finally implies that there exists $x_n \to x$ and $x_n \perp x$; this means $0 = \langle x_n, x\rangle \to \langle x, x\rangle = \|x\|$. So $x = 0$.

2.1.3 Gram-Schmidt orthogonalization procedure

Algorithm 2.1.6 (Gram-Schmidt) *Let $\{x_i\}_1^{\infty}$ be a linearly independent system. Consider $e_1 = x_1/\|x_1\|$, and inductively, $e_n = \frac{x_n - y_n}{\|x_n - y_n\|}$ for $y_n = \sum_1^{n-1} \langle x_n, e_i\rangle\, e_i$.*
Then,

1. *$\{e_i\}_1^{\infty}$ is an orthonormal system (just check that for $m < n$, $(e_m, e_n) = 0$).*

2. *$\mathrm{span}\{x_i\}_1^n = \mathrm{span}\{e_i\}_1^n$ for every $n = 1, 2, \ldots$ The proof here is by induction: if it is true for $n-1$ then $x_n - y_n \neq 0$ by the linear independence of the $\{x_k\}$. Also obviously, $e_k \in \mathrm{span}\{x_i\}_1^n$ for $k \leq n$; and $x_n \in \mathrm{span}\{e_i\}_1^n$.*

Definition 2.1.7 *A normed space X is called a separable space if there exists a dense countable set in X.*

Corollary 2.1.8 *The Hilbert space H is separable if and only if there exists a complete orthonormal system $\{e_i\}_{i\geq 1}$.*

Proof: if H separable then there exists a countable dense subset $\{y_i\}_{i\geq 1}^{\infty}$. Choose inductively a subset $\{x_i\}_{i\geq 1}$ such that the set $\mathrm{span}\{x_i\}_{i\geq 1}$ is still dense, and it is linearly independent; now apply Gram-Schmidt to the system $\{x_i\}_{i\geq 1}$completing the proof of the one direction. If on the other hand $\{e_i\}$ is complete, then consider all finite sums $\sum \alpha_i e_i$ with rational coefficients (α_i). This is a dense countable set in H. □

Definition 2.1.9 *A sequence $\{x_i\}_{i=1}^{\infty}$ is called a basis of a normed space X if for every $x \in X$ there exists a unique series $\sum_{i\geq 1} a_i x_i$ that converges to x.*

Theorem 2.1.10 *A complete orthonormal system $\{e_i\}_{i=1}^{\infty}$ in H is a basis in H.*

Proof: For every $x \in H$, by the Bessel inequality $\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 < \infty$. By the corollary 2.1.2 the element $y = \sum_{i=1}^{\infty} \langle x, e_i\rangle e_i \in H$ exists. This implies $(y - x)\perp e_i$ for forall i. By the lemma 2.1.5 we get $y = x$. So $x = \sum_{i=1} \langle x, e_i\rangle e_i$. (The uniqueness is obvious: if $x = \sum_1^{\infty} a_i e_i$, then $\langle x, e_i\rangle = a_i$.)

Corollary 2.1.11 *Every separable Hilbert space has an orthonormal basis.* □

2.1.4 Parseval's equality

Corollary 2.1.12 (Parseval) *Let $\{e_i\}_{i\geq 1}$ be an orthonormal system. Then $\{e_i\}_{i\geq 1}$ is a basis in H if and only if for all $x \in H$,*

$$\|x\|^2 = \sum_{i\geq 1} |\langle x, e_i\rangle|^2. \tag{2.3}$$

Proof:

$$\text{“}\Rightarrow\text{”} \quad : \text{ if } x = \sum_{i\geq 1} \langle x, e_i\rangle e_i \Rightarrow \|x\|^2 = \sum_{i\geq 1} |\langle x, e_i\rangle|^2 \ , \tag{2.4}$$

$$\text{“}\Leftarrow\text{”} \quad : \|x - \sum_1^n \langle x, e_i\rangle e_i\|^2 = \|x\|^2 - \sum_1^n |\langle x, e_i\rangle|^2 \underset{n\to\infty}{\longrightarrow} 0 \ . \tag{2.5}$$

since we assume that $\|x\|^2 = \sum |\langle x, e_i\rangle|^2$.

Remark 2.1.13 *Note that the direction in 2.5 of the statement is true also for a single x. Precisely, if for some $x \in H$ we have that 2.3 is true, then $x = \sum_{i\geq 1} \langle x, e_i\rangle e_i$. Therefore, if 2.3 is true for a dense subset of x's it already implies that $\{e_i\}_{i\geq 1}$ is a basis.*

Theorem 2.1.14 *Any two separable infinite dimensional Hilbert spaces H_1 and H_2 are isometrically equivalent; meaning that there exists a linear isomorphism $T : H_1 \to H_2$ such that $\|Tx\| = \|x\|$ and, moreover, $(Tx, Ty)_{H_2} = (x, y)_{H_1}$ for every x and y in H_1.*

Proof: We will build such a T for a given H (in place of H_1) and ℓ_2 (in place of H_2). Take an orthonormal basis $\{f_i\}_1^\infty$ of H. For every $x \in H$, $x = \sum_1^\infty (x, f_i) f_i$ and $\|x\|^2 = \sum_1^\infty |(x, f_i)|^2$. Let $\{e_i\}_1^\infty$ be the natural basis of ℓ_2. Then

$$Tx = \sum_1^\infty (x, f_i) e_i \in \ell_2 \ . \tag{2.6}$$

Check the isometry!

Examples

1. $\{\frac{1}{\sqrt{2\pi}} e^{int}\}_{-\infty}^{\infty}$ is an orthonormal basis of $L_2[-\pi, \pi]$.
2. Similarly, $\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nt}{\sqrt{\pi}}, \frac{\sin nt}{\sqrt{\pi}}\}$ is an orthonormal basis of $L_2[-\pi, \pi]$.

We will show now one example of use of Parseval's equality: For an interval $I = [a, b]$ we denote by $L_2(I^2)$ the space of the square integrable functions of two variables with norm:

$$\|f(t, \tau)\| = \sqrt{\int_I \int_I |f(t, \tau)|^2 dt\, d\tau}. \tag{2.7}$$

Let $\{\varphi_i(t)\}_1^\infty$ be an orthonormal basis of $L_2[a, b]$. Then the system

$$\{\varphi_i(t)\varphi_j(\tau) = \psi_{ij}(t, \tau)\}_{i,j=1}^\infty \tag{2.8}$$

is an orthonormal basis of $L_2([a, b]^2)$

Proof: Note that the system $\{\psi_{ij}\}$ is orthonormal and define $a_{ij} = \int\int_{I^2} f(t, \tau)\overline{\varphi_i(t)\varphi_j(\tau)} dt\, d\tau$. By "Parseval equality" theorem, it is enough to prove that

$$\int\int_{I^2} |f|^2 dt\, d\tau = \sum_{ij} |a_{ij}|^2 \qquad \forall f \in L_2(I^2) \ . \tag{2.9}$$

Let $a_j(t) = \int_I f(t,\tau)\overline{\varphi_j(\tau)}dt$. By the Parseval's equality is follows that $\sum |a_j(t)|^2 = \int_I |f(t,\tau)|^2 d\tau$. Also $a_{ij} = \int_I a_j(t)\overline{\varphi_i(t)}dt$ and again by the Parseval's equality $\sum_{i=1}^{\infty} |a_{ij}|^2 = \int_I |a_j(t)|^2 dt$. Combining these equalities we get:

$$\sum_{ij} |a_{ij}|^2 = \int_I \sum_j |a_j(t)|^2 dt = \int_I \int_I |f(t,\tau)|^2 dt\, d\tau \ . \tag{2.10}$$

□

2.2 Projections; decompositions

Let L be a closed subspace of H (we write $L \hookrightarrow H$). Define a projection of $x \in H$ onto L: consider the distance $\rho(x, L) = \inf_{y\in L} \|x - y\|$. If there is $y \in L$ such that $\rho(x, L) = \|x - y\|$ (i.e. the infimum is achieved), then we write $y = P_L x$, the projection of x onto L.

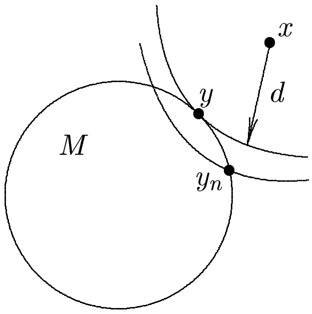
2.2.1 Separable case

Let us consider first a particular case of a separable subspace L (note that if H is separable then any of its subspaces is separable (check it!)).

Let $\{e_i\}_{i\geq 1}$ be an orthonormal basis in L. Take $y = \sum_{i\geq 1} \langle x, e_i\rangle\, e_i \in L$. Clearly $x - y \perp L$ and, because of this, for any $z \in L$: $\|x - z\|^2 = \|x - y\|^2 + \|y - z\|^2$ $(x - y \perp y - z)$. So it follows that $\inf\{\|x - z\| \mid z \in L\} = \|x - y\|$ and $y = P_L x$. Moreover such a y is unique.

2.2.2 Uniqueness of the distance from a point to a convex set: the geometric meaning

The general case is more complicated; we start with a more general problem. Let M be a convex closed set in H. Denote the distance of x to the set M with $\rho(x, M)$. *Then there exists a unique* $y \in M$ *such that* $\rho(x, M) = \|x - y\|$ (the distance is achieved at the unique element $y \in M$): indeed, let $y_n \in M$ and $\|x - y_n\| \to \rho(x, M) = d$. Such



a sequence exists since $d = \inf_{m \in M} \|x - m\|$. Now it is easy to see that $\{y_n\}$ is a *Cauchy sequence*. This follows by the parallelogram law: $2(\|x - y_n\|^2 + \|x - y_m\|^2) = \|2x - (y_n + y_m)\|^2 + \|y_n - y_m\|^2$. Since $\frac{y_n+y_m}{2} \in M$ (by the convexity of M), and $\|x - \frac{y_n+y_m}{2}\| \geq d$ we get $\|y_n - y_m\| \longrightarrow 0$ as $n, m \to \infty$. Hence there exists $y = \lim y_n \in M$ (because M-closed) and this y is *unique*. Indeed, if there exist two points y and z where the distance is achieved then we could choose $y_{2n} \to y$ and $y_{2n+1} \to z$ and then the sequence $\{y_n\}$ would not be Cauchy.

2.2.3 Orthogonal decomposition

Now consider a closed subspace L instead of M. Note that

$$\{\rho(x, L) = \rho(x, y) \text{ for some } y \in L\} \Leftrightarrow x - y \perp L. \tag{2.11}$$

Indeed, let $y \in L$ be such that $x - y \perp L$; then we proved before that it gives the distance and it is unique: $\forall\, z \in L$, $\|x - z\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2$.

In the opposite direction, if $y \in L$ is the projection $P_L x$, consider any $z \in L$: $\|x - y\|^2 \leq \|x - (y + \lambda z)\|^2 = \|x - y\|^2 - 2\mathrm{Re}\lambda(z, x - y) + |\lambda|^2\|z\|^2$. Therefore, $2\mathrm{Re}\lambda(z, x - y) \leq |\lambda|^2\|z\|^2$. Take $\lambda = t\,\overline{(z, x - y)}$, $t \in \mathbb{R}$. Then $2t|(z, x - y)|^2 \leq t^2|(z, x - y)|^2\|z\|^2$, $\forall\, t \in \mathbb{R}$; letting $t \to 0$ we see that $(z, x - y) = 0$ and hence every $z \in L$ is orthogonal to $x - y$.

We summarize what we know in the following statement:

Proposition 2.2.1 *For all $x \in H$, there exists a unique $y \in L$ such that $x - y \perp L$ and $y = P_L x$ (it gives the distance from x to L). Then (obviously) $x = x - y + y$ and $\|x\|^2 = \|x - y\|^2 + \|y\|^2$.*

Definition 2.2.2 *For $L \hookrightarrow H$ we set $L^\perp = \{x \in H \mid x \perp L\}$. This is obviously a closed subspace of H.*

Theorem 2.2.3 *For all $L \hookrightarrow H$ closed subspace of H, $L \oplus L^\perp = H$. This decomposition of H is unique.*

Proof: For all $x \in H$ and $L \hookrightarrow H$ there exists a y so that $x = x - y + y$ with $x - y \in L^\perp$ and $y \in L$ (that is, $y = P_L x$). The uniqueness of the decomposition of H is obvious.

Corollary 2.2.4 *If L is a closed subspace, then $(L^{\perp})^{\perp} = L$ (think why!)*

Exercise. Let $L_1 \hookrightarrow L_2 \hookrightarrow H$ (closed subspaces). Let $x_2 = P_{L_2}x$. Then $P_{L_1}x = x_1 = P_{L_1}x_2$ (the so-called "Theorem of the three perpendiculars"). [Indeed, define first $x_1 = P_{L_1}x_2$, then $x - x_1 = x - x_2 + x_2 - x_1$ where $x - x_2 \perp L_2$ and $x_2 - x_1 \perp L_1$. Hence $x - x_1 \perp L_1$.]

Lemma 2.2.5 *If $E \hookrightarrow H$, closed subspace and $\mathrm{codim} E = 1$, then the subspace $E^{\perp}$ is 1-dimensional.*

Proof: $E \oplus E^{\perp} = H$. If there are two vectors x_1, x_2 are linearly independent in $E^{\perp}$, then there exist two orthogonal vectors, say $y_1, y_2 \in E^{\perp}$. Now if $\alpha y_1 + \beta y_2 = z \in E \Rightarrow \alpha = \beta = 0$ (because $\langle y_i, z\rangle = 0$); thus $Y_i = y_i + E$ are linearly independent in H/E for $i = 1, 2$, a contradiction.

The reader should now try the exercises from 9 to 18

2.3 Linear functionals

2.3.1 Linear functionals in a general linear space

Definition 2.3.1 *Let E be a linear space. Linear functionals are functions*

(2.12) $\quad f : E \to \mathbb{R}$ *or* $\mathbb{C}$ *such that* $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

Note that $\ker f [\equiv H_f] = \{x \in E \mid f(x) = 0\}$ is a linear subspace.

Examples:

1. In the linear space c_0 consider the functional f defined by $f(x) = \sum_1^{\infty} a_i b_i$. where (b_i) satisfy $\sum |b_i| < \infty$, that is, $(b_i) \in \ell_1$. It is common to identify the the functional f with the element (b_i) of ℓ_1.

2. In the linera space ℓ_p consider the functional f defined by $f(x) = \sum a_i b_i$ where (b_i) is in ℓ_q. Notice that $|f(x)| \le \sum |a_i| \cdot |b_i| \le \|x\|_{\ell_p} \cdot \|f\|_{\ell_q} < \infty$. The functional f is thus identified with the element (b_i) of ℓ_q.

3. In the linear space $C[0;1]$ consider the functionals

(a) $F(x) = \int_0^1 x(t)f(t)dt$ where f is integrable; note that $|f(x)| \leq \max_t |x(t)| \int_0^1 |f(x)|dt$.

(b) $\delta_\alpha(x) = x(\alpha)$; $|\delta_\alpha(x)| \leq \|x\|_C$.

4. Re-write this item$H : (x, y) = f_y(x)$. [So, in $L_2[a, b]$: $\langle x, y\rangle = \int_a^b x(t)\overline{y(t)}dt$].

For a space E denote with $E^\#$ the space of all linear functionals on E.

Theorem 2.3.2 *Consider $f \not\equiv 0$.*

1. $\operatorname{Codim} H_f = 1$*;*

2. *If* $\ker f = \ker g$ *(g is another linear functional) then there exists $\lambda \neq 0$ such that $\lambda f = g$.*

3. *Let $L \hookrightarrow E$ and* $\operatorname{codim} L = 1$*. Then there exists $f \in E^\#$ such that* $\ker f = L$*.*

Proof:

1. Take x', $f(x') \neq 0$; let $x_0 = x'/f(x')$ and note that $f(x_0) = 1$. then for every $x \in E$, $y = x - f(x)x_0 \in \ker f$, and consequently $x = f(x)x_0 + y$, ($y \in \ker f$) and this decomposition is unique. Hence $E/H_f = \{\lambda(x_0 + H_f) : \lambda \in \mathbb{C}\}$ which implies $\dim E/H_f = 1$.

2. Take $x = f(x)x_0 + y$ and apply g:

$$g(x) = f(x)g(x_0) \Rightarrow g = g(x_o) \cdot f. \tag{2.13}$$

3. $\dim E/L = 1 \Rightarrow E/L = \{\lambda\mathfrak{X}_0\}$ where $\mathfrak{X}_0 = x_0 + L$ and $\forall\, x \in E$ there is a unique representation $x = \lambda x_0 + y$, $y \in L$. Define $f(x) = \lambda$. Then $f(L) = 0$ and f is a linear functional.

□

2.3.2 Bounded linear functionals in normed spaces. The norm of a functional

Let now $X = (E, \|\cdot\|)$; we call $f \in X^{\#}$ a bounded functional if there exists C such that $|f(x)| \le C\|x\|$ (i.e. f is bounded on bounded sets). Let X^* be a set of all bounded functionals. This is a linear space [$\alpha f + g$ is a bounded linear functional if f and $g \in X^*$]. Define a norm: for $f \in X^*$, let $\|f\|^* = \sup_{x\neq 0} |f(x)|/\|x\|$ [Check that it is a norm.] So

$$|f(x)| \le \|f\|^* \cdot \|x\|. \tag{2.14}$$

(We usually write $\|f\|$ instead of $\|f\|^*$.)

Fact: f is a bounded functional if and only if f is a continuous functional. ["$\Rightarrow$"$|f(x_n) - f(x_0)| = |f(x_n - x_0)| \le \|f\| \cdot \|x_n - x_0\| \longrightarrow 0$ as $x_n \to x_0$.

"$\Leftarrow$" Let $f(x_n) \to 0$ for $x_n \to 0$. If f is not bounded, then for every $n \in \mathbb{N}$ there exists $x_n, \|x_n\| = 1$ and $|f(x_n)| > n$. But in this case $|f(\frac{x_n}{n})| \ge 1$, where $\frac{x_n}{n} \to 0$ a contradiction].

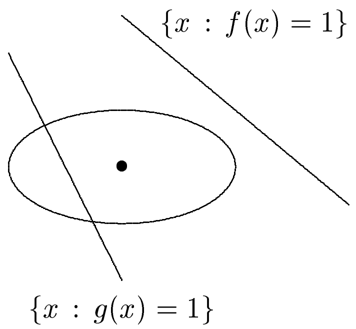
Remark: If a linear functional f is continuous at $x = 0$ then it is continuous at any x.

Note that if f is continuous then $\mathrm{ker} f$ is a closed subspace. [It is non-trivial and we don't prove that the inverse is true: if $f \in X^{\#}$ and $\mathrm{ker} f$ is closed subspace then f is continuous.]

Let us return to the definition of the norm of a linear functional. Because of the homegenuity of its definition we may use different normalizations resulting to different expressions for the norm:

$$\begin{aligned}
\|f\| &= \sup_{x\neq 0} \frac{|f(x)|}{\|x\|} \\
&= \sup\{|f(x)| \ : \ \|x\| \le 1\} \\
&= \sup_{f(x)=1} \frac{1}{\|x\|} \\
&= \frac{1}{\inf_{f(x)=1} \|x\|}.
\end{aligned}$$

Let us interpret the last expression. Note that the quantity $\inf_{f(x)=1} \|x\| =: \rho_f$ is the distance of the hyperplane $\{x \ : \ f(x) = 1\}$



to 0. So the norm $\|f\|$ is $1/\rho_f$. This means that the functional f on the picture has norm less than 1 but the functional g has norm more than 1. It also means that a functional has norm equal to 1 if and only if the hyperplane $\{x : f(x) = 1\}$ "supports" the unit ball $\mathcal{D}(X)$. Note here (an exercise) that this does not mean that the supporting hyperplane and the unit ball $\mathcal{D}(X)$ must have a common point (the picture here is misleading because it is drawn in a finite (2-)dimensional space).

2.3.3 Bounded linear functionals in a Hilbert space

We now return to study linear functionals in a Hilbert space. The following theorem describes the space of all bounded linear functionals on a Hilbert space.

Theorem 2.3.3 (Riesz Representation) *For every $\varphi \in H^*$, there exists $y \in H$ such that $\varphi(x) = (x, y)$. (Any continuous linear function on a Hilbert space is represented by some element y of the same space satisfying $\varphi(x) = (x, y)$.) Moreover $\|\varphi\|^* = \|y\|$.*

Proof: Let $\ker\varphi = L$, $\operatorname{codim} L = 1$ and since $L \oplus L^\perp = H$, $L^\perp$ is 1-dimensional. So $L^\perp = \{\lambda\hat{y}\}$ for some $\hat{y} \in H$, $\hat{y} \neq 0$; $\hat{y}$ defines a linear functional by $\hat{y}(x) = (x, \hat{y})$. Then $\ker\hat{y} = \{\hat{y}\}^\perp = (L^\perp)^\perp = L$. So $\ker\hat{y} = \ker\varphi$. By the theorem 2.3.2 part 2 we get $\varphi = \lambda\hat{y}$ (and $y = \lambda\hat{y}$ represents φ).

Now, $\|y\|^* = \sup_{x\neq 0} \frac{|(x,y)|}{\|x\|} \leq \|y\|$ and the equality holds for $x = y$. Hence $\|y\| = \|y\|^*$. □

2.3.4 An Example of a non-separable Hilbert space:

Let $H = \overline{\operatorname{span}_{\lambda\in\mathbb{R}}\{e^{i\lambda t}\}}$ with $(f, g) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x)\overline{g(x)}\, dx$. Note that if $\lambda_1 \neq \lambda_2$ then $e^{i\lambda_1 t} \perp e^{i\lambda_2 t}$. So H has uncountable set of pairwise orthonormal elements. This implies the non-separability of H.

Note that any continuous function with finite support "represents" a zero element of H (because $\langle f, f\rangle = 0$ for such a function). So, one should be carefull when describing this space as a space of functions.

2.4 Exercises

1. If L is a linear space, $\dim L = n$ $(< \infty)$ and $\{y_i\}_1^n$ is a basis of L then, for $x \in H$, there exists exactly one $y = \sum_1^n \alpha_i y_i \in L$ such that $x - y \perp L$. So the system of equations in y (for $j = 1, , \dots, n$)

$$0 = (x - y, y_j) = (x, y_j) - \sum_1^n \alpha_i (y_i, y_j) \tag{2.15}$$

has a solution, thus $g(y_1, \dots, y_n) \equiv \det|(y_i, y_j|) \neq 0$ (there is a geometric sense to this: the absolute value of the determinant of say 2 vectors is the area of the parallelogram they define) and it is called the "Gram determinant". Then we have,

$$\operatorname{dist}(x, L) = \left[\frac{g(y_1, \dots, y_{n_1} x)}{g(y_1, \dots, y_n)}\right]^{1/2} . \tag{2.16}$$

2. Let H be an inner product space.
   (a) Describe all pairs of vectors x, y for which $\| x + y \| = \| x \| + \| y \|$.
   (b) Describe all pairs of vectors x, y for which $\| x + y \|^2 = \| x \|^2 + \| y \|^2$.

3. Prove that in an inner product space holds

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

4. Let $C[0, 1]$ be the vector space of all continuous complex-valued functions on $[0, 1]$. Introduce a norm $\| \cdot \|$ on $C[0, 1]$ by $\| \xi \| = \max_{t \in [0,1]} |\xi(t)|$. Show that it is impossible to define an inner product on $C[0, 1]$ such that the norm it induces is the same as the given norm.

5. Let $w = (w_1, w_2, \cdots)$, where $w_i > 0$. Define $l_2(w)$ to be the set of all sequences $\xi = (\xi_1, \xi_2, \dots)$ of complex numbers with $\sum_{i=1}^\infty w_i |\xi_i|^2 < \infty$. Define an inner product on $l_2(w)$ by $\langle \xi, n \rangle = \sum_{i=1}^\infty w_i \xi_i \overline{n_i}$. Show that $l_2(w)$ is a Hilbert space.

6. (a) Find a vector w such that $(1, \frac{1}{2^2}, \frac{1}{3^3}, \dots) \in l_2(w)$.

(b) Find a vector w such that the set of all $\xi = (\xi_1, \xi_2, \ldots)$ with $|\xi_n| < n^n$ is in $l_2(w)$.

7. Check that the following sets are closed subspaces of ℓ_2:

   (a) $A = \{(\xi_1, 2\xi_1, \xi_3, 4\xi_3, \xi_5, \xi_6, \xi_7, \ldots) \mid \sum_{j=1}^{\infty} |\xi_j|^2 < \infty\}$

   (b) $B = \{(\xi_1, 0, \xi_3, 0, \xi_5, 0, \ldots) \mid \sum_{j=1}^{\infty} |\xi_{2j-1}|^2 < \infty\}$

   (c) $C = \{(0, \xi_2, 0, \xi_4, 0, \xi_6, \ldots) \mid \sum_{j=1}^{\infty} |\xi_{2j}|^2 < \infty\}$

8. Prove that if L is a closed subspace of a Hilbert space M, then $(L^{\perp})^{\perp} = L$.

9. Let E be a closed subspace of M and codim $E = 1$. Prove that dim $E^{\perp} = 1$.

10. (a) Prove that for any two subspaces of a Hilbert space M: $(L_1 + L_2)^{\perp} = L_1^{\perp} \cap L_2^{\perp}$.

    (b) Prove that for any two closed subspaces of a Hilbert space M: $(L_1 \cap L_2)^{\perp} = \overline{L_1^{\perp} + L_2^{\perp}}$.

11. Let $L_0 = \{\varphi \in L_2[-a, a]/\varphi(t) = -\varphi(-t) a.e.\},\ L_E = \{\varphi \in L_2[-a, a]/\varphi(t) = \varphi(-t) a.e.\}$.

    (a) Show that both sets are closed infinite demensional subspaces of $L_2[-a, a]$.

    (b) Show that L_0 and L_E are orthogonal

    (c) Show that L_E is the orthogonal complement of L_0.

    (d) For $f \in L_2[-a, a]$, find its projections into L_0 and L_E.

    (e) Find the distances from $f(t) = t^2 + t$ to L_0 and to L_E. Find the distances from any $f \in L_1[-a, a]$ to L_0 and L_E.

12. Find the Fourier coefficients of the following functions:

    (a) $f(t) = t$?

    (b) $f(t) = t^2$?

    (c) $\cos at,\ a \in \mathbb{R} \setminus \mathbb{Z}$ ($\mathbb{Z}$ is the set of intergers)?

    (d)
$$f(t) = \begin{cases} 1; & t \geq 0 \\ -1, & t < 0 \end{cases}.$$

(e) $f(t) = |t|$.

(f) Use the Parseval equality to prove that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

(g) Find $\sum_{n=1}^{\infty} \frac{1}{n^4}$.

13. Let $f(x)$ be a differentiable 2π-periodic function in $[-\pi, \pi]$ with derivative $f'(x) \in L_2[-\pi, \pi]$. Let f_n for $n \in \mathbb{Z}$ be the Fourier coefficients of $f(x)$ in the system $\{e^{inx}/\sqrt{2\pi}\}$. Prove that $\sum_{n\in\mathbb{Z}} |f_n| < \infty$.

14. Prove that the system $\sin nx$ for $n = 1, 2, \ldots$ is complete in $L_2[0, \pi]$.

15. Prove that the system $\sin\Big((2n-1)x\Big)$ for $n = 1, 2, \ldots$ is complete in $L_2[0, \pi/2]$.

16. (a) Prove that the system $\{1, t^3, t^6, \ldots\}$ is complete in the space $L_2[0, 1]$.

    (b) Prove that the system $\{1, t^2, t^4, t^6, \ldots\}$ is complete in the space $L_2[0, 1]$. Is it complete in the space $L_2[-1, 1]$?

17. Let $x_n = (0, 0, \ldots, 0, 1, 2, 0, \ldots)$ where the numbers 1 and 2 appear in the n and the $n + 1$ position, and $y_n = (1, 1, \ldots, 1, 0, 0, \ldots)$ where the first zero appears at the $n + 1$ position. Considering these vectors in ℓ_2, prove that for all $j \in \mathbb{N}$, $y_j \notin \overline{\mathrm{sp}}\{x_1, x_2, \ldots\}$.

18. Let $x_1 = (1, 0, 0, \ldots)$, $x_2 = (a, b, 0, \ldots)$, $x_3 = (0, a, b, 0, \ldots), \ldots$, where $|a/b| > 1$.

    (a) Check that $\overline{\mathrm{sp}}\{x_1, x_2, \ldots\} = \ell_2$.

    (b) Show that any finite system of these vectors is linearly independent.

    (c) Find $a_1, a_2, \ldots \in \mathbb{C}$ such that $\sum_{j=1}^{\infty} a_j x_j$ converges to zero.

19. Let $x_1 = (1, 0, 0, \ldots)$, $x_2 = (a, b, 0, \ldots)$, $x_3 = (0, a, b, 0, \ldots), \ldots$, where $|a/b| > 1$.

    (a) Check that $\overline{\mathrm{sp}}\{x_1, x_2, \ldots\} = \ell_2$.

    (b) Show that any finite system of these vectors is linearly independent.

    (c) Show that one can not find $a_1, a_2, \ldots \in \mathbb{C}$ not all zeros, such that $\sum_{j=1}^{\infty} a_j x_j$ converges to zero.

20. Determine which of the following systems are orthogonal bases in ℓ_2 and which are not:

    (a) $(1, 2, 0, 0, \ldots),\ (0, 0, 1, 2, 0, 0 \ldots),\ (0, 0, 0, 0, 1, 2, 0, \ldots), \ldots$

    (b) $(1, -1, 0, 0, \ldots),\ (1, 1, 0, 0, \ldots),\ (0, 0, 1, -1, 0, 0, \ldots),$
    $(0, 0, 1, 1, 0, 0, \ldots), \ldots$

21. Let H be the Hilbert space which is the complement of

$$M = \left\{ f \in C[0,1] \mid \exists f'(t) \ :\ \|f\|^2 = \int_0^1 |f'(t)|^2\, dt + \int_0^1 |f(t)^2\, dt \right\}$$

    and let $\phi(f) = f(0)$ for all $f \in H$.

    (a) Prove that $\phi \in H^*$.

    (b) Find $g \in M$ such that $\langle f, g\rangle = \phi(f)$ for any $f \in H$.

22. Let H be the Hilbert space which is the complement of

$$M = \{f \in C[-1,1] \mid \exists f'(t) \ :\ \|f\|^2 = \int_{-1}^1 |f'(t)|^2 dt + \int_{-1}^1 |f(t)|^2 dt < \infty\}$$

    and $\phi(f) = f(1/2)$ for all $f \in H$.

    (a) Prove that $\phi \in H^*$.

    (b) Find a $g \in M$ so that $\langle f, g\rangle = \phi(f)$ for any $f \in H$.

    Hint: Find g in the form

$$g(t) = \begin{cases} g_1(t), & -1 \le t \le 1/2 \\ g_2(t), & 1/2 \le t \le 1 \end{cases}$$

23. (a) Is the subspace

$$M = \{x = (x_1, x_2, \ldots) \in \ell_2 \ :\ \sum_{n=1}^{\infty} \tfrac{1}{n} x_n = 0\}$$

    closed in ℓ_2?

    (b) Is the subspace

$$M = \{x = (x_1, x_2, \ldots) \in \ell_2 \ :\ \sum_{n=1}^{\infty} \tfrac{1}{\sqrt{n}} x_n = 0\}$$

    closed in ℓ_2?

(c) Is the subspace $M = \{x(t) \in L_2[0,1] : \int_0^1 \frac{x(t)}{t}\,dt = 0\}$ closed in $L_2[0,1]$?

(d) Is the subspace $M = \{x(t) \in L_2[1,\infty) : \int_1^\infty \frac{x(t)}{t}\,dt = 0\}$ closed in $L_\infty[1,\infty)$?

Chapter 3

The dual space X*

3.1 Hahn-Banach theorem and its first consequences

We start with the study of the space X^* of all bounded linear functionals on a normed space X which we already introduced in section 2.3. Recall that the space X^* is equipped with the norm

$$\|f\| = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|} \tag{3.1}$$

and we call this norm the dual norm (i.e., dual to the original norm of X) and the space X^* the dual space (i.e., dual to the space X).

Statement: For any normed space X the dual space X^* is always complete, i.e. a Banach space.

Try this as an exercise now; but it will be proved later in a more general setting.

Theorem 3.1.1 (Hahn-Banach) *Let $E \hookrightarrow X$ be a subspace and $f_0 \in E^*$. Then there exists an extension $f \in X^*$ such that $f|_E = f_0$ (i.e. $f(x) = f_0(x)$ for $x \in E$) and $\|f\|_{X^*} = \|f_0\|_{E^*}$ that is,*

$$\sup_{\substack{x\neq 0\\ x\in X}} |f(x)|/\|x\| = \sup_{x\in E\setminus\{0\}} |f_0(x)|/\|x\|.$$

We will learn in this part of t he course to use this theorem without proving it. The theorem will be proved in chapter 10 in a more general setting.

Corollary 3.1.2 *1. For all $x_0 \in S(X) = \{x \in X : \|x\| = 1\}$ (the unit sphere of X), there exists $f_0 \in X^*$ such that $\|f_0\|_{X^*} = 1$, and $f_0(x_0) = 1$. [Consider the 1-dimensional subspace $E_0 = \{\lambda x_0\}$ and the functional $\varphi(\lambda x_0) = \lambda$. Then $\|\varphi\|_{E_0^*} = 1$. By the Hahn-Banach Theorem there exists an extension f_0 with the desired properties.]*

2. *For all $x_0 \in X$, there exists $f_0 \neq 0 \in X^*$ such that $f_0(x_0) = \|f_0\| \cdot \|x_0\|$. [use (1)]*

3. *For all $x_1 \in X$, for all $x_2 \in X$ such that $x_1 \neq x_2$, there exists $f \in X^*$ satisfying $f(x_1) \neq f(x_2)$ [use (2) for $x_0 = x_1 - x_2$].*

4. *X^* is a total set, meaning that if $f(x) = 0$ for every f thus $x = 0$.*

5. *Let $L \hookrightarrow X$ be a subspace of a Banach space X, and $x \in X$, $\operatorname{dist}(x, L) = d > 0$. Then ther exists $f \in X^*$ such that $\|f\| = 1$, $f(L) = 0$ and $f(x) = d$.*

Proof: First consider $L_1 = \operatorname{span}\{x, L\}$, that is

$$L_1 = \{\lambda x + y \mid \lambda \in \mathbb{R}, y \in L\}. \tag{3.2}$$

Define $f_0(z(= \lambda x + y)) = \lambda \cdot d$. Check the linearity (because $z = \lambda x + y$ can be written in a unique way); $f_0(L) = 0$ and $f_0(x) = d$. Now $\|z\| = |\lambda| \cdot \|x + \frac{y}{\lambda}\| \geq |\lambda| \cdot d = |f_0(z)| \Rightarrow \|f_0\|_{L_1^*} \leq 1$. Also there exists $y_n \in L$ so that $\|x + y_n\| \to d$. Hence $d = |f_0(x + y_n)| \leq \|f_0\| \cdot \|x + y_n\| \to d \cdot \|f_0\| \Rightarrow \|f_0\| \geq 1$.

Now consider the extension f of f_0 with $\|f\|_{X^*} = \|f_0\|_{L_1^*}$ (whose existence is guaranteed by the Hahn-Banach Theorem). □

Consider for any $L \hookrightarrow X$, $L^\perp = \{f \in X^* \mid f(L) = 0\} \hookrightarrow X^*$.

Corollary 3.1.3 *Let L be a closed subspace. Then consider $L^\perp \hookrightarrow X^*$ and $(L^\perp)^\perp = \{x \in X \mid f(x) = 0 \; \forall f \in L^\perp\}$. Then $(L^\perp)^\perp = L$.*

Proof: Clearly $L \hookrightarrow (L^\perp)^\perp$ (just check the definitions).

Now, for every $x \notin L$ and L closed we have that $d(x, L) = d > 0$. By the fifth item above there exists f such that $f(L) = 0$ (i.e. $f \in L^\perp$) and $f(x) \neq 0$. Hence $x \notin (L^\perp)^\perp$. □

Proposition 3.1.4 (Biorthogonal system) *Let $\{x_1, \ldots, x_n\} \subset X$ be a linearly independent subset of X. Then, there exists $f_1, \ldots, f_n \subset X^*$ such that $f_i(x_j) = \delta_{ij}$.*

Proof: Fix i_0; let $\operatorname{span}\{X_i\}_{i\neq i_0} = L_{i_0}$; note that L_{i_0} is a closed subspace (non-trivial) . $x_{i_0} \notin L_{i_0}$; by the fifth item above there exists $f_{i_0} \in X^*$ such that $f_{i_0}(L_{i_0}) = 0$, $f_{i_0}(x_{i_0}) = 1$. □

3.2 Dual Spaces

In section 2.3 we show a few examples that we now revisit using the terminology of the dual space.
Examples:

1. On the space c_0 of null sequences with the norm $\|x\| = \max |a_i|$ we define the linear functional $f = (b_i) \in \ell_1$. by setting $f(x) = \sum a_i b_i$. Then $|f(x)| \leq \sum |a_i b_i| \leq \max_{1\leq i\leq\infty} |a_i| \cdot \sum |b_i| = \|x\|_{c_0} \cdot \|f\|_{\ell_1}$. So, $\|f\|^* \leq \|f\|_{\ell_1}$.

   Now, let $f \in c_0^*$; define $f(e_n) = a_n$ ($e_n = (0, \dots, 0, 1, 0, \dots) \in c_0$ where the 1 occures in the n-th position). Take $y_n = \sum_1^n (\operatorname{sign} a_i) e_i$, $\|y_n\|_{c_0} = 1$:

   $$\text{(3.3)} \qquad \|f\|_{c_0^*} \geq f(y_n) = \sum_1^n |a_i| \quad (\forall n \in \mathbb{N}) \Rightarrow \|f\|_{c_0^*} \geq \|f\|_{\ell_1} \ .$$

   Thus $\ell_1 = (c_0)^*$.

2. $\ell_1^* = \ell_\infty$ (an exercise).

3. For $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$ we have $(\ell_p)^* = \ell_q$:

   Again, first check that if $f = (b_i) \in \ell_q$ then f defines a linear functional on ℓ_p by the following formula: if $x = (a_i)$

   $$\text{(3.4)} \qquad f(x) = \sum_1^\infty a_i b_i$$

   and

   $$\text{(3.5)} \qquad |f(x)| \leq |\sum a_i b_i| \leq (\sum |a_i|^p)^{1/p} \cdot (\sum |b_i|^q)^{1/q} \ .$$

   So $\|f\|_{\ell_p^*} \leq \|f\|_{\ell_q}$.

   Now, let $f \in \ell_p^*$. Consider $f(e_n) = c_n$. Take $y_n = \sum_1^n (\operatorname{sign} c_i) \cdot |c_i|^{q-1} e_i$.

   $$\text{(3.6)} \qquad \|y_n\|_{\ell_p} = \left(\sum_1^n |c_i|^{(q-1)p} \right)^{1/p} = \left(\sum_1^n |c_i|^q \right)^{1/p} \ .$$

$$\|f\|_{\ell_p^*} \cdot \|y_n\|_{\ell_p} \geq |f(y_n)| = \sum_1^n |c_i|^q$$
$$= \left(\sum_1^n |c_i|^q\right)^{1/p} \cdot \left(\sum_1^n |c_i|^q\right)^{1/q} = \|y_n\|_{\ell_p} \cdot \left(\sum_1^n |c_i|^q\right)^{1/q} \tag{3.7}$$

thus $\|f\|_{\ell_p^*} \geq \|(c_n)\|_{\ell_q}$ which gives the inverse inequality. Hence $(\ell_p)^*$ can be isometrically realized as the space ℓ_q.

4. Similarly $L_p^* = L_q$ $(1 < p < \infty)$.

3.3 Exercises:

1. Let E be an n-dimensional normed space, then E is complete.
2. If E is a finite dimensional subspace of X, then E is a closed subspace.
3. Prove that for $p > 1$, $\ell_p^* = \ell_q$ where $1/p + 1/q = 1$, i.e. there exists a one-to-one correspondence $f \leftrightarrow y$ for $f \in \ell_p^*$, $y \in \ell_q$ such that $\|f\|_{\ell_p^*} = \|y\|_{\ell_q}$
4. Prove that $\ell_1^* = \ell_\infty$.
5. (a) Is it true that for all $f \in X^*$ such that $\|f\| = 1$ there exists $x \in X$ so that $\|x\| = 1$ and $f(x) = 1$?

   (b) What is the answer if X is reflexive?
6. Prove that if $E \hookrightarrow X$ then $(X/E)^* = E^\perp$.
7. Prove that if $F \hookrightarrow X$ then $F^* = X^*/F^\perp$.

Chapter 4

Bounded linear operators

LET X and Y be Banach spaces, $T : X \to Y$, a linear map (operator) defined on X. T is called bounded if there exists C such that $\|Tx\|_Y \leq C\|x\|_X$ for all $x \in X$. If T is bounded, define $\|T\| = \sup_{x \neq 0} \|Tx\|/\|x\|$. One may check that this quantity defines a norm (check it!). We write $L(X \to Y)$ for the linear space of bounded operators with the above norm.

4.1 Completeness of the space of bounded linear operators

Theorem 4.1.1 *Let X be a normed space and Y be a complete normed space. Then $L(X \to Y)$ is a Banach space (i.e. complete).*

Proof: Let $\{A_n\}_{n=1}^{\infty}$ be a Cauchy series in $L(X \to Y)$ so

(4.1) $$\forall \epsilon \ \exists N \text{ such that } \forall m, n \geq N \quad \|A_n - A_m\| \leq \epsilon.$$

This implies that for every $x \in X$ and $M, n \geq N$

$$\begin{aligned} \|A_n(x) - A_m(x)\|_Y &= \|(A_n - A_m)(x)\|_Y \\ &\leq \|A_n - A_m\| \cdot \|x\| \\ &\leq \epsilon\|x\| \end{aligned}$$

Therefore: for all $x \in X$, the sequence $\{A_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence in Y. Since Y is a Banach space, it has a limit we can call

$A(x) \in Y$ and thus we define for all $x \in X$, $A(x) = \lim_{n\to\infty} A_n(x)$. A is a linear operator and it is also bounded since:

$$\|A(x)\|_Y \leq \sup_{n\in N} \|A_n(x)\| \leq \|x\|_X \cdot \sup_{n\in N} \|A_n\| \Rightarrow \|A\| \leq \sup_{n\in N} \|A_n\|, \tag{4.2}$$

thus $A \in L(X \to Y)$. Now we still must show that $A_n \to A$. If we assume otherwise, that is $\|A_n - A\| \not\to 0$, then there exist $\varepsilon > 0$ and $\{A_{n_k}\}_{k=1}^{\infty} \subseteq \{A_n\}_{n=1}^{\infty}$ such that for every $k \in \mathbb{N}$ we have $\|A_{n_k} - A\| \geq 2\epsilon$. Therefore for every $k \in \mathbb{N}$, we can choose $x_k \in X$ such that $\|x_k\| = 1$ and $\|A_{n_k}(x_k) - A(x_k)\| \geq \epsilon$. Recall that $\{A_n\}_{n=1}^{\infty}$ is a Cauchy sequence, so we can choose $\nu \in N$ such that for all $m, n_k \geq \nu$ we have $\|A_{n_k}(x_k) - A_m(x_k)\| \leq \frac{\epsilon}{2}$ and this implies

$$\begin{aligned} \epsilon &\leq \|A_{n_k}(x_k) - A(x_k)\| \\ &\leq \|A_{n_k}(x_k) - A_m(x_k)\| + \|A_m(x_k) - A(x_k)\|. \end{aligned}$$

Hence for all $m \geq \nu$

$$\|A_m(x_k) - A(x_k)\| \geq \frac{\epsilon}{2} \tag{4.3}$$

contradicting the definition of A (we must have $A_m(x_k) \to A(x_k)$). □

Note (i) A is a bounded operator if and only if A is a continuous operator (i.e. $Ax_n - Ax \to 0$ for $x_n \to x$)

(ii) $\ker A = \{x \mid Ax = 0\}$ is a closed subspace.

(iii) Theorem 4.1.1 implies that for any normed space X the dual space X^* is complete. Indeed, take Y to be the field $\mathbb{R}$ or $\mathbb{C}$ depending over which field our original space X is).

4.2 Examples of linear operators

1. In $C[0,1]$ define $Af = \int_0^1 K(t,\tau)f(\tau)d\tau$ (for a continuous function K of two variables). A is linear and $\|Af\|_{C[0,1]} \leq \max|f| \cdot \max_t \int_0^1 |K(t,\tau)|d\tau$. So $\|A\| \leq \max_t \int_0^1 |K(t,\tau)|d\tau$. In fact one may show that
$$\|A\| = \max_t \int_0^1 |K(t,\tau)|d\tau. \tag{4.4}$$

2. In $L_2[0,1]$ for $K(t,\tau) \in L_2([0,1]^2)$, define the operator
$$K : L_2[0,1] \mapsto L_2[0,1]$$

with

$$Kf = \int_0^1 K(t,\tau)f(\tau)d\tau. \tag{4.5}$$

The function of two variables $K(T,\tau)$ is called the *kernel* (or the *kernel function*) of the operator K. Check (as an exercise) that

$$\|K\|_{\mathrm{op}} \le \|K(t,\tau)\|_{L_2(I^2)}\ . \tag{4.6}$$

3. For every bounded linear $A : H \to H$ operator we have: $\|A\| = \sup\{|(Ax,y)| \mid \|x\| \le 1\ ,\ \|y\| \le 1\}$.

4. Let $k(t)$ be a continuous function on $[a,b]$. In $L_2[a,b]$ define the operator A by $Af = k(t)\cdot f(t)$. Then A is a bounded linear operator and
$$\|A\| = M = \max_{a\le t\le b} |k(t)|. \tag{4.7}$$

5. The shift operator in ℓ_2 defined by: $Tx = (0, a_1, \ldots, a_n, \ldots)$ for $(a_n) \in \ell_2$ satisfies $\|Tx\| = \|x\|$ and $\|T\| = 1$.

6. Let $(a_{ij})_{i,j=1}^{\infty}$ be an infinite matrix and $K^2 = \sum_{i,j=1}^{\infty} |a_{ij}|^2 < \infty$. Then the operator A defined in ℓ_2 by
$$A((\alpha_i)_{i=1}^{\infty}) = \left(\beta_i = \sum_{j=1}^{\infty} a_{ij}\alpha_j\right)_{i=1}^{\infty} \tag{4.8}$$
is a bounded linear operator. Check that $\|A(\alpha_i)\|_{\ell_2} \le K\cdot\|(\alpha_i)\|_{\ell_2}$.

7. Let H be a separable Hilbert space and $\{e_i\}_{i=1}^{\infty}$ be an orthonormal basis; Let $A : H \to H$ be a bounded operator. Then for $x = \sum_{i\ge1}(x,e_i)e_i$ we have $Ax = \sum(x,e_i)Ae_i$. Moreover $Ae_j = \sum(Ae_j,e_i)e_i$. Thus $Ax = \sum_i\big(\sum_j \alpha_j(Ae_j,e_i)\big)e_i$. Hence, the sequence (α_i) maps to $(\beta_i = \sum_j(Ae_j,e_i)\alpha_j)_{i=1}^{\infty}$. Consequently, we see that the example 6 may not be applicable to the matrix $((Ae_j,e_i))$.

The reader should now try the exercises from 1 to 5

4.3 Compact operators

$A : X \to Y$ is a *compact* operator if and only if for every bounded sequence $x_n \in X$ the sequence $\{Ax_n\}$ has a convergent subsequence.

4.3.1 Compact sets

In order to be able to work with compact operators we should first understand well the notion of compact set.

A set $K \subseteq X$ is called a *compact set* if and only if for every sequence $x_n \in K$ there exists a subsequence $x_{n_i} \to x \in K$ (for some $x \in K$). K is *relatively compact* (or *precompact*) if every sequence $x_n \in K$ has a Cauchy subsequence x_{n_i}. For example, if X is complete and K relatively compact then $\overline{K}$ is compact.

Example Any bounded set in $\mathbb{R}^n$ is a relative compact.

We will use the following statement which is standard in the calculus courses:

Theorem 4.3.1 (Arzelá) *Let $M \subset C[a,b]$; M is relatively compact (in $C[a,b]$) if and only if M is*

1. *uniformely bounded [meaning bounded set in $C[a,b]$] and*
2. *equicontinuous: $\exists\, \omega_M(\varepsilon) \searrow 0$ for $\varepsilon \to 0$ such that if $|x_1 - x_2| < \varepsilon$ then $|f(x_1) - f(x_2) < \omega_M(\varepsilon)$ for any $f \in M$.*

Note that if F is any metric precompact space (instead of the interval $[a,b]$) then the same theorem is true for $M \subseteq C(F)$.

Examples:

1. Let $M = \{x(t) \in C[a,b] \mid |x(t)| < C_1 \text{ and } |x'(t)| < C_2\} \subseteq C[a,b]$ (for some constants C_1 and C_2). Then M is relatively compact (use Arzelá Theorem: $|x'(t)| < C_2$ implies the uniform continuity: $\frac{x(t_1)-x(t_2)}{t_1-t_2} = x'(\theta)$ for $t_1 < \theta < t_2$, and therefore, $|x(t_1) - x(t_2)| \leq C_2 \cdot |t_1 - t_2|$).
2. The operator $Ax = \int_0^t x(\tau)d\tau$ on $C[0,1]$ is a compact operator.
3. The embedding operator: $A : C_1[a,b] \to C[a,b]$, $A(x) = x$ is a compact operator ($x \in C_1[a,b] : \|x\|_{C_1} = \max_t |x(t)| + \max_t |x'(t)|$).
4. Let $K(t,\tau)$ be a continuous function of two variables on $[0,1]^2$. Then the operator $Kx = \int_0^1 K(t,\tau)x(\tau)d\tau : C[0,1] \to C[0,1]$ is a compact operator (check it!). [Weaker conditions imposed on $K(t,\tau)$ would give the same result. For example, $K(t,\tau)$ may be piecewise continuous with a few discontinuity curves $\tau = \varphi_k(t)$, $k \in \mathbb{N}$.]

To build more examples we need some properties of compact sets and operators.

Definition 4.3.2 *Let A be any metric space and $\mathfrak{A} \subset A$. We call $\mathfrak{A}$ an ε-net of A if and only if for all $x \in A$ there exists $y \in \mathfrak{A}$ such that the distance from x to y is less than ε.*

Lemma 4.3.3 *M is relatively compact if and only if for every $\varepsilon > 0$ there exists a finite ε-net in M.*

Proof: " $\Rightarrow$ " assuming that there exists $\varepsilon_0 \geq 0$ such that there is no finite ε_0-net of M, then we can choose $x_1, x_2 \in M$ such that $\|x_1 - x_2\| \geq \varepsilon_0$. In this way for every n we can choose $x_n \in M$ such that $\|x_n - x_1\|, \|x_n - x_2\| \ \ldots \ \|x_n - x_{n-1}\| \geq \varepsilon_0$. Such an x_n exists for every n, otherwise $\{x_1, x_2, \ldots, x_{n-1}\}$ is a finite ε_0-net of M. Therefore no sub-sequence of $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence, which means that M is not pre-compact; a contradiction.

" $\Leftarrow$ " For all $k \in N$ we take $\varepsilon_k \leq \frac{1}{2k}$. M has a finite ε_k-net for every k (by the assumption). Let $\{x_n\}_{n=1}^{\infty} \in M$ and consider a finite ε_1-net. It divides the sequence between a finite number of balls around the net's points. So there is at least one ball which contains an infinite subsequence from the original sequence. Let us mark the sub-sequence contained within this ball by $\{x_n^{(1)}\}_{n=1}^{\infty}$. In a similar way an ε_k-net divides the sequence $\{x_n^{(k-1)}\}_{n=1}^{\infty}$ between a finite number of balls and there is one ball which contains an infinite subsequence of the previously chosen sequence $\{x_n^{(k-1)}\}$. Let us call it $\{x_n^{(k)}\}_{n=1}^{\infty}$.

We know that the ball's radius is less than ε_k hence for all $m, n \in N$ we have

$$\|x_m^{(k)} - x_n^{(k)}\| \leq 2\varepsilon_k \leq \frac{1}{k}.$$

The subsequence $\{x_n^{(n)}\}_{n=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Indeed, for all $\eta \geq 0$ we can choose $\nu \in N$ such that $\frac{1}{\nu} \leq \eta$ and then

$$\forall m, n \geq \nu \quad \|x_n^{(n)} - x_m^{(m)}\| \leq \frac{1}{\nu} \leq \eta \tag{4.9}$$

which imples that M is relatively compact.

4.3.2 The space of compact operators

Proposition 4.3.4 *The set $K(X) \equiv K(X \to X)$ of compact operators on X satisfy:*

(i) $K(X \to X)$ is a linear subspace of $L(X \to X)$ (check it!).

(ii) $K(X)$ is a two-sided ideal of $L(X \to X) \equiv L(X)$.

(iii) $K(X)$ is a closed subspace of L:

Proof: (ii) Let $A \in K(X \to X)$, $B \in L(X \to X)$ and $\{x_n\}_{n=1}^{\infty}$ bounded sequence in X. We want to show that $AB, BA \in K(X \to X)$.

The operator AB: B is bounded hence $\{B(x_n)\}_{n=1}^{\infty}$ is bounded also. A is compact hence $\{A(B(X_n))\}_{n=1}^{\infty}$ has a converging subsequence thus AB is compact.

The operator BA: A is compact hence the sequence $\{A(x_n)\}_{n=1}^{\infty}$ has a converging subsequence $\{A(x_{n_k})\}_{k=1}^{\infty}$. B is bounded, so it follows that the sequence $\{B(A(x_{n_k}))\}_{k=1}^{\infty}$ is also converging thus BA is compact.

(iii) Let $A_n \to A$ (meaning $\|A_n - A\| \to 0$) and $A_n \in K$. It is enough to prove that $A\mathcal{D}(X)$ (the image of the unit ball) is a precompact. Thus we have to find for all $\varepsilon > 0$ a finite ε-net.

The reader should now try the exercises 6,7,8

For all $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $\|A_n - A\| < \varepsilon/2$ and $A_n\mathcal{D}(X)$ is a precompact (A_n is a compact operator). Then take an $\varepsilon/2$ net $\{x_i\}_1^N$ of $A_n\mathcal{D}(X)$. Check (an exercise) that $\{x_i\}_1^N$ is an ε-net for $A\mathcal{D}(X)$. [Additional delicate point to think is that $\{x_i\}_1^N$ may not be in $A\mathcal{D}(X)$; but it is not important.]

4.4 Dual Operators

Let $A : X \to Y$ be a bounded operator. Then $\varphi(Ax) = f(x)$ ($x \in X$; $\varphi \in Y^*$) is a linear function on X. Moreover,

$$|f(x)| \leq \|\varphi\|_{Y^*} \cdot \|A\| \cdot \|x\|. \tag{4.10}$$

Thus $f \in X^*$. Hence, we have an operator $\varphi \mapsto A^*\varphi = f;\quad A^* : Y^* \to X^*$ is linear (obvious) and

$$\|A\| \quad = \quad \sup_{\|x\|=1} \|A(x)\|_Y = \sup_{\|x\|=1} \sup_{|\varphi\|=1} |\varphi(A(x))|$$

$$
\begin{aligned}
&= \sup_{\|\varphi\|=1}\sup_{\|x\|=1}|\varphi(A(x))| = \sup_{\|\varphi\|=1}\sup_{\|x\|=1}|A^*(\varphi)(x)| \\
&= \sup_{\|\varphi\|=1}\|A^*(\varphi)\|_{X^*} = \|A^*\| \qquad (4.11)
\end{aligned}
$$

Theorem 4.4.1 *$A : X \to Y$ is compact implies $A^* : Y^* \to X^*$ is compact.*

Proof: We show that $A^*\mathcal{D}(Y^*) = K \subset X^*$ is precompact. Indeed, we use the Arzelá Theorem. First "represent" the set $A^*(\mathcal{D}(Y^*)) = K$ as a set of continuous functions on the precompact set $T = A\mathcal{D}(X)$.

Let $f \in A^*\mathcal{D}(Y^*)$; $f(x) = (A^*\varphi)(x)$ for $x \in \mathcal{D}(X)$ and some $\varphi \in \mathcal{D}(Y^*)$. Then $f(x) \in \varphi(Ax)$ and φ is a (linear) continuous function on $T = A\mathcal{D}(X)$. Moreover,

$$
\begin{aligned}
\|f_1 - f_2\|_{X^*} &= \sup_{\|x\|\le 1}|(A^*\varphi_1 - A^*\varphi_2)(x)| = \sup_{\|x\|\le 1}|(\varphi_1-\varphi_2)(Ax)| \\
&= \sup_{y\in T}|(\varphi_1-\varphi_2)(y)|, \qquad (4.12)
\end{aligned}
$$

so $\mathrm{dist}(f_1, f_2) = \|\varphi_1 - \varphi_2\|_{C(T)}$. Thus, we may use Arzelá Theorem for the set of functions $\{\varphi \in \mathcal{D}(Y^*)\}$ on the precompact T. This set is bounded (by $\|A\|$) and equicontinuous: $\omega_\varphi(\varepsilon) = \sup_{\|x_1-x_2\|\le\varepsilon}|\varphi(x_1 - x_2)| \le \|\varphi\|\cdot\varepsilon \le \varepsilon$ (independent of φ). □

So, $A \in K(X \to Y)$ (A is a compact operator) implies that $A^* \in K(Y^* \to X^*)$ (A^* is also a compact operator).

More examples: (i) We call the quantity dim(ImA) the *rank* of the operator A and we write rk$A$ =dim(ImA). We say that the operator A has finite rank if rk$A < \infty$. For an example of an operator of finite rank consider a finite number of elements $y_i \in Y$ and $f_i \in X^*$ where X and Y are Banach spaces and $i = 1, 2, \ldots, n$ for some $n \in \mathbb{N}$. Define the operator A by setting $Ax = \sum_1^n f_i(x)y_i$. This operator has rank no greater than n and it is a bounded operator since $\|A\| \le \sum_1^n \|f_i\|\cdot\|y_i\|$. Check that its dual operator is $A^*(\cdot) = \sum_1^n(\cdot)(x_i)f_i$ (a "better notation" for A is $A = \sum f_i \otimes x_i$)).

Inverse: If A is bounded and $\mathrm{rk}\, A = n$, i.e. $\mathrm{Im}A = L$, $\dim L = n$, choose $\{x_i\}_i^n$ a basis in L and let $\{x_i^*\}_1^n \subset X^*$ be a biorthogonal system (as in the Corollary of the Hahn-Banach Theorem). Then $Ax = \sum_1^n x_i^*(Ax)x_i = \sum_1^n(A^*x_i^*)(x)x_i$. Let $f_i = A^*x_i^*$. We then have, $A = \sum_1^n f_i \otimes x_i$. (Obviously: $\|f_i\| \le \|A\|\cdot\|x_i^*\|$, so $f_i \in X^*$.) Check that every bounded operator of finite rank is a compact operator.

(ii) Consider $K(t,\tau) \in L_2(I^2)$ $(I = [0,1])$ and let

$$C = \|K(t,\tau)\|_{L_2(I^2)} = \sqrt{\int_0^1 \int_0^1 \|K(t,\tau)\|^2 dt\, d\tau}. \tag{4.13}$$

Define the operator K on $L_2[0,1]$ by $Kx = \int_0^1 K(t,\tau)x(\tau)d\tau$. Then K is a compact operator in $L_2[0,1]$ and $\|K\|_{\text{op}} \le C = \|K\|_{L_2(I^2)}$. Indeed: first check that $\|K\|_{\text{op}} \le \|K\|_{L_2(I^2)}$; for $\{\varphi_i(t)\}_1^\infty$-orthonormal basis in $L_2[0,1]$, we checked before that $\{\varphi_i(t)\varphi_j(\tau)\}$ is an orthonormal basis in $L_2(I^2)$; so $K(t,\tau) = \sum a_{ij}\varphi_i(t)\varphi_j(\tau)$. Let $K_n = \sum_{i=1}^n \sum_{j=1}^n a_{ij}\varphi_i(t)\varphi_j(\tau)$. Then $\|K_n - K\|_{L_2(I^2)} \longrightarrow 0$ as $n \to \infty$, therefore, considering the operator

$$K_n x = \int_0^1 K_n(t,\tau)x(\tau)d\tau, \tag{4.14}$$

The reader should now try the exercise 9

we also know that $\|K_n - K\|_{\text{op}} \longrightarrow 0$ as $n \to \infty$. K_n is an operator of rank $\le n$ and this means that K is approximable by finite rank operators, which are compact. This implies that the limit operator K is also compact.

4.5 Different convergences in the space $\mathrm{L}(X)$ of bounded operators

In the space of operators $L(X)$ one may define several notions of "convergence". The *norm convergence*, also called "*uniform convergence*" is defined by saying that the sequence of operators A_n converges in norm to the operator A and we write $(A_n \rightrightarrows A)$ if $\|A_n - A\| \longrightarrow 0$ as $n \to \infty$. $L(X)$ is complete and so if $\{A_n\}$ is a Cauchy sequence with respect to the norm then it always converges to a bounded operator.

An other usefull notion is that of the strong convergence: $A_n \to A$ strongly if for all $x \in X$ we have $A_n x \to Ax$. We note here that if the sequence $\{A_n\}$ is Cauchy in the strong sense, that is for all $x \in X$ the sequence $A_n x$ is Cauchy in X, then there exists $A \in L(X)$ such that $A_n \to A$ strongly. The proof of this fact will be given after the Banach-Steinhaus theorem.

Let us give an example in order to show that norm convergence and strong convergence do not coincide. Consider the projections in $L_2[0,1]$ defined by

(4.15) $$P_{\varepsilon_n} f = \begin{cases} f(t) & \text{for } t < \varepsilon_n; \\ 0 & \text{for } t \geq \varepsilon_n \end{cases}$$

It is easy to see that $P_{\varepsilon_n} \to 0$ as $\varepsilon_n \to 0$,that is P_{ε_n} converges strongly to zero, but $\|P_{\varepsilon_n}\| = 1$ and consequently it does not converge in norm to zero.

Our third notion of convergence is that of the weak convergence: A_n converges *weakly* to A, and we write $A^n \rightharpoonup A$ or $A_n \xrightarrow{w} A$, if for all $x \in X$ and for all $f \in X^*$ we have $f(A_n x) \to f(Ax)$.

Again in order to distinguish this notion from that of the strong convergence, consider the example of the shift operator A in ℓ_2. For this operator we have $A^n \rightharpoonup 0$, but $\|Ax\| = \|x\|$ and so A^n does not converge strongly.

We continue with two other main theorems of functional analysis (without proof).

Theorem 4.5.1 (Banach (on open map)) *Let X, Y be Banach spaces and let $A : X \to Y$ be a bounded linear operator one-to-one $[$i.e. $\ker A = 0]$ and onto $($meaning $\mathrm{Im} A = Y)$. Then the formally defined operator $A^{-1} : Y \to X$ is bounded.*

Theorem 4.5.2 (Banach-Steinhaus) *Let $A_\alpha : X \to Y$ be a family of bounded operators so that for every $x \in X$ there exists a constant $C(x)$ such that $\|A_\alpha x\| \leq C(x)$ (i.e. the family is pointwise bounded or bounded in the "strong" sense). Then there exists C such that $\|A_\alpha x\| \leq C\|x\|$ meaning that $\|A_\alpha\| \leq C$.*

In other words: *the strong boundness of a family implies uniform boundness.*

Examples

1. Let $Y = \mathbb{R}$ (or $\mathbb{C}$ in the complex case). Let A_α be linear functionals f_α. We have $\{|f_\alpha(x)| \leq C(x)$ for every $x \in X\}$ implies $\{\exists C : \|f_\alpha\| \leq C\}$.
2. If $\{x_\alpha\} \subseteq X$ and for all $f \in X^*$, there exists constant $C(f)$ so that

   (4.16) $$|f(x_\alpha)| \leq C(f) \ ,$$

   then there exists constant C independent of f such that $\|x_\alpha\| \leq C$.

3. Combining these applications, let $A_\alpha \in L(X \to Y)$ such that for all $x \in X$ and for all $f \in Y^*$ there exists constant $C(f;x)$ such that $|f(A_\alpha x)| \leq C(f,x)$. Then there exists constant C independent of f and x such that $\|A_\alpha\| \leq C$.

4.6 Invertible Operators

Let $A \in L(X)$. We call $B(= A^{-1})$ the inverse operator if and only if $BA = \mathrm{Id}$ and $AB = \mathrm{Id}$. In the finite dimensional case the notion of the determinant implies that $BA = Id$ is enough to deduce that A is invertible and $A^{-1} = B$. The reason is that $\det A \neq 0$ leads to a formula for A^{-1}. In the infinite dimensional case though this is not the case. One may consider for example the case of the shift operator in ℓ_2.

Properties

1. $(AB)^{-1} = B^{-1}A^{-1}$ (meaning that if A and B invertible, then also AB is invertible and it is computed by the previous formula).

2. If $\|A\| < 1$, then the inverse exists: $(I-A)^{-1} = \sum_0^\infty A^k$. Moreover $\|(I-A)^{-1}\| \leq 1/(1-\|A\|)$. Indeed, first note that $\|A^k\| \leq \|A\|^k \to 0$ as $k \to \infty$. Let $S_n = \sum_0^n A^k$. Then $S_n \to B$ (S_n-Cauchy sequence). Also $(I-A)S_n = I - A^{n+1} \longrightarrow I$ as $n \to \infty$ and $S_n(I-A) \longrightarrow I$ as $n \to \infty$.

3. Let A be an invertible operator and B be such that $\|A - B\| < 1/\|A^{-1}\|$. Then B is also invertible. Indeed, write $B = A[I - A^{-1}(A-B)]$. A is invertible and $(I - A^{-1}(A-B))$ is invertible by 2 because $\|A^{-1}(A-B)\| < 1$. Hence their product is invertible.

4.7 Exercises

1. Let A be an operator on $C[0,1]$ which is given by $(Af)(x) = a(x)f(x)$ where $a(x)$ is a continuous function on $[0,1]$. Prove that A is bounded and compute its norm.

2. Let $(w_j)_{j=1}^\infty$ be a sequence of complex numbers. Define an operator D_w on ℓ_2 by $D_w x = (w_1x_1, w_2x_2, \ldots)$ for $x = (x_1, x_2, \ldots) \in \ell_2$. Prove that D_w is bounded if and only if $(w_j)_{j=1}^\infty$ is bounded and in this case $\|D_w\| = \sup_j |w_j|$.

3. Let $\ell_2(\mathbb{Z})$ be the Hilbert space of all sequences $(x_j)_{j=-\infty}^{\infty}$ with
$$\sum_{j=-\infty}^{\infty} |x_j|^2 < \infty$$
and the usual inner product. Define an operator S on $\ell_2(\mathbb{Z})$ by
$$S(x_j)_{j=-\infty}^{\infty} = (x_{j-1})_{j=-\infty}^{\infty}.$$
   (a) Prove that $\|Sx\| = \|x\|$ for any $x \in \ell_2(\mathbb{Z})$.
   (b) Give a formula and a matrix of representation of the operators S^n for $n \in \mathbb{Z}$.

4. Given an infinite matrix $(a_{ij})_{i,j=1}^{\infty}$, where $\sum_{i=1}^{\infty} \sum_{i=1}^{\infty} |a_{ij}|^2 < \infty$, define $A : l_2 \to l_2$ by $A(x_1, x_2, \ldots) = (y_1, y_2, \ldots)$, where
$$(a_{ij}) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \end{pmatrix} \quad , i.e. y_i = \sum_{i=1}^{\infty} a_{ij}x_j.$$
Prove that the operator A is a bounded linear operator on l_2 and $\| A \|^2 \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|^2$.

5. Let A be an operator on ℓ_2 given by the matrix $(a_{jk})_{j,k=1}^{\infty}$ (with respect to the standard basis), where for some fixed $m, n \in \mathbb{N}$ we have that $a_{jk} = 0$ for $j - k < -m$ or $j - k > n$ (so that A has only a finite number of non-zero diagonal entries).
   (a) Prove that A is bounded if and only if
$$\sum_{k=-m}^{n} \sup_j |a_{j,j-k}| < \infty.$$
   (b) Prove that
$$\|A\| \leq \sum_{k=-m}^{n} \sup_j |a_{j,j-k}|.$$
(These kind of matrices are called *band* matrices.)

6. Let H_1 and H_2 be Hilbert spaces. Define $H = H_1 \oplus H_2$ to be the Hilbert space consisting of all pairs (u_1, u_2) with $u_1 \in H_1$ and $u_2 \in H_2$ with

$$(u_1, u_2) + (v_1 + v_2) = (u_1 + v_1, u_2 + v_2) \tag{4.17}$$
$$\lambda(u_1, u_2) = (\lambda u_1, \lambda_2) \tag{4.18}$$

and an inner product defined by

$$\langle (u_1, u_2), (v_1, v_2) \rangle = \langle u_1, v_1 \rangle_{H_1} + \langle u_2, v_2 \rangle_{H_2}.$$

Given $A_1 \in L(H_1)$ and $A_2 \in L(H_2)$ define A on H by the matrix

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

i.e. $A(u_1, u_2) = (A_1 u_1, A_2 u_2)$. Prove that A is in $L(H)$ and that $\|A\| = \max(\|A_1\|, \|A_2\|)$.

7. Which of the following operators $K : L_2[a, b] \to L_2[a, b]$ have finite rank and which do not?

   (a) $(Kf)(t) = \sum_j^n = 1\varphi_j(t) \int_a^b \psi_j(s) f(s) ds$.

   (b) $(Kf)(t) = \int_a^t \varphi(s) ds$.

8. Let $(w_j)_{j=1}^\infty$ be a sequence of complex numbers. Define an operator D_w on l_2 by $D_w x = (w_1 x_1, w_2 x_2, \ldots)$. Prove that D_w is compact if and only if $\lim_{j\to\infty} w_j = 0$.

9. Let $(a_j)_{j=1}^\infty$ be a sequence of complex numbers with $\sum_{j=1}^\infty |a_j| < \infty$. Define an operator on l_2 by the matrix

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \ldots \\ a_2 & a_3 & \ldots & \ldots \\ a_3 & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

Prove that A is compact.

10. Let $T : L_p(-\infty, \infty) \mapsto L_p(-\infty, \infty)$ for $1 \le p < \infty$ with $(Tf)(t) = f(t+1)$. Find the operator T^*.

11. Let D_w be as in exercise 2 and let $\inf |w_j| > 0$ and $\sup_j |w_j| < \infty$. Which of the following equalities or inequalities hold for any w?

    (a) $\|D_w\| = \frac{1}{\|D_w^{-1}\|}$; (b) $\|D_w\| \geq \frac{1}{\|D_w^{-1}\|}$; (c) $\|D_w\| \leq \frac{1}{\|D_w^{-1}\|}$;
    (d) $\|D_w\| < \frac{1}{\|D_w^{-1}\|}$; (e) $\|D_w\| > \frac{1}{\|D_w^{-1}\|}$

12. Let the operator D_w be as in the previous problem. Prove that D_w is invertible if and only if $\inf_j |w_j| > 0$. Give a formula for D_w^{-1}.

13. Let K be an operator of finite rank on the Hilbert space H. For $\phi \in H$ assume

    $$K\phi = \sum_{i=1}^{m} \langle \phi, \phi_i \rangle \psi_i.$$

    Suppose that $\psi_i \in (\text{span}\{\phi_1, \dots \phi_m\})^{\perp}$ for $i = 1, 2, \dots, n$. Prove that $I + \lambda K$ is invertible for any λ and find its inverse.

14. Let H be a Hilbert sapce and let $B.C, D \in L(H)$. On $H^{(3)} = H \oplus H \oplus H$ define A by the matrix

    $$A = \begin{pmatrix} 0 & D & B \\ 0 & 0 & C \\ 0 & 0 & 0 \end{pmatrix}.$$

    Prove that:

    (a) $A \in L\left(H^{(3)}\right)$;
    (b) $I - \lambda A$ is invertible for any $\lambda \in \mathbb{C}$ and find its inverse.

    (The norm of $H^{(3)}$ is defined by

    $$\|h\| = \sqrt{\sum_{i=1}^{3} \|h_i\|^2}$$

    for $h = (h_i)_{i=1}^{3} \in H \oplus H \oplus H$.)

15. Given $A_{jk} \in L(H)$ for $j, k = 1, 2$ define on $H^{(2)} = H \oplus H$ an operator A by

    $$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

    Prove that A is compact if and only if each A_{jk} is compact.

16. Suppose that $A, B \in L(H)$ and AB is compact. Which of the following statement must be true?

    (a) Both A and B are compact.

    (b) At least A or B are compact.

17. Let $\ell_2(\mathbb{Z})$ be the Hilbert space of all sequences $(\xi_j)_{j=-\infty}^{\infty}$ that satisfy $\sum_{j=-\infty}^{\infty} |\xi_j|^2 < \infty$ and the usual inner product. Define an operator S on $\ell_2(\mathbb{Z})$ by $S(\xi_j)_{j=-\infty}^{\infty} = (\xi_{j-1})_{j=-\infty}^{\infty}$.

    (a) Prove that S is invertible. What is it's inverse?

    (b) Give a formula and a matrix of representation of the operators $(S^{-1})^n$ for $n \in \mathbb{Z}$.

18. Let $\mu = (\mu_k)_{k=1}^{\infty}$ be a sequence of complex numbers with $\sup_k |\mu_k| < 1$. Prove that the following system of equations has a unique solution in ℓ_2 for any $\{\eta_k\}$ in ℓ_2. Find the solutions for $\eta_k = \delta_{1k},\ \mu_k = 1/2^{k-l}$.

    (a) $\xi_k - \mu_k \xi_{k+1} = \eta_k,\ k = 1, 2, \ldots$

    (b) $\xi_k - \mu_k \xi_{k-1} = \eta_k,\ k = 2, 3, \ldots$

Chapter 5

Spectral theory

LET $A : X \to X$ be a bounded operator. A complex number $\lambda \in \mathbb{C}$ is called a *regular point* of A if and only if there exists $(A - \lambda I)^{-1}$ as a bounded operator from $X \mapsto X$ The rest $\lambda \in \mathbb{C}$ form a set which is called the "Spectrum of A" and we denote it by $\sigma(A)$. Thus $\sigma(A) \subseteq \mathbb{C}$.

From 4.6 it follows that the set of regular points is open; therefore, $\sigma(A)$ is a closed set.

5.1 Classification of spectrum

The points in the spectrum of an operator A can be categorized as follows:

(i) The point-spectrum σ_p is the set of the eigenvalues of the operator A, that is $\lambda \in \sigma_p$ if and only if there exists $x \in X \setminus 0$ and $Ax = \lambda x$. In other words λ is an eigenvalue and x is an eigenvector of A, corresponding to the eigenvalue λ. This is obviously equivalent to saying that $\ker(A - \lambda I) \neq 0$. So, next we assume that $\ker(A - \lambda I) = 0$, meaning that $A - \lambda I : X \to X$ is one-to-one correspondence between X and $\mathrm{Im}(A - \lambda I)$. By the Theorem of Banach, if $\mathrm{Im}(A - \lambda I) = X$ then there exists the bounded inverse $(A - \lambda I)^{-1}$. Such λ's are called *regular*.

So, in our classification of $\sigma(A)$, if $\lambda \notin \sigma_p$, but $\lambda \in \sigma(A)$, we conclude that $\mathrm{Im}(A - \lambda I) \neq X$.

Lemma 5.1.1 *Let $A : X \to X$ be any bounded operator and $\lambda_i \neq \lambda_j$*

for $i \neq j|_{i>1}^{n}$ *are distinct eigenvalues. Let* $x_i \neq 0$*,* $Ax_i = \lambda_i x_i$ *(eigenvectors of different eigenvalues). Then* $\{x_i\}_1^n$*, are linearly independent.*

Proof: Let $\alpha_1 x_i + \sum_2^n \alpha_i x_i = 0$. Take a polynomial $P(\lambda)$ such that $P(\lambda_1) = 1$ and $P(\lambda_i) = 0$ for $i \geq 2$. Note that $P(A)x_i = P(\lambda_i)x_i$ ($P(\lambda_i)$ is an eigenvalue of operator $P(A)$). Appling now $P(A)$ we get

$$0 = \alpha_1 P(A)x_1 + \sum_2^n \alpha_i P(A)x_i = \alpha_1 x_1.$$

So $\alpha_1 = 0$. We repeat the same procedure for the rest of α_i's. $\square$

(ii) The *continuous spectrum* σ_c is defined as follows: $\lambda \in \sigma_c$ if and only if $\lambda \in \sigma(A) \setminus \sigma_p(A)$ and

$\mathrm{Im}(A - \lambda I)$ is dense in X.

Example: On $L_2[0,1]$ define $A : L_2[0,1] \to L_2[0,1]$ with $Ax = t \cdot x(t)$. Then $[0,1] = \sigma_c(A) = \sigma(A)$.

(iii) The *residual spectrum* is the set $\sigma_r(A) = \sigma(A) \setminus (\sigma_p \cup \sigma_c)$ (whatever remains). So for $\lambda \in \sigma_r(A)$ we have, $\overline{\mathrm{Im}(A - \lambda I)} \neq X$.

Example Consider the shift operator $Ae_i = e_{i+1}$ on ℓ_2. Trivially, $0 \in \sigma_r(A)$. In fact for all λ with $|\lambda| < 1$, $\lambda \in \sigma_r(A)$.

Remark. We agree to write $(\lambda I)^*$ for λI or $\overline{\lambda} I$. Following the standard inner product notation, that is $(Ax, y) = (x, A^* y)$, we must choose $(\lambda I)^*$ to be $\overline{\lambda} I$.

The reader should now try the exercises from 1 to 5

5.2 Fredholm Theory of compact operators

We restrict now our attention on infinite dimensional Banach spaces. Let $T : X \mapsto X$ be a compact operator. Let T_λ denote the operator $T - \lambda I$ and $\Delta_\lambda = \mathrm{Im} T_\lambda$.

Lemma 5.2.1 *Let* E_1 *be a closed subspace of* E *and such that* $E \neq E_1 \hookrightarrow E \hookrightarrow X$*. Then there exists* $y_0 \in E$*,* $\|y_0\| = 1$ *and such that the distance dist*$(y_0, E_1) \geq \frac{1}{2}$*.*

Proof: Take any $y \in E \setminus E_1$; $\rho(y, E_1) = a > 0$ (such a y exists since E_1 is closed). Let $x_0 \in E_1$ be such that $\|y - x_0\| < 2a$. Then $y_0 = \frac{y - x_0}{\|y - x_0\|}$ satisfies the statement. Indeed, $\|y_0\| = 1$ and if we assume that there exists $z \in E_1$ so that $\|y_0 - z\| \le \frac{1}{2}$ then substituting y_0 in this last inequality we get

$$\Big\| y - x_0 - z\|y - x_0\| \Big\| \le \frac{1}{2}\|y - x_0\| < a,$$

a contradiction since $\rho(y, E_1) = a > 0$. □

Corollary 5.2.2 *If* $\dim X = \infty$*, then the identity operator* $I : X \to X$ *is not compact.*

Proof: It is enough to see that the unit ball $D(x) = \{x \mid \|x\| \le 1\}$ is not a compact set. Consider any family of subspaces $E_1 \hookrightarrow E_2 \hookrightarrow \cdots \hookrightarrow E_n \hookrightarrow \cdots$ where $\dim E_n = n$. They are closed subspaces (because they are of finite dimension) and by the previous Lemma, there exists a sequence $\{y_i \in E_i\}$, $\|y_i\| = 1$ such that $\rho(y_i, E_{i-1}) \ge \frac{1}{2}$. Obviously, $\{y_i\}$ is not a Cauchy sequence and there is no Cauchy subsequence of it (why?).

1. For every compact operator T, $0 \in \sigma(T)$. Indeed, if not, then there exists T^{-1} and $T^{-1}T = I$ is a compact operator (the compact operators form an ideal), which contradicts the previous corollary. So, next we assume that $\lambda \neq 0$.

2. For every $\varepsilon > 0$, there is only a finite number of linearly independent eigenvectors corresponding to the eigenvalues λ_i, $|\lambda_i| \ge \varepsilon$. I.e. there exists a finite number of $\lambda_i \in \sigma_p$, $|\lambda_i| \ge \varepsilon$, and every λ_i has finite multiplicity.

   Proof: If not, then there exists $\{x_i\}_1^\infty$ linearly independent vectors and $Tx_i = \lambda_i x_i$, $|\lambda_i| \ge \varepsilon$. Consider $E_k = \mathrm{span}\{x_i\}_1^k \not\subset E_{k+1}$. By the last lemma, take $y_k \in E_k$, $\|y_k\| = 1$ and $\rho(y_k, E_{k-1}) \ge \frac{1}{2}$. We will show that $\{T\frac{y_k}{\lambda_k}\}$ *does not contain Cauchy subsequences* which will mean that T is not compact because y_k/λ_k is bounded (here we used $|\lambda_k| \ge \varepsilon$).

   Indeed: Let $y_k = \sum_1^k a_i x_i$. Then $T\frac{y_k}{\lambda_k} = a_k x_k + \sum_1^{k-1} \frac{a_i \lambda_i}{\lambda_k} x_i = y_k + z_k$ for some $z_k \in E_{k-1}$. Then, for any $k > n$, we have

   $$\left\| T\frac{y_k}{\lambda_k} - T\frac{y_n}{\lambda_n} \right\| = \|y_k - \underbrace{(y_n - z_k + z_n)}_{\in E_{k-1}}\| \ge \frac{1}{2} \tag{5.1}$$

   (by the choice of $\{y_k\}$). Thus there is no Cauchy subsequence of $T(y_k/\lambda_k)$.

The structure of σ_p (the point spectrum) is now clear: it is at most a sequence λ_i converging to 0 and every λ_i has finite multiplicity. Next, we will show that no other spectrum exists, besides $\lambda = 0$, for a compact operator T.

3. $\Delta_\lambda = \overline{\Delta_\lambda}$ (So, Δ_λ is always a closed subspace.)

   Let $T_\lambda x = y$; denote $E_y = \{x \mid T_\lambda x = y\}$. Clearly $E_y = x + E_0$ where $E_0 = \ker T_\lambda$. We prove first the following lemma:

   Lemma 5.2.3 *Let $\alpha(y) = \inf\{\|x\| \mid x \in E_y\}$. Then there exists constant C independent of y such that $\alpha(y) \le C\|y\|$.*

   Proof: By the homogeneity of the inequality we seek, $[\alpha(cy) = c\alpha(y)$ for $c \ge 0]$, we can assume that if it is wrong then there exists y_n such that $\alpha(y_n) = 1$ and $y_n \to 0$. Let x_n be such that $T_\lambda x_n = y_n \to 0$ and $\|x_n\| \le 2$ (because $\alpha(y_n) = 1$, we may choose x_n to be in norm close to 1). By the compactness of T there exists a subsequence x_{n_k} such that $Tx_{n_k} \to z$ (and $(T - \lambda I)x_{n_k} \to 0$). So $\lambda x_{n_k} \to z$; but $x_{n_k} \to z/\lambda = x_0 \in E_0$. Then clearly $x_{n_k} - x_0 \in E_{y_{n_k}}$ and therefore $\alpha(y_{n_k}) \to 0$ which is a contradiction since we assumed that $\alpha(y_n) = 1$. □

   Now return to the proof of the 3rd statement. Let $y_n \in \Delta_\lambda$ and $y_n \to y$. Note that $\{y_n\}$ is bounded. By the lemma above there exists x_n, $\|x_n\| < C$ such that $y_n = T_\lambda x_n (\to y)$. Then there exists x_{n_k} such that $Tx_{n_k} \to z$ which implies $x_{n_k} \to \frac{z-y}{\lambda} = x_0$, and hence $Tx_0 - \lambda x_0 = y$. Thus $y \in \Delta_\lambda$.

3*. Consider the dual operator T^* which is also compact. Let $\Delta^*_{\bar\lambda} = \mathrm{Im} I^*_{\bar\lambda} \hookrightarrow X^*$. Then also $\Delta^*_{\bar\lambda} = \overline{\Delta^*_{\bar\lambda}}$ for any $\lambda \ne 0$. It can be also shown that $\Delta^*_{\bar\lambda} = \overline{\Delta^*_{\bar\lambda}}^{w^*}$ (we omit the proof). In fact a stronger statement will be used later. This is $((\Delta^*_\lambda)_\perp)^\perp = \Delta^*_\lambda$.

4. $\Delta_\lambda = X$ implies $\ker T_\lambda = 0$. (or, equivalently, $\ker T_\lambda \ne 0 \Rightarrow \Delta_\lambda \ne X$).

   Proof: If not, then there exists $x_0 \in \ker T_\lambda$ so that $x_0 \ne 0$. Hence there exists x_1, such that $T_\lambda x_1 = x_0$ ($\mathrm{Im} T_\lambda$ is the entire space). Similarly, for every k, there exists x_k such that $T_\lambda x_k = x_{k-1}$, $k = 1, 2, \ldots$ For such $x_k : T^k_\lambda x_k = x_0 \ne 0$ but $T^{k+1}_\lambda x_k = Tx_0 = 0$. Therefore, if $N_k = \{x \mid T^k_\lambda x = 0\} = \ker T^k_\lambda$, we have $N_{k+1} \not\hookrightarrow N_k$.

By the lemma 5.2.1, there exists $y_{k+1} \in N_{k+1}$, $\|y_k\| = 1$ and $\rho(y_{k+1}, N_k) \geq \frac{1}{2}$. Then $\{Ty_k\}$ does not contain a convergent subsequence (contradicting the compactness of T). Indeed, let $k > n$, then $z = y_n + T_\lambda \frac{y_k}{\lambda} - T_\lambda \frac{y_n}{\lambda} \in N_{k-1}$. Therefore

$$\begin{aligned} \|Ty_k - Ty_n\| &= \|T_\lambda y_k + \lambda y_k - T_\lambda y_n - \lambda y_n\| \\ &= \|\lambda y_k - (\lambda y_n + T_\lambda y_k - T_\lambda y_n)\| \\ &\geq |\lambda| \cdot \frac{1}{2}. \end{aligned}$$

4^*. Similarly $\Delta^*_{\overline{\lambda}} = X^*$ implies $\ker T^*_{\overline{\lambda}} = 0$. (Just because T^* is also a compact operator.)

5. $\Delta^{\perp}_{\lambda} = \ker T^*_{\overline{\lambda}}$, and because Δ_λ is a closed subspace by (3), $\Delta_\lambda \neq X \Rightarrow \ker T^*_{\overline{\lambda}} \neq 0$, and moreover we have $\Delta_\lambda = (\ker T^*_{\lambda})^{\perp}$).

   Proof: $(T_\lambda x, f) = (x, T^*_{\overline{\lambda}} f) = 0$ if $f \perp \Delta_\lambda$ for every x. Thus $T^*_{\overline{\lambda}} f = 0$ implies $\Delta^{\perp}_{\lambda} \hookrightarrow \ker T^*_{\overline{\lambda}}$.

   Now, if $f \in \ker T^*_{\overline{\lambda}}$ then $\langle T_\lambda x, f\rangle = 0$ for all x, hence $f \perp \mathrm{Im} T_\lambda$.

5^*. $\ker T_\lambda = (\Delta^*_{\overline{\lambda}})_{\perp}$ (as the above). From this and (3^*) it follows that

(5.2) $$(\ker T_\lambda)^{\perp} = \Delta^*_{\overline{\lambda}}\,.$$

6. $\ker T^*_{\overline{\lambda}} \neq 0 \Leftrightarrow \ker T_\lambda \neq 0$ (use (4),(5) for the one direction and (4*),(5*) for the other).

7. Also (4) may be inverted: $\ker T_\lambda = 0$ implies $\Delta_\lambda = X$ (by (6) and (5)). Thus we conclude that

(5.3) $$\sigma(T) = \{0; \sigma_p\}.$$

The statement (6) above is also called *First Theorem of Fredholm* It states that,

(5.4) $$\sigma_p(T) \setminus 0 = \overline{\sigma_p(T^*)} \setminus 0$$

An interpretation of (5)–(5*) is the *Second theorem of Fredholm*: when does the equation (λ and $y \in X$ are given)

(5.5) $$Tx - \lambda x = y \text{ have a solution } x \in X?$$

Answer: Consider the homogeneous part of the dual (adjoint) equation in X^*:

(5.6) $$T^* f - \overline{\lambda} f = 0.$$

Let $\{f_1, \ldots, f_k\}$ be the maximal set of linearly independent solutions (we know that it may be at most finite). Then, a solution x of (5.5) exists if and only if $f_i(y) = 0$, $i = 1, \ldots, k$. (i.e. y is "orthogonal" to all the solutions of the homogeneous dual equation). This is equivalent to (5).

Notes: (i) Think about T as an integral operator:

$$\int_1^b K(t,\tau)x(\tau)d\sigma - \lambda x(t) = y(t) \ . \tag{5.7}$$

Also, in the theory of Integral Equations usually instead of λ, we consider the characteristic numbers μ: $-\mu Tx + x = y$, i.e. $\mu = \frac{1}{\lambda}$, and $\mu_i = \frac{1}{\lambda_i}$ is a characteristic number if and only if λ_i is an eigenvalue. Then $\lambda_i \to 0$ corresponds to $\mu_i \to \infty$.

(ii) There is also a *Third theorem of Fredholm* which we will not study: $\dim\ker T_\lambda = \dim\ker T^*_{\overline{\lambda}}$.

It may happen that $\sigma(T) = \{0\}$ in which case there are no characteristic numbers at all.

Example: The *Volterra operator*. Let $X = C[0,1]$ and $K(t,\tau)$ be a continuous function. Consider

$$x - \int_0^t K(t,\tau)x(\tau)d\tau) = y(t) \ . \tag{5.8}$$

We will show that this equation has a solution for every $y \in C[0,1]$, which means that $1 \notin \sigma_p$ (but any other $\lambda \neq 0$ may be in σ_p). Let $\max_{t,\tau} |K(t,\tau)| \le C$ and $\max_t |y(t)| \le C_1$. Write $x_n = y + \int_0^t K(t,\tau)x_{n-1}(T)dT$ (and $x_0(\tau) = 0$). Then, by induction, we assume that $|x_n - x_{n-1}| \le C_1(Ct)^{n-1}/(n-1)!$ and we have

$$\begin{aligned} |x_{n+1} - x_n| &\le \int_0^t |K(t,\tau)| \cdot |x_n - x_{n-1}| d\tau \\ &\le C_1 C^n \int_0^t \frac{\tau^{n-1}}{(n-1)!} d\tau \\ &= C_1 C^n \frac{t^n}{n!}. \end{aligned} \tag{5.9}$$

Thus, x_n is a convergent sequence in $C[0,1]$. Indeed, $x_n = \sum_1^n (x_k - x_{k-1})$. Thus $|x_n(t)| \le C_1 e^{Ct}$ so $x_n \to x \in C[0,1]$ which is a solution of the equation.

5.3 Exercises

1. Find the spectrum of the operator D_w in ℓ_2 which is defined by $D_w x = (w_1 x_1, w_2 x_2, \ldots)$ for $x = (x_1, x_2, \ldots)$.

2. Find the spectrum of the operator A in $L_2[-1,1]$ which is given by $(Af)(t) = a(t)f(t)$, where

$$a(t) = \begin{cases} t, & \text{for } 0 \leq t \leq 1 \\ 0, & \text{for } -1 \leq t < 1. \end{cases}$$

3. Let $A \in L(X)$ be an invertible operator. Prove that $\sigma(A^{-1}) = \{\lambda^{-1} \mid \lambda \in \sigma(A)\}$.

4. Let $A \in L(X)$, $\lambda \in \mathbb{C}$, and assume that there exists a sequence $\{x_n\}_{n=1}^{\infty} \subseteq X$ such that $\|x_n\| = 1$ and $Ax_n - \lambda x_n \to 0$ as $n \to \infty$. Prove that $\lambda \in \sigma(A)$.

5. Let $C_{\mathbb{R}}$ be the space of functions $x(t)$ continuous and bounded on all of the line $(-\infty, \infty)$ with norm $\|x\| = \sup_{t \in \mathbb{R}} |x(t)|$. On the space $C_{\mathbb{R}}$ we define the operator A by $(Ax)(t) = x(t+s)$ where $s \in \mathbb{R}$ is a constant. Prove that $\sigma(A) = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$.

6. Find in $C[0, \pi/2]$ or $L_p[0, \pi/2]$ —for $1 \leq p < \infty$— the solution of the equation (which will depend on λ)

$$f(x) = \lambda \int_0^{\pi/2} \cos(x-y) f(y) dy.$$

7. Decide if there exists a solution in $L_p[0, \pi/2]$ for $1 < p < \infty$, of the equation

$$f(x) - \lambda \int_0^{\pi/2} \cos(x-y) f(y) dy = 1.$$

8. Let A be an invertible operator, and K be a compact operator in a Banach space. Prove that
   (a) $\dim\ker(A+K) < \infty$;
   (b) $\operatorname{codim}\operatorname{Im}(A+K) < \infty$.

9. For what $g \in C[0,\pi]$ the integral equation

$$f(x) - \int_0^\pi \sin(x+y) f(y) dy = g(x)$$

has a solution in the space $C[0,\pi]$?

10. For what $\lambda \in \mathbb{R}$ the equation

$$F(x) - \lambda \int_a^b e^{x-y} f(y) dy = 1$$

has a solution in the space $L_p[a,b]$, for $1 \leq p < \infty$?

11. Prove that for a compact operator T the following holds:

$$\text{codim}\Delta_\lambda = \text{dimKer}(T^* - \lambda I).$$

12. Let S_r and S_l be the right shift and left shift respectively on ℓ_2, i.e.

$$\begin{aligned} S_r(x_1, x_2, \ldots) &= (0, x_1, x_2, \ldots), \\ S_l(x_1, x_2, \ldots) &= (x_2, x_3, \ldots). \end{aligned}$$

Find the spectrum of these operators.

13. Define the operator $K : L_2[0,1] \mapsto L_2[0,1]$ by

$$(Kf)(t) = \int_0^1 k(t,s) f(s)\, ds$$

where

$$k(t,s) = \begin{cases} 1, & s \leq t \\ 0, & s > t. \end{cases}$$

Find the spectrum of K.

Chapter 6

Self adjoint compact operators

We call a bounded operator $A : H \mapsto H$ a *self-adjoint* or *symmetric* operator if and only if for every x, y in H we have $\langle Ax, y\rangle = \langle x, Ay\rangle$.

We start with a few general properties of such operators.

6.1 General Properties

We present here the main properties of the self-adjoint operators.

1. The spectrum σ_p of a self-adjoint operator A satisfies $\sigma_p \subset \mathbb{R}$. Indeed, let $\lambda \in \sigma_p$ and $Ax = \lambda x$ ($x \neq 0$). Then

   (6.1) $$\lambda\|x\|^2 = \langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\lambda}\|x\|^2 \Rightarrow \lambda = \overline{\lambda}.$$

2. If $\lambda_1 \neq \lambda_2$, $\lambda_1, \lambda_2 \in \sigma_p$ and $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$ then $x_1 \perp x_2$.

   Indeed, $\lambda_1 \langle x_1, x_2\rangle = \langle Ax_1, x_2\rangle = \langle x_1, Ax_2\rangle = \overline{\lambda}_2 \langle x_1, x_2\rangle$; but λ_i are distinct reals, thus $\langle x_1, x_2\rangle = 0$.

3. A subspace L of H is called *invariant with respect to* A if and only if $A(L) \subseteq L$. For a symmetric operator A, if L is invariant then $L^\perp$ is also an invariant subspace.

   Indeed, consider $y \in L^\perp$, $x \in L$. Then $Ax \in L$ and we get:

   (6.2) $$\langle Ax, y\rangle = 0 \Rightarrow \langle x, Ay\rangle = 0 \quad \forall\, x \in L,$$

   which means $Ay \in L^\perp$.

4. If A and B are symmetric and $AB = BA$ then AB is symmetric [exercise].

5. Define $C = \sup_{x \neq 0} \frac{|\langle Ax,x\rangle|}{\|x\|^2}$. Then if A is a symmetric operator we have $C = \|A\|$.

   Proof: Using the Cauchy-Schwartz inequality

$$| \langle Ax, x\rangle | \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2, \tag{6.3}$$

   hence $\frac{|\langle Ax,x\rangle|}{\|x\|^2} \leq \|A\|$ for all $x \neq 0$ and consequently $C \leq \|A\|$, which proves the easy inequality. Now we must show the reverse. Note first that

$$\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle = 2[\langle Ax, y\rangle + \langle Ay, x\rangle]. \tag{6.4}$$

   Using the triangle inequality we get

$$| \langle Ax, y\rangle + \langle Ay, x\rangle | \leq | \langle A(x+y), x+y\rangle | + | \langle A(x-y), x-y\rangle | \tag{6.5}$$

   It follows from the definition of C (and the Parallelogram Law) that

$$\begin{aligned} | \langle Ax, y\rangle + \langle y, Ax\rangle | &\leq \frac{1}{2} C(\|x+y\|^2 + \|x-y\|^2) \\ &= C(\|x\|^2 + \|y\|^2) \ . \end{aligned}$$

   Now let x be (any) vector with $\|x\| = 1$ and $y = \frac{Ax}{\|Ax\|}$ (the case $Ax = 0$ does not give the "sup" hence we may assume that $Ax \neq 0$). Then $\|y\| = 1$ and

$$\left| \frac{\langle Ax, Ax\rangle}{\|Ax\|} + \frac{\langle Ax, Ax\rangle}{\|Ax\|} \right| \leq 2C \Rightarrow \|Ax\| \leq C \tag{6.6}$$

   for all $x \in H$ with $\|x\| = 1$.

   This means $\|A\| \leq C$. □

6. $\langle Ax, x\rangle \in \mathbb{R}$ for any $x \in H$ if and only if A is symmetric.

   Proof: "$\Leftarrow$" is obvious;

   "$\Rightarrow$" we use a standard —but important— trick: we express a bilinear form through the "correct" combination of quadratic forms:

$$\begin{aligned} \langle A(x+y), x+y\rangle &- \langle A(x-y), x-y\rangle \\ &+ i\,\langle A(x+iy), x+iy\rangle - i\,\langle A(x-iy), x-iy\rangle \\ &= 4\,\langle Ax, y\rangle \end{aligned} \tag{6.7}$$

(use (6.4) to simplify the checking of this indentity).

Changing the positions of x and y and taking complex conjugate the left side will not change (this should be carefully checked), but the right side becomes $4\langle x, Ay\rangle$, which means that $\langle Ax, y\rangle = \langle x, Ay\rangle$. □

7. Let $\mu = \sup_{\|x\|=1}\{|\langle Ax, x\rangle|\}$. Then either μ or $-\mu \in \sigma(A)$.

   Proof: Take x_n, $\|x_n\| = 1$, and $|\langle Ax_n, x_n\rangle| \to \|A\|$ $(= \mu)$ (by 5.). Let $\langle Ax_n, x_n\rangle \to \lambda$ (it may be necessary to pass to a subsequence). Clearly $\lambda = \pm\mu$. Now $0 \le \|Ax_n - \lambda x_n\|^2 = \|Ax_n\|^2 - 2\lambda\langle Ax_n, x_n\rangle + \lambda^2\|x_n\|^2 \le 2\lambda^2 - 2\lambda\langle Ax_n, x_n\rangle \to 0$ as $n \to \infty$. Therefore, the inverse operator $(A - \lambda I)^{-1}$ cannot exist and be bounded which means $\lambda \in \sigma(A)$.

   Remark 6.1.1 *If A is in addition a compact operator, then there exists a subsequence Ax_{n_k} that converges, say to y_0. This implies $\{Ax_{n_k} - \lambda x_{n_k} \to 0\}$ that the limit $\lim x_{n_k} = x_0$ $(= y_0/\lambda)$ exists. So, x_0 is an eigenvector and λ is an eigenvalue. It also means that there exists a maximum $\max_{\|x\|=1}|\langle Ax, x\rangle|$ and it is achieved on an eigenvector. Also as a consequence, if a symmetric compact operator A is not identically zero then it has a non-zero eigenvalue $\lambda_0 \neq 0$.*

Theorem 6.1.2 (First Hilbert-Schmidt theorem) *For every compact symmetric operator $T : H \to H$, $T \neq 0$, there exists a set of eigenvalues $\{\lambda_n\}_{n\ge1} \subseteq \mathbb{R}$ such that $|\lambda_1| \ge \ldots \ge |\lambda_n| \ge |\lambda_{n+1}| \ge \ldots$ (converging to zero if this sequence is infinite) and an orthonormal system $\{e_n\}_{n\ge1}$ of eigenvectors: $Te_n = \lambda_n e$, such that*

1. $\forall\, x \in H$, $x = y + \sum_{i\ge1}\langle x, e_i\rangle\, e_i$ where $y \in \ker T$,

2. $Tx = \sum_{n\ge1}\langle Tx, e_n\rangle\, e_n$ and $\forall\, z \in \overline{\mathrm{Im}T} : z = \sum_{n\ge1}\langle z, e_n\rangle\, e_n$

Proof: We will build $\{e_n\}$ by induction. First define e_1 as in the remark above ($\max_{\|x\|=1}|\langle Ax, x\rangle| = |\langle Ae_1, e_1\rangle|$ and e_1 is an eigenvector of an eigenvalue λ_1, $|\lambda_1| = \|A\| = \max|(Ax, x)|$). Let $\{e_i\}_1^n$ be defined ($Te_i = \lambda_i e_i, |\lambda_1| \ge |\lambda_2| \ge \ldots \ge |\lambda_n|$). Let $E_n = \mathrm{span}\{e_i\}_1^n$. E_n is invariant subspace of T and therefore $E_n^\perp$ is an invariant subspace of T as well. Of course T is symmetric on every invariant subspace as well, and in particular on $E_n^\perp$. Therefore, $a_{n+1} = \sup\{|\langle Tx, x\rangle| \mid \|x\| = 1, \langle x, e_i\rangle = 0, i = 1, \ldots, n\} = |\lambda_{n+1}|$ ($= \|T|_{E_n^\perp}\|$). Therefoe if $a_{n+1} \neq 0$ by the same remark there exists $e_{n+1} \in E_n^\perp$, $Te_{n+1} = \lambda_{n+1}e_{n+1}$. Of course if $a_{n+1} = 0$ then $T|_{E_n^\perp} = 0$ and we stop our induction. Consider $x_n = \sum_1^n \langle x, e_i\rangle e_i$. Then $x^n = x - x_n \in E_n^\perp$. Thus, $\|Tx^n\| \leq a_{n+1}\|x^n\| \longrightarrow 0$ as $n \to \infty$, because we know that $|\lambda_{n+1}| = a_{n+1} \to 0$ ($n \to \infty$), and $\|x^n\| \leq \|x\|$; this is under the assumption that there is an infinite sequence of $\lambda_n \neq 0$; and if not, then, after a finite number of steps, $\lambda_{n+1} = 0$ meaning $T|_{E_n^\perp} \equiv 0$. Therefore $x - \sum_{n\geq 1}(x, e_i)Te_i \in \ker T$. This proves the first item.

To prove the second item apply the operator T to the equality in (1) and use that $Ty = 0$ and the symmetry of T:

$$\begin{aligned} Tx &= \sum_{n\geq 1} \langle x, e_i\rangle Te_i = \sum_{n\geq 1} \lambda_i \langle Px, e_i\rangle e_i = \sum_{i\geq 1} \langle x, \lambda_i e_i\rangle e_i \\ &= \sum \langle x, Te_i\rangle e_i \\ &= \sum \langle Tx, e_i\rangle e_i. \end{aligned}$$

Finally, $\{e_i\}_{i\geq 1}$ is an orthogonal basis in the $\overline{\mathrm{span}}\{e_i\}_{i\geq 1} = L$ and $\mathrm{Im}T \hookrightarrow L$ implies $\overline{\mathrm{Im}T} \hookrightarrow L$ (in fact $\overline{\mathrm{Im}T} = L$). Then we know from the general theory (theorem 2.1.10) that $\{e_i\}$ is a basis for $\overline{\mathrm{Im}T}$. □

Exercise: If T is symmetric then $\ker T \perp \mathrm{Im}T$.

Corollary 6.1.3 *Let H be a separable Hilbert space. Then there exists an orthonormal basis of eigenvectors $\{\eta_k\}_{i\geq 1}^\infty$.*

Proof: Indeed, $H = \ker T \oplus \overline{\mathrm{Im}T}$ and so only the case $\ker T \neq 0$ should be added to (2). Consider the orthonormal basis $\{e_i\}_{i\geq 1}$ of $\overline{\mathrm{Im}T}$ we built above. Add to it any orthonormal basis of $\ker T$, say $\{f_i\}$. Note that $Tf_i = 0 \cdot f_i = 0$ and so f_i is an eigenvector of eigenvalue $\lambda = 0$. □

What is the statement here?

Corollary 6.1.4 *Let $\langle Tx, x\rangle = \sum_{i\geq 1} \lambda_i |\langle x, e_i\rangle|^2$. We will separate in this expression the positive and negative eigenvalues, and we denote by e_i^+ the eigenvector corresponding to the positive eigenvalue λ_i^+. Similarly, for the negative eigenvalue and the corresponding eigenvector e_i^-.*

$$\sum \lambda_i^+ |\langle x, e_i^+\rangle|^2 + \sum \lambda_i^- |\langle x, e_i^-\rangle|^2 = \langle Tx, x\rangle . \tag{6.8}$$

(of course, if there is no, say, negative λ_i's, then the second sum does not exist). Let $\lambda_1^+ \geq \lambda_2^+ \geq \dots$ and $\lambda_1^- \leq \lambda_2^- \leq \dots$

Corollary 6.1.5 *Let H be an infinite dimensional Hilbert space. As a consequence of the last corollary we have:*

(6.9) $$\max_{\|x\|=1} \langle Tx, x\rangle = \lambda_1^+ \text{ and } \min_{\|x\|=1} \langle Tx, x\rangle = \lambda_1^-$$

(or $= 0$, if there is no negative λ_i's; similarly, if there is no positive λ_i's, $\max\langle Tx, x\rangle = 0$. We use here that $\lambda_i \to 0$ as $i \to \infty$ or becomes zero after some i). Since by Bessel inequality $\|x\|^2 \geq \sum_{i\geq 1} |\langle x, e_i\rangle|^2$ we get

(6.10) $$-\lambda_1^- |\|x\|^2 \leq \langle Tx, x\rangle \leq \lambda_1^+ \|x\|^2.$$

under the assumption that both positive and negative eigenvalues exist, or otherwise put zero on the corresponding side. So, $\langle Tx, x\rangle \geq 0$ for every $x \in H$ if and only if there are no negative eigenvalues.

Corollary 6.1.6 (Minimax principle) *(of Fisher, in the finite dimensional case; of Hilbert Courant, in the infinite dimensional Hilbert space.) Let $\lambda_{n+1}^+ > 0$ (meaning that there exist at least $(n+1)$ positive eigenvalues). Then*

(6.11) $$\lambda_{n+1}^+ = \min_{\{x_1,\dots,x_n\}} \varphi(x_1, \dots, x_n)$$

where

(6.12) $$\varphi(x_1, \dots, x_n) = \sup_{\|x\|=1} \{ \langle Tx, x\rangle \mid x \in (\operatorname{span}\{x_i\}_1^n)^\perp \}$$

(A similar formula can be written for λ_i^-.)

Proof: We know by corollary 6.1.5 that $\lambda_{n+1}^+ = \max\{(Tx, x), x \perp e_i^+, i = 1, \dots, n\}$. Thus, if we will prove that $\varphi(x_1, \dots, x_n) \geq \lambda_{n+1}^+$ (for any $(x_i)_1^n$), this will imply that $\min \varphi = \lambda_{n+1}^+$.

Let us show that there exists $y \perp (x_1, \dots, x_n)$, $\|y\| = 1$ and $y \in \operatorname{span}\{e_i^+\}_1^{n+1}$. Indeed, find $y = \sum_1^{n+1} a_i e_i^+$ such that $0 = (y, x_j) = \sum_1^{n+1} a_i (e_i^+, x_j)$. This is a system of n-equations with $(n+1)$ unknowns $\{a_i\}$. Hence, there is a non-zero solution y. We may normalize it so that $\sum |a_i|^2 = 1$. Thus we built such a y. Then

$$\begin{aligned} \varphi(x_1, \dots, x_n) &\geq (Ty, y) \\ &= \sum a_i \overline{a}_j (Te_i^+, e_j^+) \\ &= \sum_1^{n+1} \lambda_i^+ |a_i|^2 \geq \lambda_{n+1}^+ . \end{aligned}$$

□

Theorem 6.1.7 (Hilbert-Schmidt, on symmetric kernels)
Let $H = L_2[a,b]$. Consider the operator

$$Kx = \int_a^b K(s,t)x(t)dt, \tag{6.13}$$

where $K(s,t) \in L_2(I^2)$, $I = [a,b]$ and $K = K^*$ (meaning, $K(s,t) = \overline{K(t,s)}$). Let $\{e_i(t)\}$ be all orthonormal eigenvectors of K from the above theorem of Hilbert-Schmidt and $Ke_i = \lambda_i e_i$. Then:

$$K(s,t) = \sum_{i\geq 1} \lambda_i e_i(s)\overline{e_i(t)} \tag{6.14}$$

(the convergence of the series and the equality is understood in the sense of $L_2(I^2)$). As a consequence

$$\sum \lambda_i^2 = \int_a^b \int_a^b |K(s,t)|^2 ds\, dt. \tag{6.15}$$

Proof: Let $\eta_i(s,t) = e_i(s)\overline{e_i(t)}$. Then $\{\eta_i\}$ is an orthonormal system in $L_2(I^2)$. Hence there exists a function

$$\phi(s,t) = \sum_{i\geq 1} \langle K, \eta_i\rangle_{L_2(I^2)} \eta_i = \sum \lambda_i \eta_i \tag{6.16}$$

and $\|\phi\|_{L_2(I^2)} = \sqrt{\sum \lambda_i^2}$. Indeed,

$$\begin{aligned}
(K,\eta_i)_{L_2(I^2)} &= \int_a^b \int_a^b K(s,t)\overline{e_i(s)\overline{e_i(t)}}dsdt && (6.17)\\
&= \int_a^b \left(\int_a^b K(s,t)e_i(t)dt\right)\overline{e_i(s)}ds && (6.18)\\
&= \langle Ke_i, e_i\rangle_{L_2(I)} = \lambda_i \langle e_i, e_i\rangle = \lambda_i. && (6.19)
\end{aligned}$$

Take $x(t), z(t) \in L_2(I)$. By the first Hilbert-Schmidt theorem we know that $Kz = \sum(Kz, e_i)e_i$. Then on one hand:

$$\langle Kz, x\rangle = \int_a^b \int_a^b K(s,t)z(t)\overline{x(s)}dt\, ds = \left\langle K, x(s)\overline{z(t)}\right\rangle_{L^2(I^2)}, \tag{6.20}$$

and on the other hand,

$$\langle Kz, x\rangle = \left\langle \sum (Kz, e_i)e_i, x\right\rangle = \sum \lambda_i \langle z, e_i\rangle \langle e_i, x\rangle . \tag{6.21}$$

This can be written in the $L_2(I^2)$ scalar product as:

$$\sum \lambda_i \langle \eta_i, x\overline{z} \rangle_{L_2(I^2)} = \langle \phi, x\overline{z} \rangle_{L_2(I^2)} . \tag{6.22}$$

Consequently, for every function $x(s)$ and $z(t)$ in $L_2(I)$ we have

$$\langle \phi, x\overline{z} \rangle = \langle K, x\overline{z} \rangle \tag{6.23}$$

or equivalently

$$\langle \phi - K, x\overline{z} \rangle = 0. \tag{6.24}$$

Note that set $\{x(s)\overline{z(t)} \ : \ x \in L_2(I^2), \ z \in L_2(I^2)\}$ is complete in $L_2(I^2)$ [*why*?]. Therefore $\phi - K = 0$ (in the sense of $L_2(I^2)$). $\square$

Remark 6.1.8 *The integral operators defined by the kernel functions $K(t,s)$ from $L_2(I^2)$ lead to a very small subclass of compact symmetric operators. Their set of eigenvalues must tend to zero so quickly that $\sum \lambda_i^2 < \infty$. For example there is no such operator with eigenvalues $\lambda_n = 1/\sqrt{n}$.*

Remark 6.1.9 *Note that if $K(x,t)$ is a continuous function then it is easy to check (check it!) that the eigenvectors $e_i(t)$ are continuous.*

Theorem 6.1.10 (Mercer) *Let $K(s,t)$ be continuous and $\lambda_i \geq 0$ (no negative eigenvalues, which means $\langle Kx, x\rangle \geq 0 \ \forall\, x$.) Then $K(s,t) = \sum_{i\geq 1} \lambda_i e_i(s)\overline{e_i(t)}$ and this series converges absolutely and uniformly.*

We omit the proof.$\square$

Corollary 6.1.11 *Under the above conditions: $\sum \lambda_i = \int_a^b K(t,t)dt$ and $K(t,t) \geq 0$.*

Indeed, by Mercer's theorem $K(t,t) = \sum \lambda_i |e_i(t)|^2$. Recalling that

$$\int_a^b |e_i(t)|^2 dt = 1. \tag{6.25}$$

and integrating both sides we get: $\sum \lambda_i = \int_a^b K(t,t)dt$

6.2 Exercises

1. Let A be in $L(H)$, where H is a Hilbert space. Define on $H^{(2)} = H \oplus H$ the operator B by
$$B = \begin{pmatrix} 0 & iA \\ -iA^* & 0 \end{pmatrix}.$$
Prove that $\|A\| = \|B\|$ and that B is itself adjoint.

2. Let χ and ϕ be given vectors in a Hilbert space H. When does there exist a selfadjoint operator A on $L(H)$ such that $A\chi = \phi$? When is A of rank 1?

3. The operator $K : L_2[0,1] \mapsto L_2[0,1]$ is given by
$$(Kf)(t) = \int_0^1 k(t,s)f(s)\,ds,$$
where $k(t,s) = \min\{t,s\}$ for $0 \leq t,s \leq 1$.
   (a) Prove that K is a compact self adjoint operator.
   (b) Find the spectrum of K.
   (c) Find $\|K\|$.
   (d) Prove that K is positive and find the sum of it's eigenvalues.

4. The same as in the previous exercise for the case
$$k(t,s) = \begin{cases} 1-t, & 0 \leq s \leq t \leq 1 \\ 1-s, & 0 \leq t \leq s \leq 1. \end{cases}$$

5. The operator $K : L_2[0,1] \mapsto L_2[0,1]$ is given by
$$(Kf)(t) = \int_0^1 k(t,s)f(s)\,ds,$$
where $k(t,s) = \max\{t,s\}$ for $0 \leq t,s \leq 1$.
   (a) Prove that K is a compact self adjoint operator.
   (b) Find the spectrum of K.
   (c) Is K a positive operator?

Chapter 7

Self-adjoint bounded operators

7.1 Order in the space of symmetric operators

Definition 7.1.1 *An operator A is called non-negative (and we write $A \geq 0$) if and only if $(Ax, x) \geq 0$ for all $x \in H$. This of course implies that A is symmetric by the sixth propert of section 6.1. Also $A \leq B$ means that*

1. *both A and B are symmetric*

2. *$B - A \geq 0$*

7.1.1 Properties

1. $-I \leq A \leq I$ implies $\|A\| \leq 1$. Indeed, the inequalities mean that A is symmetric and $\sup_{\|x\| \leq 1} |\langle Ax, x \rangle| \leq 1$. Now use the fifth property of section 6.1).

2. $A \geq 0$ and $A \leq 0$ $A = 0$ (again by property 5 of section 6.1).

3. Let $A \geq 0$. Then $|\langle Ax, y \rangle|^2 \leq \sqrt{\langle Ax, x \rangle} \cdot \sqrt{\langle Ay, y \rangle}$ (this follows from the fact that $\langle x, y \rangle = \langle Ax, y \rangle$ is a quasi-inner product and by the Cauchy-Schwartz inequality for such products (see the exercise immediately after the theorem 2.1.1).

4. If C is symmetric and $A \leq B$ then $A + C \leq B + C$.

5. If A is symmetric then $A^{2n} \geq 0$. Indeed,

$$\left\langle A^{2n}x, x\right\rangle = \langle A^n x, A^n x\rangle = \|A^n x\|^2 \geq 0. \tag{7.1}$$

If $A \geq 0$ then also $A^{2n+1} \geq 0$ because

$$\left\langle A^{2n+1}x, x\right\rangle = \langle AA^n x, A^n x\rangle \geq 0. \tag{7.2}$$

It follows that for any polynomial $P(\lambda)$ with nonnegative coefficients $P(A) \geq 0$.

Theorem 7.1.2 (On the convergence of monotone sequences of operators) *Let $A_0 \leq A_1 \leq \ldots \leq A_n \ldots \leq A$. Then there exists a strong limit of $(A_n)_n$ (i.e. there exists a bounded operator B and $A_n x \to Bx$ for all $x \in H$).*

Proof: For every symmetric operator A there is a number C such that $A \leq C \cdot I$. So, changing the sequence to $0 \leq (A_n - A_0)/C_1 \leq I$ (where C_1 is such that $A - A_0 \leq C_1 \cdot I$) we can assume, without loss of generality that our original sequence already satisfies

$$0 \leq A_n \leq I. \tag{7.3}$$

For $n > m$ define $A_{mn} = A_n - A_m \geq 0$. Also $\|A_{mn}\| \leq 1$ since $0 \leq A_{mn} \leq I$. Then using the inequality of propert 3 above it follows that for any x and $y = A_{mn}x$ we have

$$\begin{aligned} \|A_m x - A_n x\|^2 &= |\langle A_{mn}x, A_{mn}x\rangle| \qquad (7.4)\\ &\leq |\langle A_{mn}x, x\rangle| \cdot |\left\langle A_{mn}^2 x, A_{mn}x\right\rangle| \\ &\leq |\langle A_{mn}x, x\rangle| \cdot \|x\|^2, \qquad (7.5)\end{aligned}$$

again because $\|A_{mn}x\| \leq \|x\|$ and $\|A_{mn}^2 x\| \leq \|x\|$. Thus,

$$\|A_m x - A_n x\|^2 \leq |\langle A_n x, x\rangle - \langle A_m x, x\rangle| \cdot \|x\|^2 \longrightarrow 0 \tag{7.6}$$

as $n > m \to \infty$ for all $x \in H$ because the sequence $\langle A_n x, x\rangle$ is monotone, increasing and bounded, and therefore it converges.

Hence, $\{A_n x\}$ is a Cauchy sequence and the limit $\lim A_n x =: Ax$ exists. Obviously Ax depends linearly on x. Also $0 \leq \langle A_n x, x\rangle \leq \langle x, x\rangle$ so it follows that $0 \leq \langle Ax, x\rangle \leq \|x\|^2$ which implies that A is a bounded operator.

Proposition 7.1.3 (The Main Proposition) *Let A be such that*

$$m \cdot I \leq A \leq M \cdot I \tag{7.7}$$

for some $m, M \in \mathbb{R}$ and let P be a polynomial satisfying $P(z) \geq 0$ for all $z \in [m, M]$. Then $P(A) \geq 0$.

The main point in the proof of this proposition is the following lemma:

Lemma 7.1.4 *If $A \geq 0$, $B \geq 0$ and $AB = BA$, then $AB \geq 0$.*

This is nontrivial; the symmetry of AB is trivial, but not the positiveness. This lemma follows immediately from the next one.

Lemma 7.1.5 *Let $A \geq 0$. Then there exists an operator X (and it is unique) such that $X^2 = A$ and $X \geq 0$. We write $\sqrt{A}$ for X. Moreover, $\forall\, B$ such that $AB = BA$ it is also true that $\sqrt{A}B = B\sqrt{A}$.*

Note that the lemma 7.1.5 implies the lemma 7.1.4. Indeed,

$$(\sqrt{A}\sqrt{A}\,Bx, x) = (B\sqrt{A}\,x, \sqrt{A}\,x) \geq 0. \tag{7.8}$$

Proof of Lemma 7.1.5. We want to find $X \geq 0$ such that $X^2 = A$. We may assume that $0 \leq A \leq I$. Let $B = I - A$ and $Y = I - X$. Then $A = I - B$, $X = I - Y$, $0 \leq B \leq I$ and the equation to be solved is $(I - Y)^2 = I - 2Y + Y^2 = I - B$, i.e. $Y = \frac{1}{2}(B + Y^2)$. We solve it by approximating its solution through the sequence Y_n:

$$Y_{n+1} = \frac{1}{2}(B + Y_n^2) \quad \text{and} \quad Y_0 = 0\ . \tag{7.9}$$

We can see by induction on n that:

(a) $Y_n \geq 0$ and Y_n is a polynomial with non-negative coefficients of B for all $n \in \mathbb{N}$ (this is straightforward).

(b) $Y_n \leq I$. Indeed, $Y_{n-1} \leq I \Rightarrow Y_n \leq I$ because $B \leq I$ and $Y_{n-1}^2 \leq I$. We must explain the last fact: it follows from the inequalities $0 \leq Y_{n-1} \leq I$ that $\|Y_{n-1}\| \leq 1$. Then $\left\langle Y_{n-1}^2 x, x\right\rangle = \langle Y_{n-1}x, Y_{n-1}x\rangle \leq \|x\|^2$.

(c) **Statement:** $Y_{n+1} - Y_n = \frac{1}{2}(Y_n^2 - Y_{n-1}^2) = \frac{1}{2}(Y_n + Y_{n-1})(Y_n - Y_{n-1})$ (we use here that $Y_nY_{n-1} = Y_{n-1}Y_n$ because they are ((by (a)) polynomials of the same operator B). Now, $Y_1 - Y_0 = \frac{1}{2}B$, and assuming by induction that $Y_n - Y_{n-1}$ is a polynomial of B with nonnegative coefficients we derive the same conclusion for $Y_{n+1} - Y_n$. So, by induction, $Y_{n+1} - Y_n$ is a polynomial of B with nonnegative coefficients and as a result

$$Y_{n+1} - Y_n \geq 0 \qquad \text{(for every } n = 0, 1, \ldots). \tag{7.10}$$

Hence, by theorem 7.1.2, $Y_n \to Y_\infty$ for some operator Y_∞ (strongly). Clearly, $Y_\infty = \frac{1}{2}(B + Y_\infty^2)$ meaning that $X = I - Y_\infty$ is $\sqrt{A}$. Also $0 \leq Y_\infty \leq I$ implies $X \geq 0$. X is a (strong) limit of polynomials of B (and also of A); therefore for all C such that $AC = CA$ we have

$$P(A)C = CP(A) \Rightarrow XC = CX. \tag{7.11}$$

Corollary 7.1.6 . *If $A \geq 0$ and $\langle Ax, x\rangle = 0$ then $Ax = 0$*

Proof: Indeed taking $X = \sqrt{A}$ we have $\langle X^2x, x\rangle = 0$ implies $\langle Xx, Xx\rangle = 0$ and this implies $Xx = 0$. Thus $Ax = 0$. □

We do not need the fact that the positive square root of A is unique in the proof of lemma 7.1.4. However as a usefull exercise let us show it. If $X_1 \geq 0$ and $X_1^2 = A$, then $X_1 = X$. Indeed:

(i) $X_1A = X_1^3 = AX_1 \Rightarrow XX_1 = X_1X$

(ii) If $y = (X - X_1)x \Rightarrow 0 = \langle (X + X_1)y, y\rangle = \underbrace{\langle Xy, y\rangle}_{\geq 0} + \underbrace{\langle X_1y, y\rangle}_{\geq 0} \Rightarrow Xy = 0$ and $X_1y = 0 \Rightarrow X^2 = XX_1 = A$

(iii) $\|(X - X_1)x\|^2 = \langle (X - X_1)^2x, x\rangle = 0 \Rightarrow X = X_1$.

Now we will prove the *Main Proposition* from lemma 7.1.4. $P(z) \geq 0$ for $z \in (m, M)$ implies

$$P(z) = c \prod_{\alpha_i \leq m} (z - \alpha_i) \prod_{\beta_i \geq M} (\beta_i - z) \cdot \prod [(z - \gamma_i)^2 + \delta_i^2]$$

for some $c > 0$. Obviously $A - \alpha_iI \geq 0$, $\beta_iI - A \geq 0$ and $(A - \gamma_iI)^2 + \delta_i^2I \geq 0$. Since all these operators are pairwise commutative, their products are also ≥ 0 by lemma 7.1.4. □

Corollary 7.1.7 *If $mI \leq A \leq MI$ and $P_1(t), P_2(t)$ are real polynomials and $P_1(t) \leq P_2(t)$ for all $t \in [m, M]$ then $P_1(A) \leq P_2(A)$.*

7.2 Projections (projection operators)

Let E be a linear space. A linear operator $P : E \to E$ is called a *projection* if and only if $P^2 = P$. Define $E_1 = \mathrm{Im}P$ and $E_2 = \ker P$.

7.2.1 Some properties of projections in linear spaces

1. $P|_{E_1} = \mathrm{Id}_{E_1}$ (i.e. $\forall x \in E$, $Px = x$).
   Indeed, for all $x \in E_1$ there exists $y \in E$ such that
   $$Py = x \Rightarrow P^2y = Px \Rightarrow x = Py = P^2y = Px \Rightarrow x = Px.$$

2. $I - P = Q$ is a projection ($Q^2 = Q$) and

   (7.12) $$\mathrm{Im}P = \ker(I-P)\ , \qquad \ker P = \mathrm{Im}(I-P).$$

   Indeed,
   (7.13) $P(I-P) = 0$ implies $\mathrm{Im}(I-P) \subset \ker P$;

   also if $x \in \ker P$ we get $(I-P)x = x$ and this means $\ker P \subset \mathrm{Im}(I-P)$.

3. Let $E_1 = \mathrm{Im}(P), E_2 = \ker(P)$ then $E_1 + E_2 = E$ and $E_1 \cap E_2 = 0$ (i.e. $E_1 \dot{+} E_2$ is a *direct* sum and E is a direct decomposition on E_1 and E_2). Indeed, $PE + (I-P)E = E$ and $\{Px = 0, (I-P)x = 0\}$ imply $x = 0$.

4. Let $T : E \to E$ be any linear operator, $E_1 \dot{+} E_2 = E$ and P be a projection onto E_1 parallel to E_2. Then $PT = TP$ if and only if E_1 and E_2 are invariant subspaces of T.

   Proof: $TE_1 = TPE_1 = PTE_1 \hookrightarrow E_1$. Thus $T : E_1 \to E_1$. Similarly for E_2: use instead of P operator $Q = I - P$. The other direction is left to the reader as an exercise. □

Now, let P be a projection in a Hilbert space H and assume that it is also a symmetric operator: $\langle Px, y\rangle = \langle x, Py\rangle$.
Then,

1. P is an orthoprojection ($E_1 \perp E_2$): for all x and y in H,

   (7.14) $$\langle Px, (I-P)y\rangle = \left\langle x, (P-P^2)y\right\rangle = 0$$

   i.e. $\mathrm{Im}P \perp \mathrm{Im}(I-P)$.

2. $0 \le P \le I$ since $\langle Px, x\rangle = \left\langle P^2 x, x\right\rangle = \|Px\|^2 \le \|x\|^2$ (from item 1), hence $0 \le P \le I$.

3. Let $E_i = \mathrm{Im} P_i$, i.e. $E_i = P_i H$. If $P_1 P_2 = 0$ then $P_2 P_1 = 0$, $E_1 \perp E_2$ and $P_1 + P_2$ is an orthoprojection onto $E_1 \oplus E_2$. Indeed, $\langle P_1 H, P_2 H\rangle = \langle P_2 P_1 H, H\rangle = 0$; $0 = (P_1 P_2)^* = P_2 P_1$; $(P_1 + P_2)^2 = P_1 + P_2$ and $\mathrm{Im}(P_1 + P_2) = E_1 + E_2$.

4. Let $P_1 P_2 = P_2 P_1 = P$. Then P is an orthoprojection (obvious) and $E = \mathrm{Im} P = E_1 \cap E_2$ $(E_i = P_i H)$.

   Indeed, obviously $E_1 \cap E_2 \hookrightarrow E$ (P_1 and P_2 are equal to $\mathrm{Id}_{E_1 \cap E_2}$ when restricted to $E_1 \cap E_2$). Also $Px \in E_1$ (because $P = P_1 P_2$) and $Px \in E_2$ (because $P = P_2 P_1$). So $E \hookrightarrow E_1 \cap E_2$.

5. If $P_1 P_2 = P_1$ then $E_1 \hookrightarrow E_2$ [$P_1 = P_1^* = P_2 P_1 = P_1 P_2$ then apply item 4 above] and $P_1 \le P_2 : (P_2 - P_1)^2 = P_2 - P_1$ and so ≥ 0.

   Moreover, $P_1 \le P_2$ implies $P_1 P_2 = P_1$. Indeed,

$$
\begin{aligned}
\|P_1(I - P_2)x\|^2 &= \left\langle P_1 \overbrace{(I - P_2)x}^{y}, \overbrace{(I - P_2)x}^{y} \right\rangle \\
&\le \langle P_2(I - P_2)x, (I - P_2)x\rangle \\
&= 0 \qquad (7.15)
\end{aligned}
$$

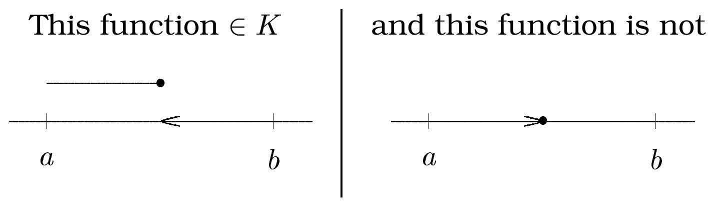
   which implies $P_1 - P_1 P_2 = 0$.

Chapter 8

Functions of operators

LET $mI \leq A \leq M \cdot I$; AND $a < m \leq M < b$. Let $K[a,b]$ be the set of piecewise continuous bounded functions and such that they are monotone decreasing limits ($\searrow$) of continuous functions. Examples:

(8.1)



(Such functions are semicontinuous from above (meaning that for all $t \in [a,b]$, $\overline{\lim}_{t_n \to t} x(t_n) = x(t)$) which is equivalent to saying that all sets $\{t : x(t) \geq a\}$ ($\forall\, a \in \mathbb{R}$) are closed sets.)

Lemma 8.0.1 *Let $\varphi(t) \in K[a,b]$. Then there exists a sequence of polynomials $P_n(t) \searrow \varphi(t)$ as $n \to \infty$ $\forall\, t \in [a,b]$.*

Proof: First, it is given that $\exists \varphi_n(t) \in C[a,b]$ such that $\varphi_n(t) \searrow \varphi(t)$. Also, by Weierstrass theorem $\forall\, n$ $\exists P_n(t)$-polynomial such that

$$|P_n(t) - [\varphi_n(t) + \frac{3}{2^{n+2}}]| \leq \frac{1}{2^{n+2}}. \tag{8.2}$$

Then $P_{n+1}(t) \leq \varphi_{n+1}(t) + \frac{1}{2^{n+1}} \leq \varphi_n(t) + \frac{1}{2^{n+1}} \leq P_n$. So $P_n(t)$ is non-increasing and obviously $P_n(t) \searrow \varphi(t)$ (because $\varphi_n(t)$ does). $\square$

The Lemma gives us the possibility to define for every $\varphi \in K$, an operator $\varphi(A)$:

Definition 8.0.2 (Defining $\varphi(A)$) *Let $P_n(t) \searrow \varphi(t)$ for $\forall\, t \in [a,b]$. Then $P_n(A) \geq P_{n+1}(A) \geq \dots$ and it is bounded (because $\varphi(t) \geq -T \Rightarrow P_n(A) \geq -T \cdot I$). So, by (vi) the strong limit of $\lim P_n(A)$ exists (call it B) (later it will be called $\varphi(A)$).*

We would like to call such a limiting operator B as $\varphi(A)$. In this case though, we must prove correctness (consistence) of such a definition. This means that B should depend only on $\varphi(t)$ and not on the specific sequence $p_n(t) \searrow \varphi(t)$. So, we should prove that if another sequence of polynomials $Q_n(t) \searrow \varphi(t)$ ($\forall\, t \in [a,b]$) then the strong limit of $Q_n(A)$ is the same B.

We prove a stronger statement needed below:

Lemma 8.0.3 *Let $Q_n(t) \searrow \psi(t) \in K$ ($\forall t \in [a,b]$) and $P_n(t) \searrow \varphi(t) \in K$. Let $\psi(t) \leq \varphi(t)$ $\forall\, t \in [a,b]$ Then $\lim_{n\to\infty} Q_n(A) = B_1 \leq B_2 = \lim_{n\to\infty} P_n(A)$.*

(So, if $\psi(t) = \varphi(t) \Rightarrow B_1 \leq B_2$ and $B_2 \leq B_1$, which implies $B_1 = B_2 = \varphi(a)$)

Proof: $\forall\, n, \forall\, t \in [a,b]$ $\exists N_0(t)$ such that for every $N \geq N_0(t)$

$$Q_N(t) < P_n(t) + \frac{1}{n}\ .(*) \tag{8.3}$$

This implies that $\exists$ open interval $I(t)$ around t where $(*)$ is also satisfied. So, we have a covering of $[a,b]$ by open intervals. Choose (by Heine-Borel Theorem) a finite subcovering $\{I(t_i)\}_{i=1}^{M}$. Then $\forall\, n$ $\exists N_0 = \max_{1\leq i\leq M} N_0(t_i)$ and for every $N > N_0$

$$Q_N(t) < P_n(t) + \frac{1}{n} \qquad \text{for every } t \in [a,b]\ . \tag{8.4}$$

Then letting $N \to \infty$ we get $B_1 \leq P_n(A) + \frac{1}{n} I$ (n is fixed here). Letting $n \to \infty$ and we have $B_1 \leq B_2$. □

So, we define a correspondence $\varphi \in K \mapsto \varphi(A) \in L(H)$

8.1 Properties of this correspondence ($\varphi_i \in K$)

(i) $\varphi_1 + \varphi_2 \mapsto \varphi_1(A) + \varphi_2(A)$ that is, $(\varphi_1 + \varphi_2)(A) = \varphi_1(A) + \varphi_2(A)$. [Indeed $P_n^{(i)} \searrow \varphi_i$ $i = 1, 2$. Then $P_n^{(1)} + P_n^{(2)} \searrow \varphi_1 + \varphi_2$ and the choice of polynomials tending decreasingly to $\varphi_1 + \varphi_2$ does not influence the limiting operator.]

(ii) For $c > 0$, $(c\varphi_1)(A) = c \cdot \varphi_1(A)$

(iii) $(\varphi_1 \cdot \varphi_2)(A) = \varphi_1(A) \cdot \varphi_2(A)$ right now this makes sense *only* for $\varphi_1 \geq 0$ and $\varphi_2 \geq 0$ because otherwise $\varphi_1 \cdot \varphi_2$ may not belong to the class K.]

(iv) $\varphi_1 \geq \varphi_2 \Rightarrow \varphi_1(A) \geq \varphi_2(A)$ (this was proved before).

We consider now the *linear class* of functions K–K of the form $\psi = f - \varphi$ where f, $\varphi \in K$. Then we write $\psi(A) = f(A) - \varphi(A)$ (by definition) and we trivially check that if $\psi = f_1 - \varphi_1 = f_2 - \varphi_2$, then $f_1 + \varphi_2 = \varphi_1 + f_2 \Rightarrow f_1(A) + \varphi_2(A) = \varphi_1(A) + f_2(A) \Rightarrow f_1(A) - \varphi_1(A) = f_2(A) - \varphi_2(A)$ and hence $\psi(A)$ is defined correctly. We may complete now the property (iii) above:

Let C_i be constants such that $\varphi_1 + C_1 \geq 0$ and $\varphi_2 + C_2 \geq 0$. Then define $\varphi_1 \cdot \varphi_2 = (\varphi_1 + C_1)(\varphi_2 + C_2) - C_1\varphi_2 - C_2\varphi_1 - C_1C_2$, and $(\varphi_1\varphi_2)(A)$ is defined through this identity and equals $\varphi_1(A) \cdot \varphi_2(A)$.

We are now ready to derive the *spectral decomposition* of a self-adjoint (=symmetric) bounded operator in H.

Consider the function

$$e_\lambda(t) = \begin{cases} 1 & \text{for } t \leq \lambda \\ 0 & \text{for } t > \lambda \end{cases} \tag{8.5}$$

$e_\lambda(t) \in K[a, b]$ and define $E_\lambda = e_\lambda(A)$. Then,

(i) $E_\lambda^2 = E_\lambda$ (because $e_\lambda(t) \cdot e_\lambda(t) = e_\lambda(t)$) and E_λ is symmetric (because $e_\lambda(t)$ is a real valued function, so $(E_\lambda x, x) \in \mathbb{R}$).

Thus E_λ *is an orthoprojection* and E_λ is symmetric.

Moreover $E_a = 0$ and $E_b = 1$ (because one compares it with the 0-function and with the identically 1 function).

(ii) E_λ is continuous (with respect to λ) from the right (in the strong sense): Indeed, let $P_n(t) \geq e_{\lambda+\frac{1}{n}}(t)$ and $P_n(t) \searrow e_\lambda(t)$ then $P_n(A) \geq E_{\lambda+\frac{1}{n}} \geq E_{\lambda+\alpha_n} \geq E_\lambda$ ($1/n \geq \alpha_n \geq 0$) and as $n \to \infty$, $P_n(A) \searrow E_\lambda$. So $E_{\lambda+\alpha_n} \underset{\alpha_n \to 0}{\longrightarrow} E_\lambda$].

(iii) $E_\lambda \cdot E_\mu = E_\lambda$ ($\lambda < \mu$), because $e_\lambda(t) \cdot e_\mu(t) = e_\lambda(t)$.

A family $\{E_\lambda\}$ with such properties is called "spectral family" or a "decomposition of identity".

(iv) $E_\lambda A = AE_\lambda$ (because E_λ is a limit of polynomials of A).

Therefore (by the property (iv) of linear projections) $\mathrm{Im}E_\lambda = H_\lambda$ is an invariant subspace of A.

8.2 The main inequality

Let $\lambda_1 < \lambda_2$. Then

$$\lambda_1[e_{\lambda_2}(t) - e_{\lambda_1}(t)] \le t \cdot [e_{\lambda_2}(t) - e_{\lambda_1}(t)] \le \lambda_2[e_{\lambda_2}(t) - e_{\lambda_1}(t)] \ . \tag{8.6}$$

inserting A in the place of t we get

$$\lambda_1(E_{\lambda_2} - E_{\lambda_1}) \le A(E_{\lambda_2} - E_{\lambda_1}) \le \lambda_2(E_{\lambda_2} - E_{\lambda_1}) \ (*) \tag{8.7}$$

Observe that $E_{\lambda_2\lambda_1} \equiv E_{\lambda_2} - E_{\lambda_1}$ is an orthoprojection. Let $H_{\lambda_2\lambda_1} = \mathrm{Im}E_{\lambda_2\lambda_1}$. It is an invariant subspace of A and (for $x \in H_{\lambda_2\lambda_1}$) we have $\lambda_1 I_{H_{\lambda_2\lambda_1}} \le A|_{H_{\lambda_2\lambda_1}} \le \lambda_2 I_{H_{\lambda_2\lambda_1}}$. Therefore, for $\forall\, \lambda \in [\lambda_1, \lambda_2]$:

$$\|A - \lambda I\|_{H_{\lambda_2\lambda_1}} \le \varepsilon = \lambda_2 - \lambda_1 \tag{8.8}$$

Thus, our operator is close to a constant operator on this subspace.

We are going to build now an *integral*.

Consider a partition of $(a, b) : a < \lambda_0 m \le \ldots \le M < \lambda_n < b$, with norm of partition $\Delta = \max|\lambda_{i+1} - \lambda_i| < \varepsilon$. Choose (any) $\mu_i \in [\lambda_i, \lambda_{i+1}]$.

Adding $(*)$ we have

$$\begin{aligned} \sum_0^{n-1} \lambda_k(E_{\lambda_{k+1}} - E_{\lambda_k}) &\le A\left(\sum_0^{n-1} E_{\lambda_{k+1}} - E_{\lambda_k}\right) \\ &\le \sum_0^{n-1} \lambda_{k+1}(E_{\lambda_{k+1}} - E_{\lambda_K}). \end{aligned} \tag{8.9}$$

Then

$$-\varepsilon I \le \sum_0^{n-1}(\lambda_k - \mu_k)(E_{\lambda_{k+1}} - E_{\lambda_k}) \tag{8.10}$$

$$\le A - \sum_0^{n-1} \mu_k(E_{\lambda_{k+1}} - E_{\lambda_k}) \tag{8.11}$$

$$\le \sum_0^{n-1}(\lambda_{k+1} - \mu_k)(E_{\lambda_{k+1}} - E_{\lambda_k}) \tag{8.12}$$

$$\le \varepsilon I \ , \tag{8.13}$$

since $-\varepsilon \le \lambda_k - \mu_k$ and $\lambda_{k+1} - \mu_k \le \varepsilon$. Consequently, by the property (i) of symmetric operators if $-\varepsilon I \le T \le \varepsilon I \Rightarrow \|T\| \le \varepsilon$ and we have

$$\left\| A - \sum_0^{n-1} \mu_k(E_{\lambda_{k+1}} - E_{\lambda_k}) \right\| \le \varepsilon \tag{8.14}$$

for *any* partition of the interval with the norm of partition $\leq \varepsilon$ and *any* choice of μ_k inside the intervals of the partition.

Then, there is a limit (in the *norm* of operators) when $\varepsilon \to 0$ and the natural name for this limit is "integral". So, we define a notion of *integral*:

$$A = \int_{m-0}^{M} \lambda dE_\lambda = \int_{-\infty}^{\infty} \lambda dE_\lambda \tag{8.15}$$

(meaning that we should take (any) $a < m$ as the low boundary, but, because of continuity from the right of E_λ, we may take the upper bound to be M; note that $E_\lambda \equiv I$ for $\lambda \geq M$ and $E_\lambda \equiv 0$ for $\lambda < m$).

Theorem 8.2.1 (Hilbert) *For every A self-adjoint bounded (or $\equiv$ symmetric bounded) operator H there exists a spectral decomposition ($\equiv$ spectral family) E_λ $(\lambda \in \mathbb{R})$ of orthoprojections, such that,*

(i) $E_\lambda = 0$ $(\lambda < m)$ *[we assume that $MI \leq A \leq M \cdot I$]* $= I$ $(\lambda \geq M)$

(ii) $E_{\lambda+0} = E_\lambda$ *(continuous from the right)*

(iii) $E_{\lambda_1} \leq E_{\lambda_2}$ *for* $\lambda_1 \leq \lambda_2$

(iv) $A = \int_{-\infty}^{\infty} \lambda dE_\lambda$

(v) *E_λ are strong limits of polynomials of A and therefore they commute with any operator B which commutes with A*

(vi) $\|Af\|^2 = \int \lambda^2 d(E_\lambda f, f)$

(vii) *a family $\{E_\lambda\}$ that satisfies (i)-(iv) is unique.*

Proof: We proved above (i)-(v). We will not prove (vii) (uniqueness) in this course.

But we will prove (vi): just return to the definition (description) of $\int \lambda dE_\lambda$ and observe that $\Delta E_{\lambda_i} \equiv E_{\lambda_{i+1}} - E_{\lambda_i}$ are pairwise orthogonal orthoprojections: $\Delta E_{\lambda_i} \perp \Delta E_{\lambda_j}$ for $i \neq j$ $(\Delta E_{\lambda_i} \cdot \Delta E_{\lambda_j} = 0)$. Moreover, $\|\Delta E_{\lambda_i} f\|^2 = (\Delta E_{\lambda_i} f, f) \equiv (E_{\lambda_{i+1}} f, f) - (E_{\lambda_i} f, f)$. Therefore, for any partition of (a, b) the Riemann–Stiltjies integral sum

$$\left| \sum \mu_i^2 (\Delta E_{\lambda_i} f, f) - \|Af\|^2 \right| < \varepsilon \tag{8.16}$$

which proves (vi). Note that $(E_\lambda f, f)$ is a monotone function of λ for any f and $\int \lambda^2 d(E_\lambda, f, f)$ can be understood as Riemann–Stiltjies integral.

Let us finish with one additional fact.

Proposition 8.2.2 (Fact) *If φ is a continuous function on $[m-0, M]$ then $\varphi(A) = \int_{m-0}^{M} \varphi(\lambda) dE_\lambda$ and the integral exists (converges) in the operator norm (i.e. "uniformly") and $\|\varphi(A)x\|^2 = \int \varphi^2(\lambda) d(E_\lambda x, x)$.*

Thus, we should prove two things: that the integral converges to some operator and that this operator was called before $\varphi(A)$.

Let $\varphi(t) = t^k$. Then $A^k \leftarrow (\sum \mu_i \Delta E_{\lambda_i})^k = \sum \mu_i^k \Delta E_{\lambda_i} \to \int \lambda^k dE_\lambda$.

Hence, $A^k = \int_{m-0}^{M} \lambda^k dE_\lambda$. From this follows that for any polynomial $P(\lambda)$ we have $P(A) = \int_{m-0}^{M} P(\lambda) dE_\lambda$.

Let $\varphi \in C[a,b]$. For a given $\varepsilon > 0$, find a polynomial $P(t)$ such that $|\varphi(t) - P(t)| < \min\{\varepsilon, \frac{\varepsilon}{b} - a\}$ $(\forall\, t \in [a,b])$ and

Then (i): $\|\varphi(A) - P(A)\| \leq \varepsilon$

Also for a suitable partition of $[a,b]$, define the corresponding Riemann integrable sums for $\int P(t)$ as $\sum(P)(\mu_i)\Delta_i$ and as $\sum \varphi(\mu_i)\Delta_i$ for $\int \varphi$ [we take φ in the same points μ_i as for $P(t)$]. Then, (ii) $|\sum(P)(\mu_i)\Delta_i - \sum(\varphi(\mu_i)\Delta_i| \leq \varepsilon \Rightarrow \|\sum(P(A)) - \sum'(\varphi(A))\| \leq \varepsilon$ [now it is integral sums for operators]. Also, because we proved the theorem for polynomials $P(A)$, we have

(iii) $\|P(A) - \sum(P(A))\| \leq \varepsilon$

All together we have (joining (i), (ii) and (iii)):

$$\left\|\varphi(A) - \sum (\varphi(A))\right\| \leq 3\varepsilon\ . \tag{8.17}$$

Consequently, the integral sums for $\varphi(A)$ converge in norm to an operator that was defined earlier as $\varphi(A)$.

Examples: (1) Let A be a compact operator; $Ax = \sum \lambda_k (x, e_k) e_k$. Then $E_\lambda x = \sum_{\lambda_k \leq \lambda} (x, e_k) e_k$ for $\lambda < 0$ and

$$E_\lambda x = x - \sum_{\lambda_k > \lambda} (x, e_k) e_k \quad \text{for} \quad \lambda > 0\ . \tag{8.18}$$

(2) $Ax(t) = t \cdot x(t) \Rightarrow \varphi(A)x(t) = \varphi(t) \cdot x(t)$ and $E_\lambda x(t) = e_\lambda(t) \cdot x(t)$.

We may define the operator $\varphi(A)$ for a larger class of functions: Let $\varphi(\lambda)$ be a measurable and bounded integrable function with respect to $\sigma(\lambda : x, y) = (E_\lambda x, y)$ for any $x, y \in H$. Then, by definition,

$$(\varphi(A)x, y) = \int \varphi(\lambda) d\sigma(\lambda; x, y)\ \ (**) \tag{8.19}$$

Note that $\sigma(\lambda; x, y)$ is of bounded variation (for every x, y). However, returning to $\varphi \in C[a,b]$ we see that, given A, x, y, we have (by

$(**)$) a linear continuous functional on $C[a,b]$ defined by $\sigma(\lambda; x, y)$. Then it is *known* that such functional defines a function $\sigma(\lambda; x, y)$ of bounded variation in the unique way under the normalization conditions: $\sigma(a) = 0$ and semicontinuous from the right.

Hence $(E_\lambda x, y)$ is uniquely defined for every x, y, which implies that $E_\lambda x$ is uniquely defined for every x. Thus $\{E_\lambda\}$ is uniquely defined by the conditions (i)-(iv) of the theorem of Hilbert.

8.3 Simple spectrum

We say that A has a *simple spectrum* if $\exists x_0 \in H$ called a generator such that $\{\Delta E_\lambda x_0 \equiv (E_{\lambda_2} - E_{\lambda_1})x \mid \forall\, \lambda_2 > \lambda_1\}$ is a complete set in H. It is easy to see that this is equivalent to the fact that $\{\varphi(A)x_0\}$ is a complete [set] for a family of all continuous functions on $[a, b]$.

We say that two opererators A_1 and A_2 are *unitary equivalent* if $\exists U$-unitary and $A_1 = U^{-1}A_2U$. In fact, we use this notion also in the case $A_i : H_i \to H_i$, $i = 1, 2$ and $U : H_1 \to H_2$ being an isometry *onto*.

Theorem 8.3.1 *Let A be a self-adjoint (bounded) operator with a simple spectrum and generator x_0. Let $\sigma(\lambda) = (E_\lambda x_0, x_0)$ where $\{E_\lambda\}$ is a spectral family of orthoprojectors defined by A. Then A is unitary equivalent to operator $T : L^2_{\sigma(\lambda)} \to L^2_{\sigma(\lambda)}$ where $T\varphi(\lambda) = \lambda \cdot \varphi(\lambda)$.*

Proof: Consider first any continuous function $\varphi(\lambda) \in C[a, b]$ (where $\mathrm{supp}\sigma(\lambda) \subset [a, b]$). Note that a set of all such functions is dense in $L^2_{\sigma(\lambda)}$ (in fact, it follows from the definition of $L^2_{\sigma(\lambda)}$).

Consider the map $U : \varphi \to y_\varphi = \varphi(A)x_0 = \int \varphi(\lambda) dE_\lambda x_0$. Then

$$\|y_\varphi\|^2 = \int_a^b \varphi^2 d(E_\lambda x_0, x_0) = \|\varphi\|^2_{L_2\sigma(\lambda)} \tag{8.20}$$

It is easy to check that the condition of simplicity of spectrum implies that $\{y_k\}$ is a dense set of H.

Thus, the linear map U is extended from a dense set $\{\varphi \mid$ continuous functions$\}$ to the completion $L^2_{\sigma(\lambda)}$ and we built an isometry $U : L^2_{\sigma(\lambda)} \to H$ (onto, because the image of an isometry is a complete space). It remains to check $UAU^{-1}\varphi = \lambda\varphi(\lambda)$:

$$A\varphi(A)x_0 = \int \lambda\varphi(\lambda) dE_\lambda x_0\ . \tag{8.21}$$

So $Ay_\varphi \mapsto \lambda\varphi(\lambda)$.

Chapter 9

Spectral theory of unitary operators

$U : H \to H$ is unitary if $(Ux, Uy) = (x, y)\ \forall\, x, y \in H$ and $\mathrm{Im} U = H$ (otherwise, we talk about *isometry*).

Properties 1. $UU^* = I = U^*U$ (unitary) 2. Linearity of U is a consequence of $(U_x, U_y) = (x, y)\ \forall\, x, y \in H$. 3. If U is linear and $(U_x, U_x) = (x, x)\ \forall\, x$ then U is unitary.

Example $L_2(-\infty, \infty)$:

(9.1) $$\tilde{f}(\tau) = \frac{1}{\sqrt{2\pi}} \lim_{N\to\infty} \int_{-N}^{N} f(t) e^{it\tau} dt \equiv Ff \quad \text{(Unitary)}$$

$F^{-1}\tilde{f} = f$ [substitute i with $-i$].
[regularization: $Ff = \frac{1}{\sqrt{2\pi}} \frac{d}{d\tau} \int_{-\infty}^{\infty} \frac{e^{-it\tau}-1}{-it} f(t) dt$].

9.1 Spectral properties

(1) $Ux = \lambda x \Rightarrow |\lambda| = 1$;

(2) $Ux_i - \lambda_i x_i$ and $\lambda_1 \neq \lambda_2$ then $(x_1, x_2) = 0$ [Indeed: $(x_1, x_2) = (Ux_1, Ux_2) = \lambda_1 \overline{\lambda_2} (x_1, x_2) \Rightarrow (x_1, x_2) = 0$].

We say that a subspace $E \hookrightarrow H$ reduces A iff

(9.2) $$A : E \to E \quad \text{and} \quad A : E^\perp \to E^\perp$$

(that is, both E and $E^\perp$ are invariant subspaces).

(3) If $U : E \to E$ and $U^{-1} : E \to E$ then $U : E^\perp \to E^\perp$. (Just use $(Ux, y) = (x, U^{-1}Y)$.)

Consider now a polynomial $P(z)$ of z and z^{-1} for $|z| = 1$ (i.e. for $z = e^{it}$, $t \in \mathbb{R}$):

$$P(z) = \sum_{-m}^{n} a_k e^{ikt} = \sum a_k z^k \Rightarrow P(U) = \sum a_k U^k \ . \tag{9.3}$$

Properties of this correspondence:

(i) $(P_1 + P_2)(U) = P_1(U) + P_2(U)$ (linearity)

(ii) $(P_1 \cdot P_2)(U) = P_1(U) \cdot P_2(U)$ (multiplicativity)

(iii) $P(U)^* = \sum \overline{a}_k U^{-k} = \overline{P}(U^{-1}) = \overline{P(z)}|_{z=U}$.

(iv) If $P(z) \geq 0$ for $|z| = 1$, then $P(U) \geq 0$.

Proof: We start with a lemma.

Lemma 9.1.1 *Let $P(z) \geq 0$, $|z| = 1$; then there is another polynomial $Q(z)$ so that $P(z) = Q(z) \cdot \overline{Q(z)} = |Q(z)|^2$. (Note $\overline{Q(z)} = \overline{Q}(\overline{z}) = \overline{Q}(1/z)$.)*

Proof: Let $P(z) > 0$ (we may consider $P(z) + \varepsilon > 0$ for the original $P(z)$ to create a strict inequality and then $\varepsilon \to 0$). Then $z^m P(z) = c \prod_{|\alpha_i|<1}(z - \alpha_i) \prod_{|\beta_j|>1}(z - \beta_i)$. Next remember that $|z| = 1$. $\overline{P(z)} = \overline{z^{-m}} \overline{c} \prod (\frac{1}{z} - \overline{\alpha}_i) \prod (\frac{1}{z} - \overline{\beta}_j) \equiv P(z)$ (because it is real for $|z| = 1$). Note that $\overline{z^{-m}} = z^m$ and the number of all roots is $n + m$.

So $\overline{P(z)} = \frac{z^m}{z^{n+m}} c_1 \prod (z - \frac{1}{\alpha_i}) \prod (z - \frac{1}{\beta_j}) = P(z)$ (for some c_1). Since it is still the same polynomial with the same roots we see that there is correspondence $\beta_i = \frac{1}{\alpha_i}$ ($\forall\, i = 1, \ldots, k$), $m + n = 2k$ and $n = m = k$. Define $Q(z) = \prod_{|\alpha_i|<1}(z - \alpha_i)$. Then $Q(z) \cdot \overline{Q(z)} = \frac{c_2}{z^n} \prod_{|\alpha_i|<1} \prod_{|\beta_i|>1} = c_3 P(z) > 0$ meaning $c_3 > 0$. □

Returning to prove we write (iv), $P(U) = Q(U)\overline{Q}(U^{-1}) = Q(U) \cdot Q(U)^* \geq 0$.

We continue as in the case of self-adjoint operators. Let $\varphi(z) = \varphi(e^{-it}) \geq 0$ (remember: $|z| = 1$) and $P_n(z) \searrow \varphi(e^{it})$ (for every t) ($P_n(z)$ are trigonometric polynomials). Then we define $\varphi(u) \geq 0$ as the strong limit of $P_n(U)$. We extend the definition of $\varphi(u)$ for functions $\varphi \in K_1 = \{c_1\varphi_1 + c_2\varphi_2\}$ for any complex numbers x_1 and c_2. We check, of course, consistence of our definitions considering the unique decomposition $\psi(z) = \mathrm{Re}\psi(z) + i\mathrm{Im}\psi(z)$ [$\equiv \psi_1 + i\psi_2$] and then $\psi(U) = \psi_1(U) + i\psi_2(U)$.

All the properties of the correspondence $\psi(e^{it}) \mapsto \psi(U)$ can be checked as in the case of self-adjoint operators. The one which should be checked in addition is:

$$\varphi(U)^* = \overline{\varphi(z)}|_{z=U} \tag{9.4}$$

(because it is true for approximating polynomials).

To build a "spectral family of projections" for a unitary operator we consider the functions

(9.5) $$\psi_\lambda(e^{it}) = \begin{cases} 1 & \text{for } 0 < t \le \lambda \\ 0 & \text{for } \lambda < t \le 2\pi \end{cases}$$

and $\psi_0(e^{it}) \equiv 0$, $\psi_{2\pi} \equiv 1$. Let also

(9.6) $$\tilde{\psi}_0(t) = \begin{cases} 1 & \text{for } t = 0 \\ 0 & \text{for } 0 < t < 2\pi \end{cases}$$

Then $\psi_\lambda(U) = E_\lambda$ are orthoprojections. To prove continuity from the right of this family, introduce $\tilde{E}_0 = \tilde{p}si_0(U)$ and prove continuity from the right of the family $\{E_\lambda + \tilde{E}_0\}$. Then return to $\{E_\lambda\}$. We build spectral integral. First functional inequality is

(9.7) $$e^{it} - \sum_1^n e^{i\varphi_j}[\psi_{t_j}(e^{it}) - \psi_{t_{i-1}}(e^{it})] = \chi(e^{it}) \ ,$$

(9.8) $|\chi(e^{it})| \le |e^{it} - e^{i\varphi_k}| \le |t - \varphi_k| \le \varepsilon$ (where $t_{k-1} \le t \le t_k$).

then $\overline{\chi(e^{it})} \cdot \chi(e^{it}) \le \varepsilon^2$ which implies

(9.9) $$\left\| U - \sum_{j=1}^n e^{i\varphi_j}[E_{t_j} - E_{j-1}] \right\| \le \varepsilon \ .$$

Thus,

(9.10) $$U = \int_0^{2\pi} e^{it} dE_t \quad \text{and} \quad U^n = \int_0^{2\pi} e^{int} dE_t,$$

for $n = 0, \pm 1, \pm 2, \dots$ (For negative powers take the dual operators $U^* = U^{-1} = \int_0^{2\pi} e^{-it} dE_t$). Similarly, for continuous functions $\varphi(e^{it})$

(9.11) $$\varphi(U) = \int_0^{2\pi} \varphi(e^{it}) dE_t \ .$$

Chapter 10

The Fundamental Theorems.

WE STUDY THE following structure: a linear space L over the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$. Usually we assume that the dimension of L is infinite. We endorse L with a norm $\|\cdot\|$ and set $X = (L; \|\cdot\|)$. We assume that X is a complete linear space with norm and we call such a space a *Banach space*.

A linear functional $f : X \mapsto \mathbb{R}$ (or $\mathbb{C}$ if X is over $\mathbb{C}$) is called a "linear functional" or just a "functional". The set of all linear functionals is a linear space denoted with $X^{\#}$ (or $L^{\#}$ emphasizing that only the linear structure is involved). A subspace of $X^{\#}$ consisting of bounded (i.e. continuous) functionals is called the dual space X^* and it is a normed space under the norm

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}. \tag{10.1}$$

The dual space of any normed space X is a Banach space.

Let L be the set (linear space) of all bounded linear maps (operators) from X to Y. Again, this is a Banach space under the norm

$$\|A : X \mapsto Y\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \tag{10.2}$$

if Y is complete.

We will deal now with the three most fundamental theorems of Functional Analysis. Some of them are based on the notion of *category* (Baire category).

10.1 The open mapping theorem

Let M be a complete metric space (not necessarily linear). We call a subset $A \subseteq M$ to be of first category if A is a union of countably many sets $A = \cup_i B_i$ and the B_i's are nowhere dense (meaning that the closure $\overline{B}_i$ does not contain any interior points) for every i. A set C which is not of first category is called a set of second category. Note that if a closed set A is of second category then $A^\circ \neq \emptyset$.

Theorem 10.1.1 (Baire-Hausdorff) *Every complete metric space M is a set of second category.*

Proof: Assume $M = \cup_1^\infty A_n$ and that every A_n is nowhere dense. Then there exists $x_1 \in M \setminus \overline{A}_1$ meaning that there exists $\varepsilon_1 > 0$ and a ball $B_1 = \mathcal{D}(x_1; \varepsilon_1)$ of radius ε_1 and center x_1 so that $B_1 \subseteq (\overline{A}_1)^{\mathrm{c}}$. Similarly since A_2 is nowhere dense there exists $x_2 \in (B_1 \setminus A_2)^\circ$, that is, there exists $0 < \varepsilon_2 < \varepsilon_1$ and a ball $B_2 = \mathcal{D}(x_2, \varepsilon_2) \subseteq B_1 \setminus \overline{A}_2$. We continue in this manner thus producing a sequence $\{x_n\}_{n=1}^\infty$ and $\{0 < \varepsilon_n \searrow 0\}$ such that

$$\mathcal{D}(x_n, \varepsilon_n) \subseteq \mathcal{D}(x_{n-1}, \varepsilon_{n-1}) \setminus \overline{A}_n. \tag{10.3}$$

Then, by the completence of the space M the limit $x_0 = \lim x_n$ exists and $x_0 \notin \overline{A}_j$ for every $j = 1, 2, \ldots$. Thus $M \neq \cup A_n$, a contradiction. □

Definition 10.1.2 *A set $K \subseteq X$ is called* perfectly convex *if and only if for every bounded sequence $x_i \in K$ and for every sequence of reals $\alpha_i \geq 0$ such that $\sum_{i=1}^\infty \alpha_i = 1$ we have that $\sum_{i=1}^\infty \alpha_i x_i \in K$.*

We also define K° to be the interior of K, $\overline{K}$ the closure of K and $\overset{\mathrm{c}}{K} = \{x \in K \mid \forall y \in X\ \exists \alpha > 0 \text{ and } \lambda y + (1-\lambda)x \in K \text{ for } 0 \leq \lambda \leq \alpha$ is the *kernel* of K (sometimes called the "center" of K).

Note that for any $x \in X$ if we put $K_x = K - x$ we have,

$$(K_x)^\circ = K^\circ - x, \qquad (\overline{K_x})^\circ = (\overline{K})^\circ - x, \tag{10.4}$$

$$\overset{\mathrm{c}}{K_x} = \overset{\mathrm{c}}{K} - x, \qquad \overset{\mathrm{c}}{\overline{K_x}} = \overset{\mathrm{c}}{\overline{K}} - x. \tag{10.5}$$

so $K - x$ is perfectly convex if and only if K is perfectly convex.

Theorem 10.1.3 (Livshič) *If K is perfectly convex in a Banach space X then*

$$K^\circ = \overset{\mathrm{c}}{K} = \overset{\mathrm{c}}{\overline{K}} = (\overline{K})^\circ. \tag{10.6}$$

Proof: We will show that

$$(\overline{K})^\circ \subseteq K^\circ \subseteq \overset{c}{K} \subseteq \overset{c}{\overline{K}} \subseteq (\overline{K})^\circ. \tag{10.7}$$

Of course the middle two relations are trivial. We will show the first and the last one. For the first one it is enough to prove that $0 \in (\overline{K})^\circ$ implies $0 \in K^\circ$. Let $\mathcal{D}$ be such an ε-ball that $\mathcal{D} \subseteq \overline{K}$. This implies

$$\mathcal{D} \subseteq \overline{K \cap \mathcal{D}} \subseteq K \cap \mathcal{D} + \frac{1}{2}\mathcal{D}. \tag{10.8}$$

Then, for every $\alpha > 0$

$$\alpha\mathcal{D} \subseteq \alpha(K \cap \mathcal{D}) + \frac{\alpha}{2}\mathcal{D}. \tag{10.9}$$

It follows that for *any* $y \in \frac{1}{2}\mathcal{D}$ we may write $y = \frac{1}{2}x_1 + y_1$ where $x_1 \in K \cap \mathcal{D}$, $y_1 \in \frac{1}{4}\mathcal{D}$ and again using (10.9) $y_1 = \frac{1}{4}x_2 + y_2$ with $x_2 \in K \cap \mathcal{D}$ and $y_2 \in \frac{1}{8}\mathcal{D}$, and so on. Therefore, $y = \sum_{n=1}^\infty \frac{1}{2^n}x_n \in K$ (since K is perfectly convex). Thus, we proved the first inclusion.

To prove the last one, we deal again only with the case $0 \in \overset{c}{\overline{K}}$. It follows from the definition of the kernel of K that

$$X = \bigcup_{n=1}^\infty n\Big[\overline{K} \cap (-\overline{K})\Big]. \tag{10.10}$$

By Baire's Theorem X is of second category meaning that one of the sets in the above union is a set of second category; so, $\overline{K} \cap (-\overline{K})$ is of second category and consequently it has an interior point, say x_0; i.e. for some $\varepsilon > 0$

$$\mathcal{D}(x_0; \varepsilon) \subseteq \Big[\overline{K} \cap (-\overline{K})\Big], \tag{10.11}$$

where $\mathcal{D}(x_0; \varepsilon)$ is the ε-ball centered at x_0. It follows that

$$\frac{1}{2}\mathcal{D}(x_0; \varepsilon) - \frac{1}{2}\mathcal{D}(x_0; \varepsilon) \subseteq \overline{K}, \tag{10.12}$$

meaning that $0 \in (\overline{K})^\circ$. □

Corollary 10.1.4 (Open Mapping Theorem) *Let X and Y be Banach spaces and $A : X \mapsto Y$ be a bounded linear operator* onto Y. *Then A is an open map meaning that for every open set $\mathcal{O} \subseteq X$ its image $A(\mathcal{O})$ is an open subset of Y.*

Proof: If $\mathcal{D}_x$ is an open ball in X then $A(\mathcal{D}_x)$ is perfectly convex in Y (we use here that our map is onto). Obviously $0 \in (A\overset{c}{\mathcal{D}})$implies by the previous theorem that $0 \in (A\mathcal{D})^\circ$ and this proves the corollary (again we proved only that $0 \in (A\overset{c}{\mathcal{D}})$but similarly we prove that for all $x \in \overset{c}{D}$we have $Ax \in (A\overset{c}{\mathcal{D}})$. □

Important partial case: Let $A \in L(Y \mapsto X)$ be onto and one-to-one. Then there is $A^{-1} \in L(Y \mapsto X)$ and A is an isomorphism between these two spaces. In other words, A can be considered the identity map (from the linear point of view) between X and Y (by just renaming in Y Ax as x). This implies that if a constant C exists such that $\|x\|_Y \leq C\|x\|_X$ for all $x \in X$ and the linear space is complete with respect to both norms it follows that there exists an other constant C_1 such that $\|x\|_X \leq C_1\|x\|_Y$ for all $x \in X$.

10.2 The Closed Graph Theorem

Let $A : X \mapsto Y$ be a linear operator. The set

$$\Gamma(A) = \{(x; Ax)\}_{x \in \mathrm{Dom}A} \subseteq X \times Y \tag{10.13}$$

is called the *graph* of A. We say that A is a closed graph operator if $\Gamma(A)$ is a closed set in $X \times Y$; this means that whenever $x_n \in \mathrm{Dom}A$, $x_n \to x$ and $Ax_n \to y$ then $x \in \mathrm{Dom}A$ and $Ax = y$.

Theorem 10.2.1 (Banach) *Let $A : X \mapsto Y$ be a closed graph operator and Dom$A = X$. Then A is a continuous (i.e. bounded) operator.*

Proof:: $A^{-1}\mathcal{D}_Y$ is perfectly convex [indeed: take any bounded set $\{x_i\} \subseteq A^{-1}\mathcal{D}_Y$ i.e. there exists $y_i \in \mathcal{D}$ and $Ax_i = y_i$; let $\alpha_i \geq 0$ and $\sum \alpha_i = 1$; then $\sum_1^n \alpha_i y_i \to y \in \mathcal{D}$ and $\sum_1^n \alpha_i x_i \to x$; by the closed graph condition $Ax = y$ meaning $\sum \alpha_i x_i \in A^{-1}\mathcal{D}$]. Clearly $0 \in A^{-1}\overset{c}{\mathcal{D}}_Y$ implying the continuity of A. □

Examples: We say that an operator $A : X \mapsto Y$ admits a closure if and only if $\overline{\Gamma A} \subseteq X \times Y$ is the graph of a closed operator.

(a) $A : L_2[0,1] \mapsto L_2[0,1]$ and $Ax = x(0) \cdot t$ and Dom$A = C[0,1]$. Clearly this operator does not admit a closure.

(b) $Ax = \frac{d}{dt}x$ in L_2 (or C) with Dom$A = \{x \in L_2 \mid x' \in L_2\}$ (or Dom$A = \{x \in C \mid x' \in C\}$). This operator admits a closure and to understand this the following easy fact is useful:

Fact: If A has its graph closed and the inverse A^{-1} exists, then A^{-1} has also a closed graph (since the graph of A and A^{-1} are the same). Thus if $A : X \mapsto Y$ is a compact operator and $\mathrm{Ker}A = 0$ then A^{-1} is formally defined on $\mathrm{Im}A (= \mathrm{Dom}A^{-1})$ and has a closed graph.

Remark: It is also possible to prove Closed Graph Theorem from Theorem from the Open Mapping Theorem with the following direct argument:

Consider the subspace $E = \{(x; y) \mid y = Ax\} \hookrightarrow X \times Y$. If A has its graph closed then E is a closed subspace therefore complete. Define $u : E \mapsto X$ by $u(x; y) = x$. This operator is onto (because $\mathrm{Dom}A = X$) and one-to-one. Then u^{-1} is bounded. Moreover the operator $T : E \mapsto Y$ defined by $T(x; y) = y$ is continuous. Thus $A = Tu^{-1}$ is a continuous operator. □

We give next an application of the Closed Graph Theorem:

Theorem 10.2.2 (Hörmander) *Let X_0, X_1, X_2 be Banach spaces and $T_1 : X_0 \mapsto X_1$, $T_2 : X_0 \mapsto X_2$, $\mathrm{Dom}T_1 \subseteq \mathrm{Dom}T_2$, T_1 is a closed operator and T_2 admits a closure. Then there exists $C > 0$ such that $\|T_2x\| \leq C[\|T_1x\| + \|x\|]$ for all $x \in \mathrm{Dom}(T_1)$.*

Proof: Consider the closed subspace E of $X_0 \times X_1$ with

$$E = \{(x; T_1x) \mid x \in \mathrm{Dom}T_1\} \tag{10.14}$$

and set $V(x; T_1x) = T_2x$. Then V is a closed operator [indeed: $x_n \to x$, $T_1x_n \to y$ and by the closeness of T_1 it follows that $y = T_1x$ and also $T_2x_n \to z$ implies $z = T_2x$]. Also $\mathrm{Dom}V = E$. This implies that V is continuous, i.e., there exists constant C such that $\|T_2x\| \leq C\|(x; T_1x)\|$. □

10.3 The Banach-Steinhaus Theorem

We start with a late version which belongs to Zabreiko:

Theorem 10.3.1 (Zabreiko) *Let μ be a function on a Banach space X, $\mu(x) \geq 0$, $\mu(tx) \to 0$ as $0 < t \to 0$ for every $x \in X$ (continuity in every direction). Assume that μ is perfectly convex, that is, if the series $\sum_1^\infty x_i$ converges then*

$$\mu\left(\sum_1^\infty x_n\right) \leq \sum_1^\infty \mu(x_n). \tag{10.15}$$

Then μ is a continuous function.

Proof: Consider, for every $\varepsilon > 0$, the set $M_\varepsilon = \{x|\ \mu(x) \le \varepsilon\}$. Then $0 \in \overset{c}{M}_\varepsilon$ (because of the continuity in each direction) and

$$\mu\left(\sum_1^\infty \alpha_i x_i\right) \le \sum_1^\infty \alpha_i \mu(x_i) \le \varepsilon \tag{10.16}$$

for every $\alpha_i \ge 0$, $\sum_1^\infty \alpha_i = 1$ and $x_i \in M_\varepsilon$. Thus, M_ε is perfectly convex. Therefore, by the above theorem $0 \in \overset{o}{M}_\varepsilon$ meaning that there exists $\delta > 0$ such that the ball $\mathcal{D}_\delta$ of radius δ and center at zero. In other words, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|x\| \le \delta$ implies $\mu(x) \le \varepsilon$. So μ is continuous at 0. The continuity at any other point x follows from the continuity at 0: let $x_n \to x$; by convexity it follows that $|\mu(x) - \mu(x_n)| \le \mu(x - x_n) \to 0$.

Remark: If $\mu(\lambda x) = |\lambda|\mu(x)$, then $\mu(x) \le \frac{\varepsilon}{\delta}\|x\|$, i.e., there exists C such that $\mu(x) \le C\|x\|$.

Theorem 10.3.2 (Banach-Steinhaus) *Let $\{A_\alpha :\ X \to Y\}_\alpha$ be a family of bounded operators between two Banach spaces X and Y, and let $\sup_\alpha \|A_\alpha x\| \le Cx$. Then there exists C such that $\|A_\alpha x\| \le C\|x\|$ for every x and any A_α from the family. This means that $\{A_\alpha\}$ is a bounded set in $L(X \to Y)$.*

Proof: Introduce the function $\mu(x) = sup_\alpha \|A_\alpha x\|$. All of the conditions of the Zabreiko theorem are obvious. So, there exists C such that $\mu(x) \le C\|x\|$ which is the statement of the theorem. □

Corollary 10.3.3 *Let X be a Banach space over the field $\mathbb{F}$ (which in our theory is either $\mathbb{R}$ or $\mathbb{C}$). Let $A = \{f \in X^*\}$ be the set of all bounded linear functionals such that for every $x \in X$ $\{f(x)\}_{f\in A}$ is bounded, i.e., $\sup_{f\in A} |f(x)| \le Cx$. Then A is a bounded set in X^*, that is, there exists C such that $\|f\| \le C$ for every $f \in A$.*

In the w^*-topology of X^* the boundness of a set $A \subseteq X^*$ is defined by the boundness for every $x \in X$: $|f(x)| \le Cx$ for every $f \in A$; so the Corollary means that w^*-boundness implies boundness in the norm topology. Indeed, this follows if one uses the Banach-Steinhaus theorem for $\{f :\ X \to \mathbb{F}\}_{f\in A}$.

Corollary 10.3.4 *Let $A \subseteq X$ be a set in X bounded in the w^*-topology meaning that for every $f \in X^*$, $\sup_{x\in A}|f(x)| \leq C(f)$. Then A is bounded in the norm topology: there exists C such that $\|x\| \leq C$.*

Indeed, use the Banach-Steinhaus theorem for the family of linear operators

$$\{x: X^* \to \mathbb{F}\}_{x\in A}. \tag{10.17}$$

Corollary 10.3.5 *Let $S \subseteq L(X \to Y)$ be the set of bounded operators such that for every $x \in X$ and every $y \in Y^*$ we have that $|f(Tx)| \leq C(x; f)$ for all $T \in A$. Then there exists C such that $\|T\| \leq C$ for all $T \in A$.*

We now show a few examples that use the Banach-Steinhaus theorem.

1. Let H be a Hilbert space and $A : H \to H$ be a linear operator with $\mathrm{Dom}A =???$ (not yet necessary continuous) and $A = A^*$, i.e., $(Ax, y) = (x, Ay)$. Then A is continuous.

Indeed,

$$|(Ax, y)| = |(x, Ay)| \leq \|Ax\| = C(y) \tag{10.18}$$

for all $x \in X$ with $\|x\| \leq 1$. Then $\{Ax\}_{x\in\mathcal{D}(X)}$ is bounded: $\|Ax\| \leq C\|x\|$.

2. *Integration formulas.* Define $\phi(f) = \int_0^1 f(t)dt,\ f \in C[0,1]$. Let $\{t_k^{(n)}\}_{k=1}^{k_n} \subseteq [0,1]$ and let $c_k^{(n)}$ be numbers such that

$$\phi_n(f) \stackrel{def}{=} \sum_1^k c_k^{(n)} f(t_k^{(n)}) = \phi(f), \tag{10.19}$$

for all f being polynomial up to degree n. Then we have the following:

Theorem 10.3.6 (Polya) *$\phi_n(f) \to \phi(f)$ for all $f \in C[0,1]$ if and only if there exists M such that $\sum_k |c_k^{(n)}| < M$.*

Proof: Check first that $\|\phi_n\|^* = \sum_k |c_k^{(n)}|$ (the norm as a linear functional over $C[0,1]$). Then $\phi_n(f) \to \phi(f)$ on a dense set of functions (in our case for every polynomial) and the uniform boundness $\|\phi_n\| < M$ implies the convergence $\phi_n(f) \to \phi(f)$ for every $f \in C[0,1]$.

Now in the opposite direction, we use the Banach-Steinhaus Theorem: if $\phi_n(f) \to \phi(f)$ for all $f \in C[0,1]$ meaning, in particular, $|\phi_n(f)| < C(f)$, then there exists M such that $\|\phi_n\| \le M$. □

3. *Use of the Banach-Steinhaus Theorem for establishing counter examples in Analysis.* We give one example: There exists function $f \in C[\pi, \pi]$ such that $\|S_n f\| \to \infty$ as $n \to \infty$ where

$$(S_n f)(t) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin(n+\frac{1}{2})(\tau - t)}{\sin\frac{\tau - t}{2}} f(\tau) d\tau \tag{10.20}$$

Then

$$\|S_n\| = \frac{1}{2\pi} \sup_t \int_0^{2\pi} \left| \frac{\sin(n+\frac{1}{2})(\tau - t)}{\sin\frac{\tau - t}{2}} \right| d\tau \tag{10.21}$$

$$= \frac{1}{\pi} \int_0^{\pi} \left| \frac{\sin(2n+1)\tau}{\sin\tau} \right| d\tau \tag{10.22}$$

$$= \frac{1}{\pi} \sum^{2} n_0 \int_{\frac{k\pi}{2n+1}}^{\frac{k+1}{2n+1}\pi} |\cdots| d\tau \tag{10.23}$$

$$\ge \frac{\ln n}{8\pi} \to \infty \tag{10.24}$$

as $n \to \infty$. Therefore, there exists f such that $\|S_n f\| \to \infty$ as $n \to \infty$.

In the few following theorems we demonstrate the use of Banach Theorem on open map.

Theorem 10.3.7 *Let $E_i \hookrightarrow X$ be closed subspaces of a Banach space $X, E_1 \cap E_2 = 0$ and $E_1 + E_2 = X$. Then the projection $P : X \to E_1$, parallel to E_2 (i.e Ker $P = E_2$) is a bounded operator. In other words, $\exists C > 0$ such that $\|x_1 + x_2\| \ge c \max(\|x_1\| x_2\|)$ for $x_i \in E_i$.*

Proof: X/E_2 is a Banach space (because E_2 is closed) and obviously there is a natural linear isomorphism $X/E_2 \approx E_1$. Also $\|x_1 + E_2\|_{X/E_2} = \inf\{\|x_1 + y\|\|| \; y \in E_2\} \le \|x_1\|$ (for $x_1 \in E_1$). So (again because E_1 is a Banach space), X/E_2 and E_1 are isomorphic as Banach spaces and $\exists C$ such that

$$\|x_1\| \le C\|x_1 + y\| \; \forall x_1 \in E_1, y \in E_2. \tag{10.25}$$

This means the boundness of the projection.

10.4 Bases In Banach Spaces

Let X be a Banach space and $e = \{e_k\}_1^\infty$ be a linearly independent complete system in X. Introduce the linear projections

$$U_n\left(\sum_1^m a_k e_k\right) = \sum_1^n a_k e_k, \tag{10.26}$$

for $m \geq n$ and for all $a_k \in \mathbb{R}$. Of course these projections are defined only on a dense set of all finite linear combinations of e. Note that if there exists a linear functional $e_k^* \in X^*$ such that $e_k^*(e_k) = \delta_{kn}$ (called biorthogonal functionals) then we call e a *minimal system.* Clearly, in this case

$$x = \sum_1^n a_k e_k = \sum_1^n e_k^*(x) e_k \tag{10.27}$$

and $U_n x = \sum_1^n e_k^*(x) e_k$. This implies $\|U_n\| \leq \sum_1^n \|e_k^*\| \cdot \|e_k\| = c_n < \infty$. Also in the opposite direction, if U_n are bounded operators then $U_n x - U_{n-1} x = e_n^*(x) e_n$ and $\|e_n^*\| \cdot \|e_n\| \leq \|U_n - U_{n-1}\|$ and $e_n^* \in X^*$. Thus we have the following:

Fact: e is a minimal system if and only if the U_n are bounded operators.

We call e a basis of X if and only if for every $x \in X$ there is exactly one decomposition $x = \sum_1^\infty a_k e_k$. We call the basis e a *Shauder basis* if in addition e is a minimal system.

Theorem 10.4.1 (Banach) *Every basis of a Banach space is a Schauder basis.*

We will prove this theorem in the following form

Theorem 10.4.2 *Let e be a complete linearly independent system. The e is a basis of a Banach space x if and only if the projections U_n (defined above) are bounded and moreover there exists C such that $\|U_n\| \leq C$ (i.e. U_n are uniformly bounded).*

Proof: We prove first the sufficient condition: if there exists C such that $\|U_n\| \leq C$ then e is a basis (and, automatically a Schauder basis) because e is a minimal subsystem. Let

$$A = \left\{x \in X \mid x = \sum_1^\infty a_n e_n \text{ converges}\right\}. \tag{10.28}$$

This means that $U_n x \to x$ $(n \to \infty)$ for $x \in A$. Clearly, A is dense in X. We will show that A is closed, meaning that $A = X$ and $\forall x \in X$, $x = \sum_1^\infty a_n e_n$ (the uniqueness of the series is trivial by the minimality of e). Indeed, let $x_j \in A$ and $x_j \to x$. Then $U_n x_j \to x_j$ $(n \to \infty)$. For $\varepsilon > 0$ there exists j_0 such that $\|x_{j_0} - x\| < \varepsilon$. Then there exists N such that for every $n > N$, $\|U_n x_{j_0} - x_{j_0}\| < \varepsilon$. Therefore for $n > N$ we have that,

$$\begin{aligned}
\|U_n x - x\| &\leq \|U_n x - U_n x_{j_0}\| + \|U_n x_{j_0} - x_{j_0}\| + \|x_{j_0} - x\| \quad (10.29)\\
&\leq (\|U_n\| + 1)\|x - x_{j_0}\| + \|U_n x_{j_0} - x_{j_0}\| \quad (10.30)\\
&\leq (C+1)\varepsilon + \varepsilon. \quad (10.31)
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we show that $U_n x \to x$.

Now we prove that this condition is necessary. Define a new norm by $\|_1 x\|_1 = \sup_n \| \sum_1^n a_i e_i\| \geq x\|$ which is defined for every $x \in X$. Obviously, the operators $U_n x = \sum_1^n a_i e_i$ are defined and $\|_1 U_n\|_1 = 1$ Then, by the sufficient condition proved before U_n are uniformly bounded, e is a basis in the completion $\widehat{X}$ of X in the $\| \cdot \|_1$ norm (since U_n are uniformly bounded). However, for every $\hat{x} \in \widehat{X}$, $\hat{x} = \sum_1^\infty a_i e_i$ and the series being converging in $\|_1 \cdot \|_1$, also converges in $\| \cdot \|$ and define $x = \sum_1^\infty a_i e_i \in X$.

By uniqueness, we have the identity map:

$$id : X \mapsto \widehat{X}, \quad (10.32)$$

between two Banach spaces, is one-to-one, onto and $\|_1 x\|_1 \geq \|x\|$. By the Banach Theorem there exists C such that $\|_1 x\|_1 \leq C\|x\|$. This means that $\|U_n\| \leq C$.

10.5 Hahn-Banach Theorem. Linear functionals

Let X be a Banach space. $X^\#$ is the Banach space of all linear functionals and X^* is the space of continuous linear functionals or, equivalently, the space of bounded linear functionals, that is, there exists C such that $|f(x)| \leq C\|x\|$. Then $(X^*, \| \cdot \|^*)$ where $\|f\|^* = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|}$. This space is (always) complete.

Examples. $\|f\| = \frac{1}{\operatorname{dist}\{0, H_f(1)\}}$, where $H_f(1) = \{x \mid f(x) = 1\}$.

Sublinear functionals.

Let L be a linear space over $\mathbb{R}$; a real function $p(x) \geq 0$ is called sublinear if

(i) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in L$

(ii) $p(\lambda x) = \lambda p(x)$ for $\lambda \geq 0$.

Theorem 10.5.1 (Hahn-Banach) *Let $p(x) < \infty$ for all $x \in L$, $L_0 \hookrightarrow L$ be a subspace and let $f_0 \in L_0^{\#}$ be a linear functional defined on L_0. Let $f_0(x) \leq p(x)$ for all $x \in L_0$. Then there exists $f(x) \in L^{\#}$ such that*

(i) $f(x) \leq p(x)$ for all $x \in L$ and

(ii) $f|_{L_0} = f_0$

($f(x) = f_0(x)$ for all $x \in L_0$; so f is an extension of f_0).

Proof: Introduce a partial ordering $P = \{(L_\alpha \hookrightarrow L; f_\alpha)\}$ of pairs of subspaces $L_\alpha \hookrightarrow L$ with $L_0 \hookrightarrow L_\alpha$ and a linear functional $f_\alpha \in L_\alpha^{\#}$ such that $f_\alpha(x) \leq p(x)$ for all $x \in L_\alpha$ and $f_\alpha|_{L_0} = f_0$, by defining $(L_\alpha, f_\alpha) \prec (L_\beta, f_\beta)$ if and only if $L_\alpha \hookrightarrow L_\beta$ and $f_\beta|_{L\alpha} = f_0$. Note that for any linear chain (L_α, f_α) there is a supremum element $(L_\infty = \cup L + \alpha, f_\infty) \in P$. Then, by Zorn's Lemma there exists a maximal element $(L_{\alpha_0}, f_{\alpha_0})$. We have to prove that $L_{\alpha_0} = L$. If $y \in L \setminus L_{\alpha_0}$ then define the subspace $L_1 = \mathrm{span}\{y, L_{\alpha_0}\}$. Any $z \in L_1$ $z = ay + x,\ a \in \mathbb{R},\ x \in L_{\alpha_0}$. For any extension of f_{α_0} on L_1,

$$f(z) = af(y) + f_{\alpha_0}(x). \tag{10.33}$$

So, the choice of the extension is defined by the value $f(y)$.

Let $C = f(y)$. We have to determine conditions on C such that

$$f(ay + x) = aC + f_{\alpha_0}(x) \leq p(ay + x), \tag{10.34}$$

for every $x \in L_{\alpha_0}$ and $a \in \mathbb{R}$. This is, in fact, two conditions: one for $a > 0$ and one for $a < 0$.

$$C \leq p(y + \frac{x}{a}) - f_{\alpha_0}\left(\frac{x}{a}\right) \text{ for } a > 0 \tag{10.35}$$

and

$$-p\left(-\frac{x}{a} - y\right) - f_{\alpha_0}\left(\frac{x}{a}\right) \leq C \text{ for} a < 0. \tag{10.36}$$

Call $x_1 = x/a$ in (10.35) and $x_2 = x/a$ in (10.36) and consider x_1 and x_2 as different (independent) vectors. Then the condition

$$-p(-x_2 - y) - f_{\alpha_0}(x_2) \leq p(y + x_1) - f_{\alpha_0}(x_1) \tag{10.37}$$

for every x_1 and x_2 in L_{α_0} would imply the existence of C satisfying (10.35) and (10.36) and, as a consequence it would satisfy also (10.34). We rewrite (10.37) as:

$$f_{\alpha_0}(x_1 - x_2) \leq p(y + x_1) + p(-x_2 - y) \tag{10.38}$$

which follows from the fact $f_{\alpha_0}(x_1 - x_2) \leq p(x_1 - x_2) \leq p(x_1 + y) + p(-y - x_2)$ (the first inequality is satisfied for any vectors in L_{α_0} and the second is the triangle inequality satisfied for for every $y \in L$).

So, there is an extension of f_{α_0} to a subspace L_1 which satisfies all the conditions of our order, meaning that $(L_{\alpha_0}, f_{\alpha_0})$ is not a maximal element. This is a contradiction, hence $L_{\alpha_0} = L$.

Proof of the complex case: We assume now that $(L; \mathbb{C})$, $p(\lambda x) = |\lambda| p(x)$ for $\lambda \in \mathbb{C}$ and $|f_0(x)| \leq p(x)$ for every $x \in L + 0 \hookrightarrow L$. Then there exists extension $f(x) \in L^{\#}$ (complex) linear functional such that $f|_{L_0} = f_0$ and $|f(x)| \leq p(x)$ for every $x \in L$.

Indeed: Consider $\phi_0(x) = Ref_0(x)$ which is a real valued linear functional on L_0 as a linear space over $\mathbb{R}$. Then for every $x \in L_0$

$$\phi_0(x) \leq |\phi_0(x)| \leq p(x). \tag{10.39}$$

So, by the real case of the Hahn-Banach theorem it follows that there is an extension $\phi \in L^{\#}$ (considering L as a space over $\mathbb{R}$).

Note that $-\phi(x) = \phi(-x) \leq p(-x) = p(x)$. Therefore for every $x \in (L, \mathbb{R})$,

$$|\phi(x)| \leq p(x). \tag{10.40}$$

Note now the connection between the complex linear functional $f(x) = \phi(x) + i\psi(x)$, $Ref = \phi$, $Imf = \psi$ and its real part $\phi(x)$:

$$if(x) = f(ix) - \phi(ix) + i\psi(ix). \tag{10.41}$$

So, $f(x) - \psi(ix) - i\phi(ix)$ and $Imf(x) = -\phi(ix)$. Therefore, if the $\phi(x)$ above is a real-valued linear functional then

(i) $f(x) = \phi(x) - i\phi(ix)$ is a (complex-valued) linear functional over $\mathbb{C}$.

(ii) Clearly, f is an extension of f_0: $Ref|_{x\in L_0} = \phi_0$ meaning that $f_0 = Ref_0 + iImf - 0 = \phi_0 - i\phi_0(i\cdot)$.

(iii) Check now that $|f(x)| \le p(x)$: $f(x) = |f(x)|e^{i\theta(x)}$ and $f(e^{-i\theta(x)}x) = |f(x)|$.

Thus the inequality $|\phi(x)| \le p(x)$ implies $|f(x)| \le p(x)$. □

Corollary 10.5.2 *Let X be a normed space, $E_0 \hookrightarrow X$ be a subspace and $f_0 \in E_0^*$. Then there exists $f \in X^*$ such that $f|_{E_0} = f_0$ and $\|f\|_{X^*} = \|f_0\|_{E_0^*}$.*

In order to prove this, use the Hahn-Banach Theorem for $p(x) = a\|x\|$ where $a = \|f_0\|_{E_0^*}$.

Corollary 10.5.3 *For every x_0, $\|x_0\| = 1$ there exists $f_0 \in X^*$, $\|f_0\| = 1$ and $f_0(x_0) = 1$. (So f_0 is a supported functional at $x_0 \in S(X)$.)*

In order to prove this, one can consider the one-dimensional subspace $E_0 = \{\lambda x_0\}$ and the functional $f_0(x) = \lambda$ for $x = \lambda x_0$. Clearly $\|f_0\|_{E_0^*} = 1$. Then consider an extension f of f_0 with the same norm.

Corollary 10.5.4 *For every x_0 there exists $f_0 \neq 0$ such that $f_0(x_0) = \|x_0\| \cdot \|f_0\|$.*

Corollary 10.5.5 *For every $x_1 \neq x_2$ there exists $f \in X^*$ such that $f(x_1) \neq f(x_2)$. This means that X^* is a total set.*

Corollary 10.5.6 *The $\max_{f\neq 0} \frac{|f(x)|}{\|f\|}$ exists and equals $\|x\|$. This means that $X \hookrightarrow X^{**}$ and*

The Minkowski functional:

Let M be a convex set, $0 \in M \subset E$ consider

$$p_M(x) = \begin{cases} 0 & x = 0 \\ \infty & \text{if } \nexists t \in \mathbb{R}^+ \ : \ x/t \in M \\ \inf\{t \in \mathbb{R}^+ | x/t \in M\} & \text{otherwise} \end{cases} \tag{10.42}$$

Obviously $p_M(\lambda x) = \lambda p_m(x)$ for $\lambda > 0$. (homogeneity) and p_M is a convex functional: $p(x_1 + x_2) \le p(x_1) + p(x_2)$.

Indeed: Fix $\varepsilon > 0$. Take t_i such that $p(x_i) < t_i \leq p(x_i)+\varepsilon$ (of course, only the case $p(x_i) < \infty$ is non-trivial). So $x_i/t_i \in M$. Let $t = t_1 + t_2$. Then $\frac{x_1+x_2}{t_1+t_2} = \frac{t_1 x_1}{t \cdot t_1} + \frac{t_2 x_2}{t \cdot t_2} \in [\frac{x_1}{t_1}, \frac{x_2}{t_2}] \subset M$ (by convexity of M).
Therefore, $P_M(x_1 + x_2) \leq t = t_1 + t_2 \leq p(x_1) + p(x_2) + 2\varepsilon$. Note also that $0 \in M^c$ implies (and equivalent X_0).

$$p(x) < \infty \; \forall x \in E \tag{10.43}$$

Theorem 10.5.7 (on separation of convex sets) *Let M_1 and M_2 be convex sets in E and let $M_1^c \neq \emptyset$ and $M_1^c \cap M_2 = \emptyset$. Then $\exists f \in E^\sharp$ such that $f(M_1) \leq C \leq f(M_2)$ [meaning for every $x \in M_1, f(x) \leq C$ and $\forall y \in M_2, f(y) \geq C$].*

Proof: Consider the set $M = M_1^c - M_2$. Then $0 \notin M, M^c \neq \emptyset$ (because $M_1^c \neq \emptyset$). We want to build a functional $f \neq 0, f \in E$, and $f(M) \leq 0$, meaning $f(x) \leq 0$ for every $x \in M$. Let $x_0 \in M^c$. Introduce still another set $M_0 = M - X_0$. Then $0 \in M_0^c$ and the Minkowski's functional $P_{M_0}(x) \equiv p(x)$ is defined and finite for every $\forall x \in E$. Also $-x_0 \notin M_0$ (because $0 \notin M$). Consider the 1-dim space $E_0 = \{\lambda x_0\}$ and define the (linear) functional $f_0(\lambda x_0) = -\lambda$ on E_0. Since $p(-x_0) \geq 1$ (recall $-x_0 \notin M_0$) we have $f_0(\lambda x_0) \leq p(\lambda x_0)$. So, by the Hahn-Banach theorem, there exists extension $f(x) \leq p(x)$. Then $f(x) \leq 1$ for $x \in M_0$ and $f(x_0) = f_0(-x_0) = 1$. Therefore $\forall y \in M, f(y) \leq 0$.

Corollaries of Hahn-Banach theorem; continuation.

If X^* is separable space then X is also separable.

Indeed: Let $\{f_i\}$ be a dense set in $S(X^*)$- the unit space of X^*. Let $x_i \in S(X)$ such that $|f_i(x_i)| \geq \frac{1}{2}$. Consider $E = \text{span}\,\{x_i\}$. If $E = X$ then X is separable. But if $E \neq X$ then $\exists f \in X^*, \|f\| = 1, f(E) = 0$. Now, for any $\varepsilon > 0$ $\exists f_i$ and

$$\varepsilon \geq \|f - f_i\| \geq |(f - f_{i_0})(x_{i_0})| = |f_{i_0}(x_{i_0})| \geq \frac{1}{2} \tag{10.44}$$

a contradiction.

Lemma 10.5.8 (Mazur) *Let $x_n \xrightarrow{w} x_0$ [meaning that for all $f \in X^*$, $f(x_n) \to f(x_0)$]. Then $x_0 \in Conv\{x_i\}_1^\infty$.*

Proof: Indeed, if $x_0 \notin K = \overline{conv}\{x_i\}_i^\infty$ then use the previous Corollary: $\exists\ f$ and $f(x_0) < \inf\ f(x_i)$ which contradicts the weak convergence $x_n \xrightarrow{w} x_0$.

Remark: Let $K_n = \overline{Conv}\{x_i\}_n^\infty$. Then $x_0 \in K_n$ for every n and $x_0 \in \cap_1^\infty K_n$. [Check that, in fact $x_0 = \cap K_n$].

13. Let $K \subset X$ be a convex set. Then K is closed iff K is w-closed (i.e. closed in weak topology).

Indeed, K-closed iff $\forall x \notin K\ \exists\ f$ which separates x from K.

Let K be closed convex set and $x_0 \notin K$. Then $\exists\ f \in X^*$ such that $f(x_0) < \inf\ f(x) x \in K$. Indeed, $\exists d > 0$ and the ball $\mathcal{D}(x_0; d)$ with center x_0 and radius d such that $\mathcal{D}(x_0, d) \cap K = \emptyset$. Use the theorem on the separation of convex sets for $M_1 = \mathcal{D}(x_0; d)$ and $M_2 = K$.

Then K is the intersection of all layers $\{x | a \le f(x) \le \beta\} = H_f(a; \beta)$ such that $K \subset H_f(a; \beta)$. This means that K is w-closed set.

14. Let $E \hookrightarrow X$ be a closed subspace. Then

(i) $(X/E)^* = E^1 \hookrightarrow X^*$ and

(ii) $E^* = X^*/E^1$.

15. Let X be reflexive space [i.e. $X = X^{**}$]. Then every closed subspace $E \hookrightarrow X$ is also reflexive. [Obviously from $14: E^* = X^*/E^1$ and $(X^*/E^1)^* = (E^1)^1 = E$]. w^*-topology. w^*-topology is defined in the dual space $Y = X^*$. Then we define a weak topology in Y using only linear functionals from $X \hookrightarrow X^{**}$: subbasis of neibourhoods of $0 \in Y$ is defined by $\varepsilon > 0$ and $x \in X : U_{x;\varepsilon}(0) = \{y \in Y |\ |y(x)| < \varepsilon\}$

Theorem 10.5.9 (Alaoglu) . $\mathcal{D}(X^*) = \{f \in X^* |\ \|f\| \le 1\}$ *is a compact set in* w^**-topology.*

Proof: Let $I_x = [-\|x\|, \|x\|]$ and $K = \prod I_x$ be a product of interval named by elements $x \in X$ and equipped with the product (Tihonov) topology. Then K is a compact in this topology. Consider the one-to-one embedding $\mathcal{D}(X^*) \subset_i K : f \in \mathcal{D}(X^*)$ corresponds $i(f) = (f(x))\ x \in X \in K$. Note that w^*-topology on $\mathcal{D}(X^*)$ is exactly restriction of the product topology on K. Also $\mathcal{D}(X^*)$ is intersection of closed subsets in $K : \{\text{linear relations} : f(x_1 + x_2) = f(x_1) + f(x_2) + \cdots\}$

The next fact is describing convex w^*-closed sets in X^*.

Theorem 10.5.10 *If* $K \subset X^*$ *is a convex set and* w^**-closed and if* $f_0 \notin K (f_0 \in X^*)$ *then* $\exists x \in X$ *such that* $x f_0 > \sup\{\varphi(x) | \varphi \in K\}$. *Therefore,* $K = \cap_{x \in X} H^-_{a(x)}$.

Proof: We need the following lemma:

Lemma 10.5.11 *Let* $L \hookleftarrow E = \cap_1^n Kerf_i,\ f_i \in K^\sharp$, *and* $f(E) = 0$ *for some* $f \in L^\sharp$ *(i.e.* $E \subset Kerf$.*) Then* $f = \sum_1^n a_i f_i$.

Proof: Consider the space L/E. Then $F_i \in (L/E)^\sharp$ and also $f \in (L/E)^\sharp$. Note, $\{f_i\}_i^n$ is a total set on L/E. Indeed, if $[X] = x + E \in L/E$ and $f_i([x_x]) = 0 (i = 1, \ldots, n)$ means $f_i(x) = 0$ for $i = 1, \cdots, n$. Then $x \in E$ and $[X] = 0$ in the quotient space. Thus span $\{f_i\}_i^n = (L/E)^\sharp$ and $f \in$ span $\{f_i\}_1^n$ meaning $f = \sum_i^n a_i f_i$ for some members a_i. We return now to the proof of the theorem.

Since K is closed in w^*-topology and $f_0 \notin K$, then there is a neibourhood $U(f_0)$ from a sub-basis set of neibourhoods of f_0 such that $K \cap (U(f_0)) = \emptyset$. The neibourhood $U(f_0)$ is defined by $\varepsilon > 0$ and $x_1, \cdot, x_n \subset X$ such that $U(f_0) = \{\varphi \in X^* | \ |\varphi(x_i) - f_0(x_i)| < \varepsilon\} i = 1, \cdots, n$. Define $= f_0(x_i)$. Now, by Hahn-Banach Theorem, $\exists X \in X^{**}$ which separates K from $U(f_0)$: This means that $\exists a$ and $X(f) \geq a$ for every$f \in U(f_0) \supset \{\cap_i\}\varphi | x_i(\varphi) = a_i\} \overset{def}{\equiv} M$. Let $E = \cup_i^n Kerx_i$ and $M = E + f_0$. We show that $E \subset KerX$. Indeed, if not and $\exists\ \varphi_0 \in E\ KerX$; let $X(\varphi_0) = \beta \neq 0$; also $x_i(\varphi_0) = 0$. Consider $\lambda\varphi_0 + f_0 \in M$ (then $x_i(\lambda\varphi_0 + f_0) = a_i$ for $\forall \lambda \in \mathbb{R}$)

However $X(\lambda\varphi + f_0) = \lambda\beta + X(f_0)$ and λ is any which is a contradiction. By Lemma $X = \sum_1^n \alpha_i x_i \in X$.

Theorem 10.5.12 (Goldstein) *:* $\mathcal{D}(x) \subset (x^{**})$ *and it is dense in the* w^**-topology.*

Proof: Note first that the unit ball of the dual space is closed in the w^*-topology and there $\mathcal{D}(x^{**})$ is closed in w^*-topology. Let $K = \overline{\mathcal{D}(x)}^{w^*} \subset \mathcal{D}(x^{**})$. If $\exists\ X \in \mathcal{D}(x^{**}) \setminus K$ then $\exists f \in X^*$ separating X and K. This means

$$f(x) > sup\{f(x) | x \in K\} \geq \|f\|. \tag{10.45}$$

But then $\|X\| > 1$, contradiction.

Around Eberlain-Schmulian theorem.

Theorem 10.5.13 *: If* X *is reflexive, then for every bounded sequence* x_n *there exists* $x_{n_k} \overset{w}{\to} x_0 \in X$.

Proof: The sufficient condition is easy: If X is reflexive and X be separable consider a dense set $\{f_i\}_1^\infty \subset X^*$.

(Note, that X^* is separable because $X^{**} = X$ is). Let $\{x_n\}$be a bounded sequence. Then choose a subsequence $\{x_n^{(1)}\} \subset \{x_n\}$
such that

$$\exists \lim_{n\to\infty} f_1(x_n^{(1)})(= \alpha(f_1)). \tag{10.46}$$

Choose $\{x_n^{(i)}\} \subset \{x_n^{(i-1)}\} \subset \cdots \subset \{x_n\}$ such that

$$f_i(x_n^{(i)}) \to \alpha(f_i). \tag{10.47}$$

Then $f_n(x_i^{(i)}) \to a(f_n)\ \forall f_n$ as $i \to \infty$ and, since $\{x_n\}$ is bounded and f_n is dense, $f(x_i^{(i)}) \to a(f)$ for some $f \in X^*$. Note that a f is just a name the limit. However, it is easy to see that f is linearly dependent, $f \in X^*$ and $|a(f)| \leq sup|f(x_i^{(i)})| \leq \|f\| \cdot \sup \|x_n\|$. So f is a bounded linear functional, so $a \in X^{**} = X$ by reflexivity (Extend to not necessary separable X. Before we start to prove necessary condition, let us note that it is a trivial consequence of another theorem.

Theorem 10.5.14 (James) *Let X be non-reflexive Banach space. Then $\exists\ f_0 \in X^*$ such that there is no element $x \in X$ such that*

$$f_0(x) = \|f_0\| \cdot \|x\| \tag{10.48}$$

So, normalizing f_0 to satisfy $\|f_0\| = 1$, the affine hyperplane $H_{f_0}(1) = \{x \in X \mid f_0(x) = 1$ has no common point with the unit ball

$$\mathcal{D}(x) = \{x \in X \mid \|x\| \leq 1\} : H_{f_0}(1) \cap \mathcal{D}(x) = \emptyset. \tag{10.49}$$

This is extremely non-trivial fact and we will not treat it here. However let us note how James Theorem implies the remaining part of the Eberlain-Shmulian Theorem: since

$$\|f_0\| = \sup\{f_0(x) \mid \|x\| \leq 1\}, \tag{10.50}$$

take $x_n,\ \|x_n\| \leq 1$, and

$$(1 \geq) f_0(x_n) \geq 1 - \frac{1}{n}. \tag{10.51}$$

Obviously, there is no $x_0 \in X$ and subsequence $\{x_n^{(1)}\} \subset \{x_n\}$ such that $x_n^{(1)} \xrightarrow{w} x_0 \in X$ because otherwise $f_0(x_0) = 1$ and $\|x_0\| \leq 1$ which contradicts the property of f_0.

Returning to the proof of the theorem, we need a few observations and definitions.

10.6 Extremal points; The Krein-Milman Theorem

Consider a linear space L over $\mathbb{R}$. Let $L^\sharp$ be the linear space of all linear functionals on L and $F \subset L^\sharp$ be a subset of separating points of L, i.e. $\forall x_1 \neq x_2 \ \exists \ f \in F$ and $f(x_1) \neq f(x_2)$. Let K be a set in L which $W(F)$-compact (meaning a compact set in the weak topology generated by F). We call a subset $M \subset K$ an extremal set of K iff $\forall x \in M \ x = \alpha x_1 + (1-\alpha)x_2, \ 0 < \alpha < 1$, $x_i \epsilon K$ implies $x_i \in M$. A point $x_0 \in K$ is called an extremal point of K if $\{x_0\}$ is an extremal set of K. Note, that an extremal set M_1 of another extremal set M of K is itself an extremal set of K. Define Extr$K$ the set of all extremal points of K. Examples:

(i) $\mathcal{D}(c_0)$ has no extremal points.

(ii) Extr$(\mathcal{D}(C[0,1])) = \{\pm 1\}$ (so, it is two points set consisting of functions identically $+1$ or -1).

Theorem 10.6.1 (Krein-Milman) . *Let K be a convex compact set in the space $(L, W(F))$. Then*

(i) Extr$K \neq \emptyset$

(ii) $\overline{conv}$Extr$K = K$.

Moreover, we don't need the convexity condition: for every $W(F)$--compact set $K \subset L$, $\overline{conv} K = \overline{conv} Extr K$.

Proof: (i) First we use Zorn Lemma. Consider an ordering $(K_a, <)$ when K_a are extremal compact subsets of K and $K_\alpha < K_\beta$ means $K_\alpha \subset K_\beta$. Note that any linear chain of extremal sets $\{K_\alpha\}$ has a minorant element $K_0 = \cap_\alpha K_\alpha$ (obviously extremal compact set). So, there is a minimal element K_0 in this order. We want to show that K_0 contains only one point.

Assume there $K_0 \supset \{a \neq b\}$. Then $\exists f \in F$ and $f(a) = \alpha < f(b) = \beta$. Consider $V = \{y \in K_0 | \ t = \min f(x) = f(y)\}$. Note that f is a continuous function on the compact K_0, and so the minimum exists. Also $V \subset K_0, b \notin V$ and $V = K_0 \cap \{y| \ f(y) = t\}$. Therefore V is closed, meaning compact in our situation. Also it is extremal set of K_0 and, as a consequence, extremal set of K. This contradicts the minimality of K_0 and proves that $K_0 = \{a\}$ is a point so Extr$K \neq \emptyset$.

(ii) Let $E = \overline{conv}\mathrm{Extr}K$ and $E \subsetneq \overline{conv} K$. Let $a \in \overline{conv} K \setminus E \ \exists f_0 \in F$ and $\alpha = f_0(a) < \min\{f_0(y) \mid y \in E\}$ (by the separation theorem). Again

$V = \{x \in K \mid f_0(x) = \min_{y\in K} f_0(y)\}$ and $V = K \cap \{y \mid f_0(y) = Const\}$. This is a closed set (a compact subset of K) and extremal set of K. By (i) Extr$V \neq \emptyset$ and $ExtrV \subset ExtrK$. But $V \cap E = \emptyset$, a contradiction.

Examples.

1. (Birkhoff's Theorem) Let K be the set of all double stochastic matrices in $\mathbb{R}^n$:

$$K = \{(a_{ij})_{i,j=1}^n \mid a_{ij} \geq 0, \ \sum_{j=1}^n a_{ij} = 1 = \sum_{i=1}^n a_{ij}\}. \tag{10.52}$$

(Clearly this is a convex subset of the M_n-space of all $n \times n$ matrices.) Then K is a convex combination of the permutations in

$$\Pi = \Big\{(a_{ij}), \ where \ a_{ij} \in \{0,1\}\Big\}. \tag{10.53}$$

Indeed, $\mathrm{Extr}K$ is exactly the set of the permutations Π.

2. $C([0,1])$, c_0, L_1 are not dual spaces to any other Banach space. Indeed, if $X = Y^*$ for some Y then $\mathcal{D}(X)$ is compact in the w^*-topology and has a lot of extremal points:

$$\overline{\mathrm{convExtr}\mathcal{D}(X)}^{w^*} = \mathcal{D}(X). \tag{10.54}$$

(i.e. there is a topology such that the closure in this topology is $\mathcal{D}(X)$). But these spaces either do not have any extremal points or, in the case of $C([0,1])$, do not have enough.

3. Let M be the set of the probability measures on a compact set K. Then $\mathrm{Extr}M = \{\delta_x \ : \ x \in K\}$.

4. The unitary operators in $\mathbb{C}^n$ form the set of extremal points of $L(\ell_2^n \to \ell_2^n)$.

5. (Herglotz Theorem) Let K be the set of all analytic functions defined on $\{z \in \mathbb{C} \ : \ |z| < 1 \text{ and } Reu(z) > 0; \text{ with normalization } u(0) = 1\}$.

Chapter 11

Banach algebras

A BANACH SPACE A is called a *Banach Algebra* if an operation "product" $x \cdot y$ is defined for the elements of A and it is continuous with respect to every variable (x and y). We always require that A contains an identity element and that this product is associative: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$. We also require that this product is linear with respect to both variables.

The theory of Banach algebras was developed by Gelfand in the end of the thirties and in the forties.

Theorem 11.0.2 *There exists an equivalent norm $|x|$ on A such that*

$$|x \cdot y| \leq |x| \cdot |y| \quad \text{and} \quad |e| = 1. \tag{11.1}$$

Proof: Consider a map $x \mapsto T_x \in L(A \to A)$ by $T_x y = x \cdot y$. The property $T(y \cdot z) = Ty \cdot z$ defines operators in $L(A \to A)$ which is the image of this map (i.e. $A \hookrightarrow L(A \to A)$). Indeed, define $Te = x$ and then $Tz = T_x z$ $(x \cdot z)$. Note that this property is closed (even with respect to strong topology). Therefore A is a closed subspace of $L(A)$ and it is complete in both topologies. Now, we have the operator norm on A:

$$\|A_x\|_{op} \geq \left\| x \cdot \frac{e}{\|e\|} \right\| = \frac{\|x\|}{\|e\|} \text{ and } \|I\| = 1 \ (I = A_e). \tag{11.2}$$

By Banach theorem the open map $\|A_x\|_{op}$ is equivalent with $\|x\|$ and obviously $\|x \cdot y\|_{op} \leq \|x\|_{op} \|y\|_{op}$.

Note that if $\|x \cdot y\| \leq \|x\| \cdot \|y\|$ and $\|e\| = 1$ then $\| \cdot \|_{op} = \| \cdot \|$ ($\|x\| \leq \|e\| \cdot \|A_x\| = \sup_{\|y\|=1} \|x \cdot y\| \leq \|x\|$). Note also that the product map $x \cdot y$ is continuous with respect to both variables.

Examples. 1. $C[0,1]$ (obviously $\|x \cdot y\|_C \le \|x\|_C \cdot \|y\|_C$) or $C(K)$.

2. W (the Weiner ring): $z \in W$ iff $z = \sum_{-\infty}^{\infty} c_n e^{int}$ and $\|z\| = \sum_{-\infty}^{\infty} |c_n| < \infty$. $x(t) \cdot y(t) = z(t)$ (as functions on $[0, 2\pi]$) which means that $x = \sum a_n e^{int}$ and $y = \sum b_n e^{int}$ implies $z = \sum c_n e^{int}$ for $(c_n) = (a_n) * (b_n) = \left(\sum_{m=-\infty}^{\infty} a_{n-m} b_m\right)$ (convolution of sequences).

3. $L_1[-\infty, \infty]$: $x * y = \int_{-\infty}^{\infty} x(t-\tau) y(\tau) d\tau$. Similarly for $L_1[0,1]$: $x * y = \int_0^t x(t-\tau) y(\tau) d\tau$ (in order to deal with convolution for any functions from L_1, consider first continuous functions, prove that $\|x * y\|_{L_1} \le \|x\|_{L_1} \cdot \|y\|_{L_1}$ and then extend the operation using continuity for all L_1).

Let $\mathcal{O}$ be the set of invertible elements of A:

$$\mathcal{O} = \{x \in A \mid \exists x^{-1}\}. \tag{11.3}$$

Note that $\mathcal{O}$ is an open subset of A. We proved this in the part of the course that dealt with operators, when we showed that for any x with $\|x\| < 1$ there exists $(e - x)^{-1}$ and then if there exists x^{-1} and $\|y\|$ is very small then $x - y = x(e - x^{-1}y)$ is invertible.

Let I be a proper ideal (meaning an ideal which is not trivial (iėṅot equal to) or A)). Note that since it is proper it can not contain *any* invertible element. Then $\overline{I}$, the closure of I, is a proper ideal (because an open ball of radius 1 around e is not contained in I hence it is not contained in $\overline{I}$). We make now a few observations:

1. If there exists the inverse of (x, y) then both x and y are invertible. Indeed, $xy(xy)^{-1} = e$, hence $y(xy)^{-1} = x^{-1}$ (commutativity used).

2. x is invertible iff x does not belong to any proper ideal. Indeed, if x is invertible then $x \cdot A = A$ thus the minimal ideal spanned by x is all of A. If, on the other hand, x is not invertible then $x \cdot A$ is a proper ideal (if $xy_n \to e$ then there exists $(xy_n)^{-1}$ which implies the existence of x^{-1}).

Corollary 11.0.3 *If A does not have any proper ideal then A is a field.*

We call $M \subseteq A$ a *maximal ideal* if there is no proper ideal which contains the ideal M. Note that if M is maximal ideal then M is closed (if not, then its closure is an ideal that contains M).

Example $C[0,1]$. Obviously, $M_\tau = \{x \mid x(\tau) = 0\}$ or a fixed $\tau \in [0,1]$ is a maximal ideal (it is both a hyperplane and an ideal). In the opposite direction, if M is a maximal ideal then there is a $\tau \in [0,1]$ and $M = M_\tau$ (as above).

Indeed, if such a τ does not exist and $\forall \tau \in [0,1]$ there exist $x_\tau \in C[0,1]$ and $x_\tau(\tau) \neq 0$, then there is an open interval I_τ around τ and $x_\tau(I_\tau) \neq 0$. Take a finite covering $\{I_{\tau_i}\}_{i=1}^N$ and consider the function

$$x = \sum_{i=1}^{N} x_{\tau_i}^2(t) \neq 0 \text{ forevery } t \in [0,1]. \tag{11.4}$$

Thus $x \in M$ and x has is invertible, meaning $M = A$, a contradiction.

Theorem 11.0.4 (Gelfand) *For every proper ideal $I \subseteq A$ there exists a maximal ideal M such that $I \subseteq M$.*

Proof: We use the lemma of Zorn (historically this is the first use of this lemma in functional analysis). Consider the set $\mathcal{A}$ of all proper ideals $J \supseteq I$. Define $J_1 \succ J_2$ iff $J_1 \supseteq J_2$. Obviously for any chain $\mathcal{F} = \{J_i\}$ (that is for all $J_1, J_2 \in \mathcal{F}$ either $J_1 \succ J_2$ or $J_2 \succ J_1$) there is a majorizing element

$$J = \cup_{J_i \in \mathcal{J}} J_i \tag{11.5}$$

which is a proper ideal. Consequently by Zorn's lemma there exists a maximal ideal. □

Corollary 11.0.5 *x is invertible iff x does not belong to any maximal ideal M.*

(Exercise).

Let I be a closed ideal of the algebra $\mathcal{A}$ Then A/I is a Banach algebra. Indeed,

(i) $(x+I)(y+I) \subseteq xy+I$ and let $\|x+I\|_{A/I} \simeq \|x+z_1\|$, $\|y+I\|_{A/I} \simeq \|y+z_2\|$ ($z_i \in I$). Then $\|xy+I\|_{A/I} \leq \|x+z_1\| \cdot \|y+z_2\| \simeq \|x+I\|_{A/I} \cdot \|y+I\|_{A/I}$.

(ii) $\|e_{A/I}\| = 1$. Indeed, it is clear that $\|e_{A/I}\| \leq 1$ and since $\|e - x\| < 1$ implies the existence of x^{-1} we get that $\|e_{A/I}\| \geq 1$.

Note also that if $J \supseteq I$ then J is a proper ideal of A iff J/I is a proper ideal of A/I.

Corollary 11.0.6 *Let M be a maximal ideal of $\mathcal{A}$. Then $\mathcal{A}/M$ is a field and if for some closed ideal I, $\mathcal{A}/I$ is a field then $I = M$ a maximal ideal.*

For an example one can see that $C[0,1]/M = \mathbb{R}$ (or $\mathbb{C}$) were M is maximal ideal of $C[0,1]$.

11.1 Analytic functions

Let x be a function from $\mathbb{C}$ to the algebra $\mathcal{A}$, that is, $x(\lambda) \in \mathcal{A}$ for every $\lambda \in \mathbb{C}$. We say that the function $x(\lambda)$ is analytic at λ_0 if the complex derivative $x'(\lambda_0)$ exists (convergence with respect to the norm of the algebra $\mathcal{A}$). Then for any $f \in \mathcal{A}^*$, $f(x(\lambda))$ is an analytic function (this can be also taken as an equivalent definition to a function x being analytic).

An example is the the function $(z-\lambda e)^{-1}$. This function is analytic at every regular point $\lambda \in \mathbb{C}$ (meaning, at every point that the inverse element exists). For this function we have:

$$\frac{(z-\lambda_1 e)^{-1} - (z-\lambda_2 e)^{-1}}{\lambda_1 - \lambda_2} = (z-\lambda_1 e)^{-1}(z-\lambda_2 e)^{-1} \tag{11.6}$$

which gives $\left((z-\lambda e)^{-1}\right)'_\lambda = (z-\lambda e)^{-2}$.

The Cauchy integral is defined by

$$f\left(\int_\Gamma x(\lambda) d\lambda\right) = \int_\Gamma f\left(x(\lambda)\right) d\lambda \tag{11.7}$$

for all $f \in A^*$ where Γ is a rectifiable curve. Now we have the following theorem:

Theorem 11.1.1 (Cauchy) *Let $\Gamma = \partial D$ as above, D being simply connected and $x(\lambda)$ is analytic in a neighborhood of D. Then,*

$$\int_\Gamma x(\lambda) d\lambda = 0. \tag{11.8}$$

Indeed, let $\int_\Gamma x(\lambda) = y$. By the Cauchy theorem for complex functions we have that $\int_\Gamma f\left(x(\lambda 0\right) d\lambda = 0 = f(y)$. Hence $f(y) = 0$ for all $f \in \mathcal{A}$ thus $y = 0$.

Corollary 11.1.2 (Integral representation)

$$x(\lambda) = \frac{1}{2\pi i}\int_\Gamma \frac{x(\xi)}{\xi - \lambda}. \tag{11.9}$$

Note that we also have $f\left(x(\lambda)\right)' = f\left(x'(\lambda)\right)$ and the Taylor expansion is valid:

$$x(\lambda) = x(\lambda_0) + (\lambda-\lambda_0)x'(\lambda_0) + \frac{(\lambda-\lambda_0)^2}{2!}x''(\lambda_0) + \cdots \tag{11.10}$$

and the radious of convergence is the distance to the closest singular point of $x(\lambda)$.

Example: $(e-\lambda x)^{-1}=\sum_0^\infty \lambda^n x^n$, it converges for small λ but then convergence is extended to the first singularity. The radius of convergence is $R = 1/\lim_n \|x^n\|^{1/n}$. To see this first note that the above limit exists: let $a_n = \|x^n\|$ then $a_{n+m} \le a_n \cdot a_m$. Thus, $a_{mk+l} \le a_k^m \cdot a_l$. So $\limsup a_n^{1/n} \le a_k^{1/k}$. Taking now $\liminf$ both sides we are done.

Theorem 11.1.3 (Liouville) *Let $x(\lambda)$ be an analytic function for all $\lambda \in \mathbb{C}$ which is assumed to be uniformelly bounded, that is, $\|x(\lambda)\| \le C$. Then $x(\lambda)$ is constant.*

Proof: From the known Liouville theorem for analytic functions it follows that if $f \in \mathcal{A}^*$ the function $f\,(x(\lambda_0)$ is constant. Lets say that $f\,(x(\lambda_0) = c_f$. So, fixing any λ_0 we have that $f\,(x(\lambda)) = f\,(x(\lambda_0))$ for every f. Thus $x(\lambda)$ is constant. $\square$

Theorem 11.1.4 *The spectrum of every $x \in \mathcal{A}$ is not empty.*

Recall here that the spectrum $\sigma(x)$ is the set of $\lambda \in \mathbb{C}$ such that $(x-\lambda e)$ is not invertible.

Proof: If $\sigma(x) = \emptyset$ then for every $\lambda \in \mathbb{C}$

$$\|(x-\lambda e)^{-1}\| = |\lambda^{-1}| \| \left(e - \frac{x}{\lambda}\right)^{-1} \|. \tag{11.11}$$

But,

$$\left(e - \frac{x}{\lambda}\right)^{-1} \to e \tag{11.12}$$

as λ tends to infinity, and this means that $\| \left(e-\frac{x}{\lambda}\right)^{-1} \|$ is bounded. Thus $(x-\lambda e)^{-1}$ is a constant and sending λ to infinity we get $(x-\lambda e)^{-1} = 0$ which contadicts the invertibitily of $x-\lambda e$. $\square$

Corollary 11.1.5 (Gelfand-Mazur) *If a Banach algebra $\mathcal{A}$ is a field then $\mathcal{A} = \mathbb{C}$.*

Indeed, $\sigma(x) \neq$ Let $\lambda \in \sigma(x)$. Then $x - \lambda e$ is not invertible hence it equals zero as $\mathcal{A}$ is assumed to be a field. Hence we proved that for every $x \in \mathcal{A}$ there exists $\lambda \in \mathbb{C}$ such that $x = \lambda e_0$.

Corollary 11.1.6 *For any maximal ideal $M \subseteq \mathcal{A}$ there is a natural (algebraic) isomorphism $f_M : \mathcal{A}/M \simeq \mathbb{C}$.*

This implies that $\mathrm{codim} M = 1$ and thus M is a hyperplane. Also f_M is a multiplicative map, meaning that f_M is a multiplicative linear functional:

$$f_M(x \cdot y) = f_M(x) \cdot f_M(y) \tag{11.13}$$

$$f_M(e) = 1 \tag{11.14}$$

Corollary 11.1.7 *If $T : \mathcal{A} \to \mathbb{C}$ is an algebraic homomorphism onto $\mathbb{C}$ then $\mathrm{Ker} T = M$ is a maximal ideal.*

Now for such homomorphism T_M (with $\mathrm{Ker} T_M = M$ a maximal ideal) we define $x(M) = T_M(x) \in \mathbb{C}$.Thus $x(M)$ satisfies the following properties.

(i) $(x_1 + x_2)(M) = x_1(M) + x_2(M)$

(ii) $x_1 x_2(M) = x_1(M) x_2(M)$ and $e(M) = 1$

(iii) $x \in M$ iff $x(M) = 0$ and $M_1 \neq M_2$ then there exists x such that $x(M_1) \neq x(M_2)$.

Corollary 11.1.8 *(a) There exists x^{-1} iff $x(M) \neq 0$ for every maximal ideal M.*

(b) $\sigma(x) = \{x(M) : \ M \text{ is a maximal ideal}\}$

(iv) $|x(M)| \leq \|x\|$ and the norm of any multiplicative functional f_M equals one: $\|f_M\| = 1$.

Indeed, $\|e + x\| \geq 1$ for all $x \in M$ (otherwise x is invertible and does not belong to M) and

$$\|\lambda e + x\| \geq |\lambda| \tag{11.15}$$

for every $x \in M$. Thus, $f_M(\lambda e + x) = \lambda$ and $\|f_M\| = 1$.

Moreover in the opposite direction, $f(\lambda e + x) = \lambda$ is a definition of a linear functional f_M such that $\mathrm{Ker} F_M = M$ and it is multiplicative. So this is the construction of a multiplicative map.

Let W be the Wiener space that consists of elements of the form $x(t) = \sum_{-\infty}^{\infty} a_n e^{int}$ where the series is absolutely convergent, i.e., $\sum |a_n| < \infty$.

Theorem 11.1.9 (The Weiner Theorem) *If $x \in W$ then $1/x(t)$ is an absolutely convergent series with*

$$\frac{1}{x(t)} = \sum_{-\infty}^{\infty}, \tag{11.16}$$

and $\sum |b_n| < \infty$.

Proof: We first describe a maximal ideal M of W. Let $e^{it}(M) = a$. Then $e^{-it}(M) = a^{-1}$ and $\|e^{it}\|_W = \|e^{-it}\|_W = 1$. So $|a| \leq 1$ and $|a^{-1}| \leq 1$ which gives $|a| = 1$ and there exists t_0 such that $a = e^{it_0}$. By linearity and multiplicativity we get that

$$\left(\sum_{-n}^{m} c_k e^{ikt}\right)(M) = \sum_{-n}^{m} c_k e^{ikt_0}, \tag{11.17}$$

for all n, m, c_k. Thus, by continuity, for any $z \in W$ $x(M) = z(t_0)$. Therefore $M = \{z \in W \mid z(t_0) = 0\}$. Hence, $x(t)$ does not belong to any maximal ideal and there exists $x^{-1} \in W$. $\square$

Let now $\mathcal{A}$ be a Banach algebra of functions $f(\lambda)$ which are analytical for $|\lambda| < 1$ and continuous on $|\lambda| < 1$. Let

$$\|f\|_{\mathcal{A}} = \max_{|\lambda|<1} |f(\lambda)|. \tag{11.18}$$

We want to describe the set $\mathcal{M}$ of maximal ideals of $\mathcal{A}$. Let $M \in \mathcal{M}$ and $x(\lambda) = z$ be a generator function. Let $z(M) = z_0$. Since $\|\lambda\|_{\mathcal{A}} = 1$, $|\lambda_0| \leq 1$. Then $z^n(M) = \lambda_0^n$ and $x(\lambda)|_M = x(\lambda_0)$ for any $x \in \mathcal{A}$. Thus,

$$\mathcal{M} = \{z \mid |\lambda| \leq 1\}. \tag{11.19}$$

We study now the space $\mathcal{M}(\mathcal{A})$ of maximal ideals of $\mathcal{A}$. We equip $\mathcal{M}(\mathcal{A})$ with the w^*-topology (remember that $\mathcal{M} \subseteq \mathcal{D}(\mathcal{A}^*)$ the unit ball of $\mathcal{A}^*$). Also note that $x(M)$ are continuous functions on $\mathcal{M}$ (by the definition of the w^*-topology) and $(\mathcal{M}, w^*)$ is a Hausdorff space.

Theorem 11.1.10 *$(\mathcal{M}(\mathcal{A}), w^*)$ is compact.*

Proof: For every $x \in \mathcal{A}$ consider $Q(x) = \{z \in \mathbb{C} \mid |z| \leq \|x\|\}$. Clearly

$$\mathcal{M}(\mathcal{A}) \subseteq \prod_{x \in \mathcal{A}} Q(x) = K \tag{11.20}$$

is compact in the w^*-topology on $\mathcal{M}(\mathcal{A})$. Moreover,

$$\begin{aligned}\mathcal{M}(\mathcal{A}) = \quad & \{ \quad \overline{x} \in K \mid \pi_{xy}(\overline{x}) = \pi_x(\overline{x}) \cdot \pi_y(\overline{x}x);\ \pi_{ax+by}(\overline{x}) = \\ & = \quad a\pi_x(\overline{x}) + b\pi_y(\overline{x});\ \pi_e(\overline{x}) = 1\}\end{aligned}$$

where $\pi_z(\overline{x})$ is a z-coordinate of $\overline{x} = \prod_{x \in \mathcal{A}} Q(x)$.

The conditions that define $\mathcal{M}(\mathcal{A})$ are all closed conditions in the w-topology. So, $\mathcal{M}(\mathcal{A})$ is a closed subset of a compact set and hence compact in itself. □

Corollary 11.1.11 *If $(\mathcal{M}, \mathcal{T})$ is compact in some topology $\mathcal{T}$ and $x(M)$ is continuous in $\mathcal{T}$ for every $x \in \mathcal{A}$ then $\mathcal{T} = w^*$.*

Proof: Consider the map $Id : (\mathcal{M}, \mathcal{T}) \mapsto (\mathcal{M}, w^*)$. Any basic neighborhood in w^* is also a neighborhood in $\mathcal{T}$-topology because $x(M)$ is continuous, and by the Hausdorff theorem (since $\mathcal{T}$ is compact topology on $\mathcal{M}$ and w^* is a Hausdorff topology) it follows that Id is a homomorphism. □

Exercises. 1. $W:\ \mathcal{M}(W) = \mathbb{S}^1$ (with the natural topology).

2. $\mathcal{M}\left(C(S)\right) = S$ for any compact metric space S.

3. $\mathcal{M}(\mathcal{A}) = \mathcal{D}$ (the unit disk)

So, in these examples, $\mathcal{M}$ is a natural domain of functions.

11.2 Radicals

Definition 11.2.1 *Consider the homomorphism*

$$T:\ \mathcal{A} \mapsto \hat{\mathcal{A}} = \{x(\mathcal{M}) \mid \forall x \in \mathcal{A}\} \hookrightarrow C(\mathcal{M}) \tag{11.21}$$

The set

$$\mathbb{R} = \ker(T) = \{x \mid x(\mathcal{M}) = 0\} \tag{11.22}$$

is called the radical of the algebra $\mathcal{A}$.

Clearly, the radical is an ideal. Also $x \in ker(T)$ iff there is $(e - \lambda x)^{-1}$ for every $\lambda \in \mathbb{C}$ (meaning that $(e - \lambda x)^{-1}$ is an entire function). Then the corresponding series at $\lambda = 0$ converges in all $\mathbb{C}$ and the radius of convergence equals infinity. Thus we have arrived at the following theorem:

Theorem 11.2.2 *$x \in \mathbb{R}$ iff $\lim \|x^n\|^{1/n} = 0$.*

Such an x is also called *a generalized nilpotent.*

Example. Consider the space $\widetilde{L}_1[0,1] = L_1[0,1] \oplus \lambda e$. Let $x_0(t) = 1$. It is a generator of this Banach algebra and $(x_0)^n = \frac{t^{n-1}}{(n-1)!}$. Thus, $\|x_0^n\| = \frac{1}{n!}$, $\|x_0\|^{1/n} \to 0$ and for every $x \in L_1[0,1]$ is in the radical. So, $\mathbb{R}(= L_1)$ is the only maximal ideal. Existence of (non-trivial) radical is a "bad" property for the general theory but has an interesting consequence for the theory of integral equations:

consider the equation

$$u(t) - \lambda \int_0^t k(t-\tau)u(\tau)d\tau = f(t) \tag{11.23}$$

where f and k are in $L_1[0,1]$. Then for every λ there exists solution $u \in L_1[0,1]$. Indeed, the equation (11.23) can be rewritten as

$$u - \lambda(k * u) = f \text{ or } (e - \lambda k) * u = f. \tag{11.24}$$

Since k is in the radical there exists $(e - \lambda k)^{-1}$ and $u = (e - \lambda k)^{-1} * f$ (so it is a Volterra equation and $\sigma(k) = 0$).

Theorem 11.2.3 *Let $\|x\|_c = a = \sup\{|x(M) \mid M \in \mathcal{M}\}$. Then $a = \lim \|x^n\|^{1/n}$.*

Proof: $(x - \lambda e)$ is invertible for all $|\lambda| > a$ means that there exists $(e - \mu x)^{-1}$ and it is analytical in $|\mu| < \frac{1}{a}$. Thus $a \geq \lim \|x^n\|^{1/n}$ From the other side though, $\sup |x^n(M)| = a^n$ and $\|x^n\|^{1/n} \geq a$. We see that the limit $\lim \|x^n\|^{1/n}$ exists and equals a.

Next we consider Banach algebras with radical equal to zero, called *semisimple* Banach algebras.

Theorem 11.2.4 *Every algebraic isomorphism $T : \mathcal{A}_1 \mapsto \mathcal{A}_2$ between two Banach algebras $\mathcal{A}_1$ and $\mathcal{A}_2$ is also a topological isomorphism.*

We will actually prove a stronger statement:

Lemma 11.2.5 *If $A_1 \subseteq A_2$ is a subalgebra (both algebras are assumed with zero radical),and the sets of maximal ideals satisfy $\mathcal{M}(\mathcal{A}_1) = \mathcal{M}(\mathcal{A}_2)$ then $x_i \to x$ (in A_1) implies $x_i \to x$ (in A_2).*

Proof: Introduce a new norm on $\mathcal{A}_1$ by,

$$\|x\| \max\{\|x\|_1, \|x\|_2\}. \tag{11.25}$$

If x_i is a Cauchy sequence in $\|\cdot\|$ then it is also Cauchy in both norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Now, both spaces are complete, thus $x_i \to x$ (in $\|\cdot\|_1$) and $x_i \to y$ (in $\|\cdot\|_2$). But for every maximal ideal M, $x_i(M) \to a(M) = x(M) = y(M)$. This implies $x = y$. Then $x_i \to x$ in $\|\cdot\|$ as well, which proves completence of $(\mathcal{A}_1, \|\cdot\|_1)$. It follows from Banach Theorem that $\|\cdot\|$ is equivalent to $\|\cdot\|_1$ Obviously, convergence in $\|\cdot\|$ implies convergence in $\|\cdot\|_2$. □

Corollary 11.2.6 *Automorphisms of Banach algebras without radical are continuous.*

Consider now an algebra of functions $\hat{\mathcal{A}}\,(=\{x(M)\}_{x\in\mathcal{A}})$.

Problem. When $\hat{\mathcal{A}}$ is dense in $C(\mathcal{M})$? (i.e. when is it true that $\overline{\hat{\mathcal{A}}}^c = C(\mathcal{M})$?).

The example of the algebra of analytical functions on the disk $\mathcal{D}$ shows that some conditions are necessary.

Theorem 11.2.7 *If $\hat{\mathcal{A}}$ is symmetric, then $\hat{\mathcal{A}}$ is dense in $C(\mathcal{M})$.*

This is a form of the Weierstrass theorem. The straightforward consequences of this theorem are:

1. Weierstrass theorems on the density of polynomials and trigonometric polynomials.
2. Let S and T be compact metric spaces. The functions of the form $\sum_1^n x_i(s)y_i(t)$ are dense in $C(S\times T)$.
3. If $\hat{\mathcal{A}}$ is symmetric and $\|x^2\| = \|x\|^2$ then $\mathcal{A} = C(\mathcal{M})$ [Indeed, $\max|x(M)| = \|x\|$ and $\hat{\mathcal{A}}$ is dense in $C(\mathcal{M})$].

11.3 Involutions

Definition 11.3.1 *We call involution a map $x \mapsto x^*$ with the properties:*

a. $(x^*)^* = x$
b. $(\lambda x + \mu y)^* = \overline{\lambda} x^* + \overline{\mu} y^*$ *(anti-linearity)*
c. $(xy)^* = y^* x^*$.

The main example of involution for us is the dual operators for an algebra of operators in a Hilbert space.

We call x^* the conjugate of x.

Examples. 1. $C(S)$: $x(s)^* = \overline{x(s)}$

2. Le $\mathcal{A}$ be the analytical functions on $\mathcal{D}$. Then $x(\xi) \to \overline{x(\overline{\xi})}$ (meaning $\sum a_n\xi^n \to \sum \overline{a_n}\xi^n$) is an involution.

3. Let Q_0 be the set of pairwise commutative normal operators from a Hilbert space H to itself. Then consider the norm-closure of the algebraic span: $Q = \overline{\text{algspan}(Q_0; i; Q_0^*)}$. We obtain a closed subalgebra of $L(H \to H)$ such that for every $T \in Q$ we also have $T^* \in Q$. So, Q is a commutative Banach algebra with involution.

A few more definitions follow: if $x = x^*$ then x is called *self-adjoint element*. For every $x \in \mathcal{A}$, $x = y + iz$ where y and z are self-adjoint and this decomposition is unique ($\frac{x+x^*}{2}$ and $\frac{x-x^*}{2i} = z$)

Note that x is invertible iff x^* is invertible ($xx^{-1} = e$ iff $((x^{-1})^*x^* = e)$.

An algebra $\mathcal{A}$ with an involution $*$ is called symmetric iff $x^*(M) = \overline{x(M)}$. Of course if $(A, *)$ is symmetric then $\hat{\mathcal{A}}$ is symmetric.

We add now an other property of the convolution:

Definition 11.3.2 *A Banach algebra $\mathcal{A}$ with involution $*$ is called a C^*-algebra iff*

$$\|xx^*\| = \|x\| \cdot \|x^*\|. \tag{11.26}$$

Theorem 11.3.3 (Gelfand-Naimark) *If $\mathcal{A}$ is a commutative C^*-algebra then $\mathcal{A} = C(\mathcal{M})$.*

Proof: We will prove this by showing that $\|x^2\| = \|x\|^2$ and $\mathcal{A}$ is symmetric. Now,

$$\|(xx^*)^2\| = \|xx^*(xx^*)^*\| = \|xx^*\|^2 = \|x\|^2\|x^*\|^2. \tag{11.27}$$

From the other side, using commutativity we have that the same expression is equal to

$$\|x^2(x^*)^2\| = \|x^2\|\|(x^*)\|^2. \tag{11.28}$$

This implies that we have equality, that is,

$$\|x^2\| = \|x\|^2 \quad \text{forevery} x \in \mathcal{A}. \tag{11.29}$$

Now assume that $\mathcal{A}$ is not symmetric. Then there is x_0 and M_0, a maximal ideal, so that $x_0^*(M_0) \neq \overline{x_0 M_0}$. Then $\mathrm{Im}(x_0 + x_0^*)(M_0) \neq 0$ and there is an element

$$h = \frac{x_0 + x_0^* - Re[(x_0 + x_0^*)(M_0)]e}{\mathrm{Im}(x_0 + x_0^*)(M_0)}. \tag{11.30}$$

We see that $h = h^*$ and $h(M_0) = i$. Som $(h - ie)$ is not invertible. Then it follows that $(h = ie)^* = h + ie$ is not invertible meaning that there is a maximal ideal M_1 such that $h(M_1) = -i$. Thus for every $t > 0$

$$(h + tie)(M_0) = (1 + t)i \text{ and}(h - tie)(M_1) = -(1 + t)i. \tag{11.31}$$

Therefore, $\|h \pm tie\| \geq 1 + t$. Finally

$$\|h^2 + t^2 e\| = \|h + tie\| \cdot \|h - tie\| \geq (1 + t)^2 \tag{11.32}$$

and

$$\|h^2 + t^2 e\| \leq \|^2\| + t^2. \tag{11.33}$$

This is a contradiction because it is wrong that $C + t^2 \geq (1 + t)^2$ for a large t no matter what is the constant C. □

This theorem implies a spectral decomposition for a family of pairwise commutative operators. But before we show this, let us establish a few additional properties and examples of symmetric algebras with involution.

Theorem 11.3.4 *$(\mathcal{A}, *)$ is symmetric iff $(e+xx^*)$ is invertible for every $x \in \mathcal{A}$.*

Proof: It is obvious that $(\mathcal{A}, *)$ being symmetric implies that $(e + xx^*)$ is invertible. For every M,

$$(e + xx^*)(M) = 1 + |x(M)|^2 > 0. \tag{11.34}$$

In the opposite direction, let us show that if $x = x^*$ then $x(M) \in \mathbb{R}$. It is enough to prove that for any $a \in \mathbb{R}$ and $b \in \mathbb{R} \setminus \{0\}$ the inverse of $(x - (a + ib)e)$ exists. Consider,

$$(x - (a + ib)e)\,(x - (a - ib)e) = (x - ae)^2 + b^2 e = b^2[e + zz^*] \tag{11.35}$$

and hence it is invertible. It follows that $(x - (a + ib)e)$ is invertible. □

Example. Let $\mathcal{A}$ be the algebra of bounded functions on a set S with the uniform norm: $\|f\| = \sup_s |f(s)|$. Let $x, \overline{x} \in \mathcal{A}$. Then define the involution $x^* = \overline{x}$. This is a symmetric involution. Indeed, for every $x \in \mathcal{A}$, $\frac{1}{1+|x|^2} \in \mathcal{A}$: let $\|x\| = a$. Then

$$\begin{aligned} \frac{1}{1+|x|^2} &= \frac{1}{a^2+1} \frac{1}{1 - \frac{a^2-|x|^2}{a^2+1}} \\ &= \frac{1}{a^2+1} \sum_0^\infty \left(\frac{a^2-|x|^2}{a^2+1} \right)^2 \end{aligned}$$

which converges uniformly in $\mathcal{A}$. Now, since $\|x^2\| = \|x\|^2$ we have that $\mathcal{A} = C(\mathcal{M})$.

A few important concrete examples are:

1. All bounded continuous functions on $(-\infty, \infty)$.

2. B-almost periodic functions on $(-\infty, \infty)$. For example: $\{e^{ita}\}_a$ [$e^{ita} + e^{itb}$ is almost periodic].

Next we present a criterion for an algebra $\mathcal{A}$ being semisimple (i.e. with zero radical). We call a linear functional f positive iff $f(xx^*) \geq 0$ for every $x \in \mathcal{A}$ and we say that the involution is essential if for every non-zero x there exists $f \geq 0$ such that $f(xx^*) > 0$.

Theorem 11.3.5 *If the involution $*$ is essential then $\mathcal{A}$ is semisimple.*

Proof: Take a positive linear functional f such that for a given x we have $f(xx^*) > 0$.Then

(11.36) $0 \leq f[(x+\lambda e)(x+\lambda e)^*] = f(xx^*) + \lambda f(x^*) + \overline{\lambda} f(x) + |\lambda| f(e)$,

for every $\lambda \in \mathbb{C}$. Take $\lambda \in \mathbb{R}$ and then $\lambda \in i\mathbb{R}$; we obtain $f(x^*) = \overline{f(x)}$ [because $\lambda \in \mathbb{R}$ gives $\mathrm{Im} f(x^*) = -\mathrm{Im} f(x)$ and λ purely imaginary implies $Ref(x^*) = Ref(x)$]. Put now $\lambda = f(x)t$ for $t \in \mathbb{R}$. Then (for every $t \in \mathbb{R}$)

(11.37) $$t^2|f(x)|^2 f(e) + 2t|f(x)|^2 + f(xx^*) \geq 0,$$

which means

(11.38) $$f(xx^*)f(e)|f(x)|^2 \geq |f(x)|^4.$$

Thus, $|f|^2 \leq f(e)f(xx^2)$. Let $z = xx^2$. Then similarly $|f(z)|^2 \leq f(e)f(z^2)$,

$$|f(x)| \leq f(e)^{1/2} f(z)^{1/2}$$

$$\begin{aligned} &\leq \quad f(e)^{\frac{1}{2}+\frac{1}{4}} f(z^2)^{\frac{1}{4}} \\ &\vdots \\ &\to \quad f(e) \max_M |(xx^*)(M)|^{1/2}, \end{aligned}$$

as n tends to infinity. Using this inequality for x instead of $z = xx^*$ and a special positive functional f such that $f(z) > 0$ we get

$$0 < f(z) \leq f(e) \max_M |x(M)| \cdot |x^*(M)|. \tag{11.39}$$

This implies that $x(M) \neq 0$ for some M and x outside the radical. Thus the radical is zero. □

Now we return to the spectral theory of a family of commutative normal operators Q and H. Let $\mathcal{A} = \overline{\text{algspan}}[Q, e, Q^*] \hookrightarrow L(H)$ where the closure is in the strong operator topology. Clearly $\mathcal{A}$ has involution $\mathcal{A} \in T \mapsto T^* \in \mathcal{A}$ and $\mathcal{A}$ is a C^*-algebra. So, $A = C(\mathcal{M})$ which means that there is a correspondence (map) which is algebraic homomorphism:

$$C(\mathcal{M}) \longrightarrow \mathcal{A} \hookrightarrow L(H \to H). \tag{11.40}$$

For every $f \in C(\mathcal{M})$ corresponds a $T_f \in A$, $T_f : H \to H$ with the following properties:

1. If f is real-valued then T_f is self-adjoint (since involution is symmetric: $\overline{f} \mapsto T_f^* = T_{\overline{f}}$).

2. $f \geq 0$ implies that $T_f \geq 0$ (because $f = \sqrt{f}\sqrt{f}$ and thus, $T_f = T^2_{\sqrt{f}}$).

3. $(T_f x, y)$ is a three dimensional functional; in particular it is linear on $C(\mathcal{M})$ for fixed x and y. Therefore there is a measure $\mu_{x,y}$ on $\mathcal{M}$ such that:

$$(T_f x, y) = \int_{\mathcal{M}} f d\mu_{x,y} \text{ for any } f \in C(\mathcal{M}). \tag{11.41}$$

Our purpose now is to extend this correspondence to the class of Baire functions on $\mathcal{M}$. In particular:

4. Define for the characteristic function χ of a "good" subset (Borel subset) an operator by

$$(T_\chi x, y) = \int_{\mathcal{M}} \chi d\mu_{x,y} \; [= \chi(\mu_{x,y}, \chi \in C^{**}]. \tag{11.42}$$

Let us check the properties of this correspondence $\chi \mapsto T_\chi : H \to H$.

5. By multiplicativity: ($\phi f \mapsto T_{\phi f} = T_\phi \cdot T_f$) we have

$$f d\mu_{x,y} d\mu_{T_f x,y} = d\mu_{x,T_f^* y}, \tag{11.43}$$

since $\int \phi f d\mu_{x,y} = \int \phi d\mu_{T_f x,y}$ and so on...

6. Using (11.42) and (11.43) we have that

$$\chi d\mu_{x,y} = d\mu_{T_\chi x,y} \text{ and } \chi_1\chi_2 d\mu_{x,y} = \chi_1 d\mu_{T_{\chi_2}x,y} = d\mu_{T_{\chi_1}T_{\chi_2}x,y}. \tag{11.44}$$

So,

$$T_{\chi_1\cdot\chi_2} = T_{\chi_1} \cdot T_{\chi_2}, \tag{11.45}$$

thus our extension on Baire functions is multiplicative.

Lemma 11.3.6 *If $S \in L(H \to H)$ is such that for every $f \in C(\mathcal{M})$ $ST_f = T_f S$ then T_χ commutes with S for any Baire function χ.*

Proof: $(T_f Sx, y) = (T_f x, S^* y)$ which implies

$$\int f d\mu_{SX,y} = \int f d\mu_{x,S^* y} \tag{11.46}$$

hence

$$d\mu_{Sx,y} = d\mu_{x,S^* y}. \tag{11.47}$$

Then

$$\chi d\mu_{Sx,y} = \chi d\mu_{x,S^* y}. \tag{11.48}$$

Therefore $(T_\chi Sx, y) = (T_\chi x, S^* y)$ and $T_\chi S = ST_\chi$. □

Lemma 11.3.7 *If $T_{\overline{f}} = T_f^*$ for every $f \in C(\mathcal{M})$ then $T_{\overline{\chi}} = T_\chi^*$ where χ is any Baire function.*

Proof: Note that $d\mu_{x,y} = d\overline{\mu_{y,x}}$. Indeed, $\int f d\mu_{x,y} = (T_f x, y) = \overline{(T_f^* x, y)} = \overline{(T_{\overline{f}} y, x)} = \int f d\overline{\mu_{y,x}}$. Then

$$\int \overline{\chi} d\mu_{x,y} = \overline{\int \chi d\mu_{y,x}}, \tag{11.49}$$

meaning

$$(T_{\overline{\chi}} x, y) = \overline{(T_\chi y, x)} = (x, T_\chi y), \tag{11.50}$$

so, $T_{\overline{\chi}} = T_\chi^*$ □

Corollaries. 1. If χ_S is a characteristic function of $S \subseteq \mathcal{M}$ then $T_{\chi_S} = E_S$ is an orthoprojection (it is self adjoint because χ_S is real-valued function and $E_S^2 = E_S$, since $\chi_S^2 = \chi_S$). Also, E_S commutes with all operators from the algebra $\mathcal{A}$ and therefore is the orthoprojection on an *invariant* subspace for all these operators.

2. The individual spectral theorem: Let $T = T^*$ and let $\mathcal{A}$ be an algebraic envelope of T, closed in the strong topology. Let $\hat{T}$ be the corresponding function in $C(\mathcal{M})$. Let $-I \leq T \leq I$. Then $-1 \leq \hat{T}(M) \leq 1$. Take an ε-partition $-1 = \lambda_0 < \lambda_1 < \ldots < \lambda_n = 1$ and let $\hat{E}_\lambda =$?????????????????????????????????? Then,

$$\|\hat{T} - \sum_1^n \lambda_i' \left(\hat{E}_{\lambda_i} - \hat{E}_{\lambda_{i-1}}\right) \|_{C(\mathcal{M})} \leq \varepsilon, \tag{11.51}$$

for functions in C^{**}. This implies that

$$\|T - \sum \lambda_i \left(E_{\lambda_i} - E_{\lambda_{i-1}}\right) \|_{op} \leq \varepsilon \tag{11.52}$$

and

$$T = \int \lambda dE_\lambda. \tag{11.53}$$

Chapter 12

Unbounded self-adjoint and symmetric operators in H

LET A BE A LINEAR operation defined on some (linear) subset Dom$A$, the domain of A, of a Hilbert space H. Then the pair (A; Dom$A$) is called an operator. We always assume that $\overline{\mathrm{Dom} A} = H$ and we use write $\mathcal{D}_A$ instead of Dom$A$.

Let, for given $y \in H$, y^* be such that for every $x \in \mathcal{D}_A$

$$(Ax, y) = (x, y^*). \tag{12.1}$$

Then we then write $y^* = A^* y$. The condition $\overline{\mathcal{D}_A} = H$ guarantees that if such a y^* exists then is is unique: if $(Ax, y) = (x, y_1^*) = (x, y_2^*)$ then $(x, y_1^* - y_2^*) = 0$ for all $x \in \mathcal{D}_A$ hence $y_1^* = y_2^*$.

The operator A^* has a natural domain, that is, the set $\mathcal{D}_{A^*}$ containing all y for which the y^* exists.

Examples. 1. Let $H = L_2[0, 1]$ and $Ax = -x''$ with domain $\mathcal{D}_A = \{x \in H \ : \ \exists x'' \in H \text{ and } x(o) = x'(0) = 0\}$.

2. The same space and operation as above but with domain $\mathcal{D}_A = \{x \in H \ : \ x'' \in H \text{ and } x(0) = 0\}$. We will see that this is a very different operator than the previous one.

Two other examples would be if we start with the Hilbert space $L_2[0, \infty)$.

The notion of closed graph operator (or shortly, just "closed operator") will play an important role in what follows.

Definition 12.0.8 *The operator $(A; \mathcal{D}_A)$ is called a closed operator iff for any $x_n \in \mathcal{D}_A$ such that $x_n \to x \in H$ and $Ax_n \to y \in H$ it follows that $x \in \mathcal{D}_A$ and $Ax = y$ (that is, $y \in$ ImA).*

Some first properties of the dual operator A^* are the following:

1. A^* is closed operator: it is obvious that if $y_n \to y$, $y_n^* \to y^*$ and $(Ax, y_n) = (x, y_n^*)$ for every $x \in \mathcal{D}_A$ then $(Ax, y) = (x, y^*)$ which means that $y \in \mathcal{D}_{A^*}$ and $A^*y = y^*$.

2. We say that $A_1 \subseteq A_2$ if $\mathcal{D}_{A_1} \subseteq \mathcal{D}_{A_2}$ and $A_2x = A_1x$ for every $x \in \mathcal{D}_{A_1}$ (that is, $A_2|_{\mathcal{D}_{A_1}} = A_1$). It is clear that $A_1 \subseteq A_2$ implies $A_1^* \supseteq A_2^*$.

3. For any operator $(A, \mathcal{D}_A)$ we define $\overline{A}$ the closure of A (if it exists) to be the operator with the graph $gr\overline{A} = \overline{grA}$. Of course, it may happen that the set $\overline{grA} \subseteq H \times H$ is not the graph of any operator. We say that A permits the closure if the closed operator $\overline{A}$ exists.

Now if A permits the closure then $\overline{A}^* = A^*$ (obvious) and if $(A^*)^*$ exists (which means that $\mathcal{D}_{A^*}$ is dense in H) then

$$\overline{A} \subseteq A^{**}. \tag{12.2}$$

We call A a *symmetric* operator if $\overline{\mathcal{D}_A} = H$ and for every $x \in \mathcal{D}_A$

$$(Ax, y) = (x, Ay). \tag{12.3}$$

Therefore $A \subseteq A^*$ (and it is, in fact, the definition of symmetry of A).

Observe that any symmetric $B \supseteq A$ (symmetric extension) satisfies $B \subseteq A^*$ (because $B \subseteq B^* \subseteq A^*$). So, all symmetric extensions B of a symmetric operator A stay between A and A^*, that is, $A \subseteq B \subseteq A^*$.

We call A a self adjoint operator if $A = A^*$. Note that if $\overline{\mathrm{Im}A} = E \neq H$ then $\ker A^* \neq 0$ (meaning that 0 is an eigenvalue of A^*): indeed, $(Ax, y) = (x, A^*y)$ and $y \perp E$ means $A^*y = 0$.

Theorem 12.0.9 *Let A be a symmetric operator with $\mathcal{D}_A = H$. Then A is a bounded operator.*

Proof: For every $x \in H$, $(Ax, y) = (x, Ay)$. Consider a family $\{Ax\}_{x \in B_H}$. Then $\{|(Ax, y)|\}_{x \in B_H}$ is a bounded set for every y because $|(Ax, y)| = |(x, Ay)| \leq \|x\| \cdot \|Ay\|$. By the Banach-Steinhaus Theorem it follows that $\{Ax\}_{x \in B_H}$ is bounded meaning that A is a bounded operator.

Before we continue to develop the general theory let us consider a few examples.

Examples. 1a. Let A_1 be an operator on $L_2[0, 1]$ defined by the operation $Ax = i\frac{d}{dt}x$. Let

$$\mathrm{Dom}A = \{x \in L_2[0, 1] \ : \ x \text{ is smooth function}; x(0) = x(1) = 0\}.$$

Then,

$$\begin{aligned}
(Ax,y) &= (x,y^*) = \int_0^1 x\overline{y^*}d\tau \\
&= x\left(\int_0^t \overline{y^*}d\tau + c\right)|_{t=0}^1 - \int x'\left(\int_0^t \overline{y}^*d\tau + c\right)dt \\
&= \int_0^1 (ix)\left(\int_0^t \overline{(-iy^*)}d\tau + c\right).
\end{aligned}$$

So, since $\{Ax\}_{x\in \text{Dom}A}$ is dense in L_2, $y = \int_0^t \overline{-iy^*} + c$. Therefore there is $y' \in L_2$ and $y' = -iy^*$. Thus $y^*A^*y = iy'$ and $\text{Dom}_{A^*} = \{y \in H \ : \ y' \in L_2\}$ (note the lack of boundary conditions). Hence A is symmetric and $A \subseteq A^*$. A is not closed (since Dom$A$ was chosen to be "too small"). But A admits a closure and $A_1 = \overline{A}$ i.e. $A_1x = ix$ and

(12.4) $$\text{Dom}_{A_1} = \{x \in L_2[0,1],\ x' \in L_2 \text{ and } x(0) = x(1) = 0\}.$$

1b. Let A_2 be an operator defined on $H = L_2[0,1]$ and

(12.5) $$\text{Dom}A_2 = \{x \in H \mid x' \in L_2[0,1] \text{ and } x(0) = x(1)\}.$$

Thus, $A_2 \supseteq A_1$. Therefore $y^* = iy'$ ($A_2^* \subseteq A_1^*$).

$$\begin{aligned}
(Ax,y) &= \int_0^1 i\frac{d}{dt}x(t)\overline{y(t)}dt \\
&= i[x(1)\overline{y(1)} - x(0)\overline{y(0)}] + \int_0^1 x\overline{iy'}dt.
\end{aligned}$$

Now, the quantity $K = ix(1)\overline{[y(1) - y(0)]}$ must be zero because $(Ax,y) = (x,y^*)$ and there exists $x_n \to 0$ (in L_2) but $x(0) = x(1) = 1$. Then $(Ax,y) \not\to 0$ but $(x_n,y^*) \to 0$ a contradiction. So $y(1) = y(0)$ and we see that $A_2^* = A_2$; this operator is self-adjoint.

Let us return to 1a and 1b examples and compute the eigenvalues and eigenfunctions of A^*.

In the 1a example, $A^*x = ix'$ and

$$\mathcal{D}_{A^*} = \{x' \in L_2 \ : \ \text{no other conditions}\},$$

$ix' = \lambda x$. So, $x = e^{\lambda t/i}$ and a non-trivial solution exists for every λ. Moreover $\dim\ker(A^* - \lambda I) = 1$ meaning that $\text{codim}\overline{\text{Im}(A - \overline{\lambda}I)} = 1$. We call this codimension "*index of defect*". It may be shown that it is the same number (for symmetric operators) for λ, $\mathrm{Im}\lambda > 0$ and (probably another) the same for λ, $\mathrm{Im}\lambda < 0$. Thus in this example the indices are $(1,1)$. (Note that $Ax = ix'$ has no eigenvalues because the solution of the equation does not satisfy the conditions $x(0) = x(1) = 0$.)

In the 1b example now, $A^* = A$ and the conditions are $x(0) = x(1)$. This gives $1 = e^{\lambda/i}$ and $\lambda = 2\pi n$, $n = 0, \pm 1, \pm 2, \dots$. So, there are no solutions for $\mathrm{Im}\lambda > 0$ or $\mathrm{Im}\lambda < 0$ and the indices are $(0,0)$.

1c. Let $L_2[0,\infty)$: $Ax = ix'$ and $\mathcal{D}_A = \{x \in L_2 \mid x' \in L_2,\ x(0) = 0\}$. (We start first with smooth functions of finite support and then take closure.) Then the same line of computation as in example 1a implies

(12.6) $$A^*x = ix' \text{ and } \mathcal{D}_{A^*} = \{x \in L_2 \mid x' \in L_2\}.$$

Thus A is symmetric ($A \subseteq A^*$). Computing the eigenvalues of A^* we see that $x = e^{\lambda t/i}$ is an eigenvalue only if $\mathrm{Im}\lambda < 0$ because x must be in $L_2[0,\infty)$. Thus the indices are $(0,1)$.

2. The operation $Ax = -x''$.

2a. In $L_2[0,\infty)$ let

(12.7) $$\mathcal{D}_A = \{x \in L_2 \mid x'' \in L_2 \text{ and } x(0) = x'(0) = 0\}.$$

Again, we start with $\mathcal{D}_A$ which contains only smooth functions of finite support but then take closure. First in the same way as in 1a we show that $y \in \mathcal{D}_{A^*}$ implies that there exists $y'' \in L_2[0,\infty)$. Then we proceed as follows:

$$\begin{aligned}(Ax, y) &= \int_0^\infty (-x'')\overline{y}dt = -x'\overline{y}|_0^\infty + \int_0^\infty x'\overline{y'}dt \\ &= (-x'\overline{y} + x\overline{y'}|_0^\infty + (x, -y'').\end{aligned}$$

(if $y = x$ we see that $(Ax, x) \geq 0$). So, $y^* = -y''$ and

(12.8) $$\mathcal{D}_{A^*} = \{y \in L_2 \mid y'' \in L_2 \text{ no boundary conditions}\}.$$

Call this operator A_1. So, $A_1 \subseteq A_1^*$ and it is a symmetric operator but not self-adjoint.

2b. Consider the same operation in the same space $L_2[0,\infty)$, $Ax = -x''$ but now

(12.9) $$\mathcal{D}_A = \{x \in L_2 \mid x'' \in L_2[0,\infty) \text{ and } x(0) = 0\}.$$

Call the operator we obtain A_2. Obviously $A_1 \subseteq A_2$ and $A_2^* \subseteq A_1^*$. So, $y \in \mathcal{D}_{A_2^*}$ implies that there exists $y'' \in L_2[0,\infty)$. Repeating the above line of computation we have that for $y \in \mathcal{D}_{A_2^*}$

$$(x, y^*) = (Ax, y) = -x'(0)\overline{y(0)} + (x, -y''), \tag{12.10}$$

and we must have $y(0) = 0$ (otherwise we take $x_n \to 0$ in L_2 but $x_n'(0) = 1$ arriving to a contradiction). So, $A_2 = A_2^*$ and A_2 is a self-adjoint extension of A_1. Computing now indices of defect of A_1 and A_2 we have to look for eigenvalues of the dual operators A_1^* and A_2^*:

$$-x'' = \lambda x \tag{12.11}$$

and $x(t) = c_1 e^{i\sqrt{\lambda}t} + c_2 e^{-i\sqrt{\lambda}t}$. However we are looking for solutions $x(t) \in L_2[0,\infty)$. So, if $\mathrm{Im}\lambda \neq 0$ only one of the functions $e^{i\sqrt{\lambda}t}$ or $e^{-i\sqrt{\lambda}t}$ remain. Therefore the indices of A_1 are $(1,1)$. In the case $A_2^* = A_2$ there is another condition $x(0) = 0$ and no such solutions exist. Thus, the index of A_2 is $(0,0)$.

3. Let $L_2(-\infty,\infty)$ and $Ax = tx$, $\mathcal{D}_A = \{x \in L_2 \mid tx \in L_2\}$. Obviously $A = A^*$ and the operator is self-adjoint.

12.1 More Properties Of Operators

We add to Theorem 1 above a few more facts.

Theorem 12.1.1 *If A is a symmetric operator and $\mathrm{Im}A = H$ then A is self-adjoint.*

Proof: Take any $y \in \mathcal{D}_{A^*}$, $A^*y = y$. Since $\mathrm{Im}A = H$ there is $x \in \mathcal{D}_A$ and $Ax = y^*$. Let us show that $x = y$ which would mean $\mathcal{D}_{A^*} = \mathcal{D}_A$ and $A = A^*$. So, for every $z \in \mathcal{D}_A$ we have that

$$(Az, y) = (z, y^*) = (z, Ax) = (Az, x) \tag{12.12}$$

by the assumption. Thus, $y = x$.

Theorem 12.1.2 *If A is a self-adjoint operator and there is a formal inverse A^{-1} (meaning that $\ker A = 0$, i.e. A is one-to-one from $\mathcal{D}_A$ to $\mathrm{Im}A$) then A^{-1} is also self-adjoint.*

Proof: First if $\ker A = 0$ and $A = A^*$ then $\overline{\mathrm{Im} A} = H$. Indeed, if $\overline{\overline{\mathrm{Im} A}} = E \subsetneqq H$ then there is a $y_0 \neq 0$ and such that $A^* y_0 = 0$; but $A^* = A$ and $\ker A \neq 0$. So, $\mathcal{D}_{A^{-1}}$ is dense in H. To describe the dual operator $\left(A^{-1}\right)^*$ we should consider the equation $(A^{-1}x, y) = (x, y^*)$. Let $z = A^{-1}x$. Then $(z, y) = (Az, y^*)$. Since A is self-adjoint, $Ay^* = y$ and $y^* \in \mathcal{D}_A$. Thus, $y \in \mathrm{Im} A = \mathcal{D}_{A^{-1}}$ and $y^* = A^{-1}y$. We see that $(A^{-1}x, y) = (x, A^{-1}y)$. □

Note that the Theorem 3 gives us many examples of self-adjoint unbounded operators. Start with any self-adjoint compact operator A without non-trivial kernel. Then A^{-1} is an unbounded self-adjoint operator.

12.2 The Spectrum $\sigma(A)$

Similarly to the case of the bounded operators we say that $\lambda \in \mathbb{C}$ is a regular point if there exists a bounded operator $(A - \lambda I)^{-1}$. The spectrum $\sigma(A)$ consists of all non-regular points. We devide $\sigma(A)$ in:

(i) the point spectrum $\sigma_p(A)$ of eigenvalues of A, that is, $\lambda \in \sigma_p(A)$ iff there exists an $x \in \mathcal{D}_A$ such that $Ax = \lambda x$ (i.e. $\ker A \neq \emptyset$).

(ii) the continuous spectrum $\sigma_c(A)$ where $\lambda \in \sigma_c(A)$ iff $\lambda \notin \sigma_p(A)$ and $\mathrm{Im}(A - \lambda I)$ is dense in H (but not equal to H). Of course, $A - \lambda I$ is defined on $\mathcal{D}_A$.

(iii) the residue spectrum $\sigma_t(A)$ where $\lambda \in \sigma_r(A)$ iff $A - \lambda I$ is one-to-one, that is, $\lambda \notin \sigma_p(A)$ and $\overline{\mathrm{Im}(A - \lambda I)} \neq H$.

Now let $A \subseteq A^*$ (i.e., A is symmetric). Define $\mathcal{D}_A(\lambda) = \mathrm{Im}(A - \lambda I)$.

(a) $\mathcal{D}_A(\lambda)$ is not dense in h iff $\overline{\lambda} \in \sigma_p(A)$. Indeed, $\mathcal{D}_A^{\perp}$ is the span of eigenvectors of S^* for $\overline{\lambda}$.

(b) Again, $A \subseteq A^*$; then $\lambda \in \sigma_p(A)$ implies that $\lambda \in \mathbb{R}$. and if $\lambda_1 \neq \lambda_2$ are both in $\sigma_p(A)$ the the eigenvectors $Ax_i = \lambda_i x_i$ for $i = 1, 2$, are orthogonal: $(x_1, x_2) = 0$. As for bounded operators,

$$\lambda(x, x) = (Ax, x) = (x, Ax) = \overline{\lambda}(x, x) \quad (x \in \mathcal{D}_A). \tag{12.13}$$

(c) Let now $z = \lambda + i\mu$ for $\lambda, \mu \in \mathbb{R}$ with $\mu \neq 0$, and let A be a closed symmetric operator. Then $\mathcal{D}_A(z)$ is a closed subspace.

Proof of (c). Note that for symmetric A, $A_\lambda = A - \lambda I$ is also symmetric (for $\lambda \in \mathbb{R}$). For $x \in \mathcal{D}_A$ we have

$$\begin{aligned} \|Ax - zx\|^2 &= \|A_\lambda x\|^2 - (A_\lambda x, i\mu x) - (i\mu x, A_\lambda x) + \mu^2(x, x) \\ &\geq \mu^2 \|x\|^2. \end{aligned}$$

So, $(A - xI)^{-1}$ is formally defined ($z \notin \sigma_p(A)$ because $\sigma_p(A) \subseteq \mathbb{R}$) and

$$\|(A - zI)^{-1}|_{\mathcal{D}_A(x)}\| \leq \frac{1}{|\mu|}. \tag{12.14}$$

Thus, $A - zI \ : \ \mathcal{D}_A \mapsto \mathcal{D}_A(z) \hookrightarrow H$ is “onto” and bounded operator and $(A - zI)^{-1} \ : \ \mathcal{D}_A(z) \mapsto \mathcal{D}_A$ is also a bounded operator. But a bounded operator is extended on the closure $\overline{\mathcal{D}_A(z)}$. Now we recall that $A_z = A - zI$ is closed, therefore A_z^{-1} is closed which means that $\mathcal{D}_A(z) = \overline{\mathcal{D}_A(z)}$. □

Combining (a), (b) and (c) in the case of the self adjoint operator $A = A^*$ we have that if $z \in \mathbb{C} \setminus \mathbb{R}$, then $\mathcal{D}_A(z) = H$ (otherwise $\overline{z} \in \sigma_p(A^*) = \sigma_p(A)$). So, the index of A is $(0, 0)$ and $\sigma(A) \subseteq \mathbb{R}$.

Let now $\lambda \in \mathbb{R}$. We will show that $\sigma(A) = \sigma_p(A) \cup \sigma_c(A)$. If $\overline{\mathcal{D}_A(\lambda)} \neq H$ then $\lambda\sigma_p(A)$. If $\mathcal{D}_A(\lambda) = H$ then there exists the inverse $(A - \lambda I)^{-1}$ and it is bounded. Indeed, $A - \lambda I$ is self adjoint; by Theorem 3 $(A - \lambda I)^{-1}$ is self-adjoint; then by Theorem 1 ?????????????????????????????????

12.3 Elements Of The “Graph Method”

Theorem 12.3.1 *Let $\overline{\mathcal{D}_A} = H$. If A admits a closure, then A^{**} exists and $A^{**} = \overline{A}$. In the opposite direction, if A^{**} exists (meaning that $\mathcal{D}_{A^*}$ is dense in H) then A admits a closure and $A^{**} = \overline{A}$.*

Proof: Let

$$\Gamma(A) = \{(x; Ax)\}_{x \in \mathcal{D}_A} \subseteq \mathbb{H} = H \oplus H$$

be the graph of A. Consider the unitary operator $U(x; y) = (y; -x)$. Note that $U^2 = -I$. Then

$$\mathbb{H} \ominus \cup \Gamma(A) = \{(y; y^*) \ : \ (Ax, y) - (x, y^*) = \langle U(x; Ax), (y; y^*) \rangle = 0\} = \Gamma(A^*). \tag{12.15}$$

The above line means that if $U\Gamma(A)^\perp$ is a graph of some operator then this operator is A^*. Now, if $\Gamma(A)$ is closed then

$$\mathbb{H} = U\Gamma(A) \oplus \Gamma(A^*) \tag{12.16}$$

and applying U (which does not change $\mathbb{H}$) we get that

$$\mathbb{H} = U\Gamma(A^*) \oplus \Gamma(A), \tag{12.17}$$

since $U^2 = -I$ and $-\Gamma(A) = \Gamma(A)$. So, $\Gamma(A) = \Gamma(A^{**})$. Similarly in the case that $\overline{A}$ exists: $\overline{U\Gamma(A)} = U\overline{\Gamma(A)}$ and the equality $\mathbb{H} = \Gamma(A^*) = U\Gamma(\overline{A})$ implies that $\Gamma(\overline{A}) = \Gamma(A^{**})$.

Also in the inverse direction, if A^{**} exists then $\gamma(A^{**}) = \overline{\Gamma(A)}$ and $\overline{\Gamma(A)}$ is a graph of some operator called (by definition) $\overline{A}$. $\square$

12.4 Reduction Of Operator

Let $E_1 \oplus E_2 = H$ and let P be an orthoprojection onto the subspace E_1. We say that E_1 reduces A iff $P\mathcal{D}_A \subseteq \mathcal{D}_A$ and E_1, E_2 are invariant subspaces of A. Note that the linearity of $\mathcal{D}_A$ implies that $(I-P)\mathcal{D}_A \subseteq \mathcal{D}_A$ and for every $x \in \mathcal{D}_A$ $Ax = Ax_1 + Ax_2$ where $x_1 = Px$ and $x_2 = x - x_1$.

Lemma 12.4.1 *E reduces A iff*
(i) $P\mathcal{D}_A \subseteq \mathcal{D}_A$ and
(ii) $PAx = APx$ for every $x \in \mathcal{D}_A$.

(Here as before, P is the orthoprojection onto E.) The proof of this lemma is obvious.

Theorem 12.4.2 (Decomposition) *Let A be a closed operator, $H_k \hookrightarrow H$ and*

$$H = \oplus \sum_{k=1}^{\infty} H_k, \tag{12.18}$$

the orthogonal decomposition of H into the sum of subspaces H_k. Let P_k be the orthoprojection onto H_k and A is reduced by every H_k. Then $x \in \mathcal{D}_A$ iff $P_k x \in \mathcal{D}_A$ and $\sum_1^\infty \|AP_k x\|^2 < \infty$. Moreover, $Ax = \sum_1^\infty AP_k x$.

Proof: If $x \in \mathcal{D}_A$ then $P_k x \in \mathcal{D}_A$ and $P_k A = AP_k$. Moreover for all $x \in \mathcal{D}_A$ we have $Ax = \sum P_k Ax$ and

$$\|Ax\|^2 = \sum \|P_k Ax\|^2 = \sum \|AP_k x\|^2 < \infty. \tag{12.19}$$

In the opposite direction, let $\sum \|AP_k x\|^2 < \infty$; then $\sum_1^n P_k x \to x$ and $A\sum_1^n P_k x \to y$. Since H is closed $Ax = y$ and $x \in \mathcal{D}_A$. $\square$

An example of spectral decomposition is provided by the following theorem.

Theorem 12.4.3 *Let $\{E_\lambda\}_{-\infty}^{\infty}$ be a spectral family of orthoprojections E_λ, i.e., $E_\lambda \to 0$ in the strong sense as $\lambda \to -\infty$. and $E_\lambda \to I$ as $\lambda \to \infty$, $E_\lambda \leq E_\mu$ for $\lambda \leq \mu$ and $E_{\lambda+0} = E_\lambda$ (semicontinuity from the right). Consider the operator A with*

$$\mathcal{D}_A = \{x \mid \int_{-\infty}^{\infty} \lambda^2 d(E_\lambda x, x) < \infty\}$$

and

$$A = \int_{-\infty}^{\infty} \lambda dE_\lambda,$$

which means that for $x \in \mathcal{D}_A$, $Ax = \int_{-\infty}^{\infty} \lambda dE_\lambda x$. Then A is a self-adjoint operator and

$$\|Ax\|^2 = \int \lambda^2 d(E_\lambda x, x).$$

We say that A has a spectral decomposition.

Proof: All properties but the self-adjointness of A follow immediately from the theorem of decomposition (obviously the sequence $\int_{-N}^{M} \lambda dE_\lambda x$ is Cauchy). To prove that A is self-adjoint, let $P_{-N,M} = E_M - E_{-N}$ be an orthoprojection. Then for every $x, y \in \mathcal{D}_A$

$$(AP_{-N,M}x, y) = \int_{-N}^{M} \lambda d(E_\lambda x, y) = (x, \int_{-N}^{M} \lambda dE_\lambda y) \tag{12.20}$$

and $(Ax, y) = (x, Ay)$. So, A is symmetric and $AP_{-N,M} = P_{-N,M}A$. Now for all $y \in \mathcal{D}_{A^*}$

$$(AP_{-N,M}x, y) = (P_{-N,M}x, y^*), \tag{12.21}$$

which implies that $(Ax, P_{-N,M}y) = (x, P_{-N,M}y^*)$ for all $x \in \mathcal{D}_A$. Putting together (12.20) and (12.21) we get that for every $x \in \mathcal{D}_A$

$$(x, A^*P_{-N,M}y) = (x, P_{-N,M}A^*y) \tag{12.22}$$

which implies $A^*P_{-N,M}y = P_{-N,M}A^*y$ and this is (by (12.20)) equal to

$$\int_{-N}^{M} \lambda dE_\lambda y.$$

Thus, sending both N and M to infinity we get that

$$A^*y = \int_{-\infty}^{\infty} \lambda dE_\lambda y = Ay \tag{12.23}$$

and $y \in \mathcal{D}_A$ (since the convergence of the integral is in norm). □

12.5 Cayley Transform

Let A be a closed symmetric operator. Define for every $X \in \mathcal{D}_A$

$$(A + iI)x = y$$

$$(A - iI)x = z.$$

We proved before that $H_1 = \mathcal{D}_{-i} = \mathrm{Im}(A - iI)$ are closed subspaces of H. Also both operators $A \pm iI$ are one-to-one because $\pm i$ cannot be the eigenvalues of a symmetric operator. So, x defines both y and z in a unique way. Then the operator $z = Vy$ is defined. We checked that

$$\|y\|^2 = \|Ax\|^2 + \|x\|^2 = \|z\|^2$$

and V is an isometry. Note that 1 is not an eigenvalue of V because $z = y$ implies that $x = 0$ and $z = y = 0$.

We call this isometric operator $V \ : \ H_1 \mapsto H_2$ the *Cayley Transform* of A. The inverse transform is

$$x = \frac{1}{2i}(I - V)y \text{ and } Ax = \frac{1}{2}(I + V)y.$$

So, $Ax = i(I + V)(I - V)^{-1}x$.

Remark 12.5.1 *We proved before that A is self-adjoint implies that $H_1 = H_2 = H$ (the indices of defect are $(0,0)$) and V is a unitary operator.*

Theorem 12.5.2 *If A is a closed symmetric operator and the Cayley transform V is a unitary operator (i.e., $H_1 = H_2 = H$ and the indices of defect are $(0,0)$) then A is self-adjoint with a spectral decomposition $E_t F_s$ and $t = -ctg(s/2)$, where $\{F_s\}_0^{2\pi}$ is the spectral decomposition of V. This means that*

$$\mathcal{D}_A = \{x \mid \int_{-\infty}^{\infty} t^2 d(E_t x, x) < \infty\} \tag{12.24}$$

and

$$Ax = \int_{-\infty}^{\infty} t dE_t x. \tag{12.25}$$

Proof: Let $x \in \mathcal{D}_A$ and $x = \frac{1}{2i}(I - V)y$. Then

$$F_s x = \frac{1}{2i}(I - V)F_s y = \frac{1}{2i}\int_0^s \left(1 - e^{i\tau}\right) dE_\tau \tag{12.26}$$

and

$$(F_s x, x) = \frac{1}{4}\left((2I - V - V^{-1})F_s y, y\right) = \int_0^s \sin^2\frac{\pi}{2} d(F_\tau y, y), \tag{12.27}$$

where we used that $2 - e^{i\tau} - e^{-i\tau} = 4\sin^2\frac{\tau}{2}$. Similarly, for $x \in \mathcal{D}_A$ we have that

$$Ax = \frac{1}{2}(I + V)y = \frac{1}{2}\int_0^{2\pi}(1 + e^{is})dF_s x, \tag{12.28}$$

which, using (12.26), gives:

$$\begin{aligned}
Ax &= \frac{2i}{2}\int_0^{2\pi}\frac{1+e^{is}}{1-e^{is}}\frac{1-e^{is}}{2i}dF_s y \\
&= \int_0^{2\pi} i\frac{1+e^{is}}{1-e^{is}}dF_s x \\
&= -\int_0^{2\pi} ctg(s/2)sF_s x \\
&= \int_{-\infty}^{\infty} tdE_t x.
\end{aligned}$$

Now, using (12.28), for $x \in \mathcal{D}_A$ we have that

$$\begin{aligned}
\|Ax\|^2 &= (Ax, Ax) = \frac{1}{4}\left((2I - V - V^{-1})y, y\right) \\
&= \int_0^{2\pi}\frac{\cos^2(s/2)}{\sin^2(s/2)}\sin^2(s/2)d(F_s y, y)
\end{aligned}$$

and by (12.27) it follows that this equals

$$\int_0^{2\pi} ctg^2(s/2)d((F_s x, x) = \int_{-\infty}^{\infty} t^2 d(E_t x, x).$$

This means that $x \in \mathcal{D}_A$ implies $\int_{-\infty}^{\infty} t^3 d(E_t x, x) < \infty$.

The last part of the proof shows that if x is such that

$$\int_{-\infty}^{\infty} t^2 d(E_t x, x) = \int_0^{2\pi} ctg^2(s/2)d(F_s x, x) < \infty,$$

then $x \in \mathcal{D}_A$. To show this we have to find $y \in H$ such that $\frac{(I-V)y}{2i} = x$. In order to find such a y we start with the information

$$\int_0^{2\pi} ctg^2(s/2) d\sigma(s) < \infty,$$

for $\sigma(s) = (F_s x, x)$ being a monotone function of bounded variation:

$$\int_0^{2\pi} d\sigma(s) < \infty. \tag{12.29}$$

Then also

$$\int_0^{2\pi} \frac{1}{\sin^2(s/2)} d\sigma(s) < \infty. \tag{12.30}$$

Therefore there exists

$$y = \int_0^{2\pi} \frac{e^{-is/2}}{\sin(s/2)} dF_s.$$

[Indeed, consider first the $y_{\varepsilon,\eta} = \int_\varepsilon^{2\pi\eta} \frac{-e^{is/2}}{\sin(s/2)} dF_s x$ which exists because it is the integral of a continuous function; we cut off singular points. Then observe that it is a Cauchy sequence when $\varepsilon \to 0,\ \eta \to 0$ since (12.30) exists.]

Consider now $(I-V)y = \int_0^{2\pi} (1 - e^{is}) dF_s y$ and note that

$$F_s y = -\int_0^s \frac{e^{-i\tau/2}}{\sin(\tau/2)} dF_\tau x.$$

Therefore,

$$\begin{aligned}
(I-V) &= \int_0^{2\pi} \frac{(1-e^{is})e^{-is/2}}{\sin(s/2)} dF_s x \\
&= 2i \int_0^{2\pi} dF_s x \\
&= 2ix
\end{aligned}$$

and $x \in \mathcal{D}_A$. □

Corollary 12.5.3 *If* $1 \notin \sigma(V)$ *then* A *is a bounded (self-adjoint) operator.*

Return back to the construction of the Cayley transform of a symmetric operator. Let A_1 be a symmetric extension of A, $A_1 \supseteq A$ and $A_1 \neq A$. Then there exists $x_1 \in \mathcal{D}_{A_1} \setminus \mathcal{D}_A$ which means that

$$(A + iI)x_1 = y_1 \text{ and } (A - iI)x_1 = z_1 - V_1 y_1$$

and both $y_1 \notin H_1$, $z_1 \notin H_2$. So, the Cayley transform V_1 of A_1 is an isometric extension of V and does not coincide with V. This means that there exists a $y_1 \in H \setminus H_1$ and there exists a $z_1 \in H \setminus H_2$. Hence we have the following:

Fact 1. If the indices of defect are $(n, 0)$ and $(0, n)$ for $n \neq 0$ then A does not have any symmetric extension, i.e., A is a maximal symmetric (and not self-adjoint) operator.

Let $\mathcal{D}_{V_1} = H_1' = H_1 \oplus L_1$ and $\mathrm{Im} V_1 - H_2' = H_2 \oplus L_2$. So, $V_1 : H_1 \oplus L_1 \mapsto H_2 \oplus L_2$. Also, $V_1|_{H_1} = V$ and V_1 restricted on L_1 is an isometry between L_1 and L_2. In particular, $\dim L_1 = \dim L_2$. Therefore we arrive at

Fact2. If the indices of A are (m, n) and $m \neq n$ then there is **no** self-adjoint extension of A [because any self-adjoint extension A_1 has Cayley transform $V_1 : H \mapsto H$ meaning that $H_1 \oplus L_1 = H$ and $H_2 \oplus L_2 = H$ and $\mathrm{codim} H_1 = \mathrm{codim} H_2$].

Consider now the inverse question: let V be that Cayley transform of A and let V_1 be some isometric extension of V.

Does there exist a symmetric extension A_1 of A such that V_1 is the Cayley transform of A_1?

The answer is "yes" and formulas for V builds this extension: we consider an operator A_1 with

$$\mathcal{D}_A = \{x \mid x = \frac{1}{2i}(I - V)y,\ y \in \mathcal{D}_{V_1}\} \tag{12.31}$$

and for $x \in \mathcal{D}_{A_1}$, $x = \frac{1}{2i}(I - V_1)y$,

$$A_1 x = \frac{1}{2}(I + V_1)y.$$

In order to see that the operator A_1 is well defined we need to show that $1 \notin \sigma_p(V_1)$, i.e., $\ker(I - V_1) = 0$. If $1 \in \sigma_p(V_1)$ then there exists $y_0 \neq 0$ such that $y_0 = V_1 y_0$. Let us check that such a $y_0 \perp \mathcal{D}_{A_1}$, which will be a contradiction because $\mathcal{D}_{A_1} \supseteq \mathcal{D}_A$ and $\mathcal{D}_A$ is dense in H. So for any $x \in \mathcal{D}_{A_1}$ there is a y and $x = \frac{1}{2i}(I - V_1)y$ and

$$(y_0, x) = (y_0, \frac{1}{2i}(I - V_1)y) = \frac{1}{2i}[(y_0, y) - (y_0, V_1 y)] = 0, \tag{12.32}$$

because $V_1 y_0 = y_0$ and $(V_1 y_0, V_1 y) = (y_0, y)$ since V_1 is an isometry. It follows that $y_0 = 0$. It remains to show that A_1 is a symmetric operator: for any $x_1, x_2 \in \mathcal{D}_{A_1}$ we have that

$$(A_1 x_1, x_2) = -\frac{1}{4i}\left((I+V_1)y_1, (I-V_1)y_2\right) = (x_1, A_1 x_2). \tag{12.33}$$

[Indeed,

$$\begin{aligned}(A_1 x_1, x_2) &= -\frac{1}{4i}[(y_1, y_2) + (V_1 y_1, y_2) - (y_1, V_1 y_2) - (V_1 y_1, V_1 y_2)] \\ &= -\frac{1}{4i}[(V_1 y_1, y_2) - (y_1, V_1 y_2)].\end{aligned}$$

Similarly,

$$\begin{aligned}(x_1, A_1 x_2) &= \frac{1}{4i}((I-V_1)y_1, (I+V_1)y_2) &(12.34)\\ &= \frac{1}{4i}[-(V_1 y_1, y_2) + (y_1, V_1 y_2)].] &(12.35)\end{aligned}$$

As a consequence we have:

Fact 3. If the indices of a symmetric operator A are (n, n), then there exists a self-adjoint extension A_1 of A.

Indeed, if V is the Cayley transform of A and $V : H_1 \mapsto H_2$, $H = H_1 \oplus L_1$, $H = H_2 \oplus L_2$, $\dim L_1 = \dim L_2 = n$ then it is trivial to build an extension of V to a unitary operator $V_1 : H \mapsto H$. The corresponding symmetric extension A_1 of A has indices $(0, 0)$ and by the last Theorem it is self-adjoint.

Example. Consider the operation $Ay = iy'$ on the spaces $L_2[0, 1]$, $L_2[0, \infty)$, and $L_2(-\infty, \infty)$.

Index