Hilbert spaces

September 3, 2014

1 Introduction

In linear algebra one of the main notions is the concept of a *vector space*. Geometry on finite dimensional vector spaces is called *Euclidean geometry* and is based on the concept of inner product spaces which brings notions of orthogonality, distance etc.

This semester we will attempt to study linear algebra in infinite dimensions. This will bring a lot of new features that cannot be seen in finite dimensions such as the concept of continuity of linear operators (not all linear operators are continuous), completness and some new metric properties (the relation between closed and compact sets is more delicate than in finite dimensions).

Recall a well-known theorem from finite dimensions:

Theorem 1.1. *Let A be a real symmetric $n \times n$ matrix. Then there is an orthonormal basis $\{e_1, e_2, \dots, e_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of A, i.e.,*

$$Ae_j = \lambda_j e_j, \qquad j = 1, 2, \dots, n.$$

Hence if $x = \sum_{i=1}^{n} \alpha_j e_j$, then $Ax = \sum_{i=1}^{n} \lambda_j \alpha_j e_j$.

One of our main goals this semester will be to establish an analogue of this statement for infinite dimensional spaces. See section 5.7 for more.

2 Inner product and normed linear spaces

Unless said specifically otherwise all scalars will be assumed to be *complex* (in $\mathbb{C}$) unless stated specifically as *real* (in $\mathbb{R}$).

Definition 2.1. *Let X be a vector space over $\mathbb{C}$. If there exists a function $\langle .,. \rangle : X \times X \to \mathbb{C}$ such that*

- $\langle x, x\rangle \geq 0$ *and* $\langle x, x\rangle = 0$ *if and only if* $x = 0$.
- $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
- $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle, \quad \lambda \in \mathbb{C}$

- $\langle x, y\rangle = \overline{\langle y, x\rangle}$,

then $(X, \langle ., .\rangle)$ *is called a inner product space (i.p.s.).*

Examples: 1. $\mathbb{C}^n$, $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n) \in \mathbb{C}^n$ and

$$\langle x, y\rangle = \sum_{i=1}^{n} x_j \overline{y_j}.$$

2. $\ell^2 = \{x = (x_1, x_2, x_3, \dots) \in \mathbb{C}^\infty : \sum_{i=1}^{\infty} |x_i|^2 < \infty\}$ with inner product

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_j \overline{y_j}.$$

3. $M_n(\mathbb{C})$ the space of all $n \times n$ matrices with complex entries with inner product

$$\langle A, B\rangle = \operatorname{tr}(B^* A), \quad B^* = \overline{B^T}.$$

4. $C[0, 1] = \{f : [0, 1] \to \mathbb{C} : f$ is continuous on $[0, 1]\}$ with

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

5. $C^1[0, 1] = \{f : [0, 1] \to \mathbb{C} : f$ is continuously differentiable on $[0, 1]\}$ with

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}\, dx + \int_0^1 f'(x)\overline{g'(x)}\, dx.$$

Definition 2.2. *Let* $(X, \langle ., .\rangle)$ *be an i.p.s. Then for* $x \in X$ *we define*

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

The function $\|.\| : X \to [0, \infty)$ *is called a norm. It gives the notion of length of a vector* x.

Recall that if $x, y \in \mathbb{R}^n$, then $\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ and

$$-1 \le \frac{\langle x, y\rangle}{\|x\|\|y\|} \le 1.$$

Then there exists a unique $\theta \in [0, \pi]$ such that $\cos\theta = \frac{\langle x,y\rangle}{\|x\|\|y\|}$. The number θ is called an angle between vectors x and y.

Is there an analogue of this in any i.p.s ?

Lemma 2.1. *(Cauchy-Schwartz inequality) Let* $(X, \langle ., .\rangle)$ *be any i.p.s. Then*

$$|\langle x, y\rangle| \le \|x\|\|y\|, \qquad \textit{for all } x, y \in X.$$

Furthermore the equality holds if and only if vectors x, y *are linearly dependent.*

Proof. We may assume that $y \neq 0$ and $\|y\| = 1$. Indeed, the Cauchy-Schwartz inequality holds when $y = 0$. If $y \neq 0$ then $z = y/\|y\|$ has length 1. So if $|\langle x, z\rangle| \leq \|x\|$ holds then

$$|\langle x, z\rangle| = \frac{|\langle x, y\rangle|}{\|y\|} \leq \|x\|,$$

from which $|\langle x, y\rangle| \leq \|x\|\|y\|$ follows.

Assume therefore that $\|y\| = 1$. We look at $\|x - \langle x, y\rangle y\|^2$.

$$\begin{aligned}
\|x - \langle x, y\rangle y\|^2 &= \langle x - \langle x, y\rangle y, x - \langle x, y\rangle y\rangle = \langle x, x\rangle - \langle x, \langle x, y\rangle y\rangle - \\
&- \langle\langle x, y\rangle y, x\rangle + \langle\langle x, y\rangle y, \langle x, y\rangle y\rangle = \\
&= \|x\|^2 - \overline{\langle x, y\rangle}\langle x, y\rangle - \langle x, y\rangle\langle y, x\rangle + \langle x, y\rangle\overline{\langle x, y\rangle} = \\
&= \|x\|^2 - |\langle x, y\rangle|^2. \qquad (2.1)
\end{aligned}$$

From this $\|x\|^2 - |\langle x, y\rangle|^2 \geq 0$, since the lefthand side is nonnegative.

Equality holds if and only if $\|x - \langle x, y\rangle y\|^2 = 0$, i.e., $x - \langle x, y\rangle y = 0$. So vector x is a scalar multiple of y and therefore x, y are linearly dependent.

Lemma 2.2. *Let $(X, \langle ., .\rangle)$ be an i.p.s. Then $\|x\| = \sqrt{\langle x, x\rangle}$ satisfies*

- $\|x\| \geq 0$*, with equality if and only if* $x = 0$*,*
- $\|\alpha x\| = |\alpha|\|x\|$ *for all* $x \in X$ *and* $\alpha \in \mathbb{C}$*,*
- $\|x + y\| \leq \|x\| + \|y\|$ *(triangle inequality).*

Proof. The first two are a simple exercises. The triangle inequality holds, since using Cauchy-Schwartz:

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\, \langle x, y\rangle \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \qquad (2.2)
\end{aligned}$$

Lemma 2.3. *Let $(X, \langle ., .\rangle)$ be an i.p.s. Then*

- *(Pythagoras theorem) If* $\langle x, y\rangle = 0$*, then* $\|x + y\|^2 = \|x\|^2 + \|y\|^2$*.*
- *(Parallelogram law) For all* $x, y \in X$*:* $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$*.*
- *(Polarization identity) For all* $x, y \in X$*:*

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

Proof. Exercise, homework.

Definition 2.3. *Let X be a vectors space with a function $\|.\| : X \to \mathbb{R}$ such that*

- $\|x\| \geq 0$*, with equality if and only if* $x = 0$*,*

- $\|\alpha x\| = |\alpha|\|x\|$ *for all* $x \in X$ *and* $\alpha \in \mathbb{C}$,
- $\|x + y\| \leq \|x\| + \|y\|$ *(triangle inequality).*

The $(X, \|.\|)$ *is called a linear normed space (l.n.s) and function* $\|.\|$ *is called a norm.*

Corollary 2.4. *Lemma 2.2 implies that any inner product space* $(X, \langle ., . \rangle)$ *is also a normed linear space with norm*

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

Other examples: 1. $\mathbb{C}^n$ with norm

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|,$$

or norm

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

2. $C[0, 1]$ with norms

$$\|f\|_\infty = \max_{0 \leq x \leq 1} |f(x)|, \qquad \|f\|_p = \left(\int_0^1 |f(x)|^p \, dx \right)^{1/p}, \quad 1 \leq p < \infty.$$

Notice that the norm $\|.\|_2$ comes from an inner product:

$$\|f\|_2 = \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2} = \sqrt{\int_0^1 f(x)\overline{f(x)} \, dx} = \sqrt{\langle f, f \rangle}.$$

3. ℓ^p for $1 \leq p < \infty$, where

$$\ell^p = \{x = (x_1, x_2, x_3, \dots) : x_i \in \mathbb{C} \text{ and } \sum_{i=1}^{\infty} |x_i|^p < \infty\}$$

and

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

One can also define ℓ^∞ by

$$\ell^\infty = \{x = (x_1, x_2, x_3, \dots) : x_i \in \mathbb{C} \text{ and } \sup_{i \in \mathbb{N}} |x_i| < \infty\},$$

with

$$\|x\|_\infty = \sup_{i \in \mathbb{N}} |x_i|.$$

Notice that when $p = 2$ the ℓ^2 is and i.p.s. Also $\ell^p \subset \ell^q$ for $p \leq q$.

Question: When does the norm arise from an inner product? Given n.l.s $(X, \|.\|)$ does there exists $\langle ., . \rangle$ an inner product such that $\|x\|^2 = \langle x, x \rangle$.

Claim: A necessary and sufficient condition is that the parallelogram law (Lemma2.3) holds for all $x, y \in X$:

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

3 Metric space structure of normed linear spaces

Consider a normed linear space $(X, \|.\|)$. Then

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

is a metric on X, hence X is a metric space! Hence we can talk about continuity of functions on X.

Exercise: Show that

- $x \mapsto \|x\|$ is continuous map $X \to \mathbb{R}$,
- $(x, y) \mapsto x + y$ is a continuous map $X \times X \to X$,
- $(\lambda, x) \mapsto \lambda x$ is a continuous map $\mathbb{R} \times X \to X$,
- if in addition X is an i.p.s then $(x, y) \mapsto \langle x, y \rangle$ is a continuous map $X \times X \to \mathbb{C}$.

Note. Recall that convergence on a metric space is defined as follows. If (x_n) is a sequence on X, we say that $x_n \to x$ (sequence converges to x) if

$$\lim_{n\to\infty} d(x_n, x) = \lim_{n\to\infty} \|x_n - x\| = 0.$$

We say that a sequence (x_n) is Cauchy if

$$\lim_{n,m\to\infty} d(x_n, x_m) = \lim_{n,m\to\infty} \|x_n - x_m\| = 0.$$

Recall also that every convergent sequence is Cauchy, the reverse is also true if (X, d) is a **complete** metric space.

3.1 Closed, open and compact sets on $(X, \|.\|)$

Again this is a review from the 3rd year metric spaces course. Let $(X, \|.\|)$ by any inner product space.

Definition 3.1. *Set $\Omega \subset X$ is called open if for every $x \in \Omega$ there is $\varepsilon > 0$ such that*

$$B_\varepsilon(x) = \{y \in X; \|y - x\| < \varepsilon\} \subset \Omega.$$

Definition 3.2. *Set $F \subset X$ is called closed if*

$$\Omega = X \setminus F = F^c$$

is open.

Fact: Set F is closed if and only if for every convergent sequence $(x_n) \subset F$ we have that

$$x = \lim x_n \in F.$$

(We say that F contains all it limit points.)

Definition 3.3. *Set $K \subset X$ is called compact if for any sequence $(x_n) \subset K$ there is a subsequence (x_{n_k}) such that*

$$\lim_{k\to\infty} x_{n_k} \text{ exists and belongs to the set } K.$$

Recall that:

- Compact set are closed
- whenever $\dim X < \infty$ closed balls

  $$\overline{B_r(x)} = \{y \in X; \|x - y\| \leq r\}$$

  are compact
- When $\dim X = \infty$ closed balls $\overline{B_r(x)}$ are **never** compact (last exercise on sheet 2).

Example. This example shows that whether or not a set is closed depends on the considered norm. Let W be a set consisting of these continuous functions:

$$f_n(x) = nx, \quad \text{on } [0, 1/n] \qquad f_n(x) = 1, \quad \text{on } [1/n, 1],$$

for $n = 1, 2, 3, \ldots$. Then W is a closed set on the space $(C[0,1], \|.\|_\infty)$ but not on $(C[0,1], \|.\|_1)$.

Indeed, on $(C[0,1], \|.\|_1)$ we have that $f_n \to f$, where f is a constant function $1 \notin W$.

$$\|f_n - f\|_1 = \int_0^1 |f_n(x) - 1| dx = \int_0^{1/n} |f_n(x) - 1| dx \leq \frac{2}{n} \to 0.$$

Definition 3.4. *Let $A \subset X$, where X is vector space. In what follows we will denote by $lin(A)$ the linear span of A:*

$$lin(A) = \left\{ \sum_{i=1}^{N} \alpha_i x_i; \text{ for } \alpha_i \in \mathbb{C} \text{ and } x_i \in A \right\}.$$

The set $lin(A)$ is the smallest subspace of X containing A.

Definition 3.5. *Let $A \subset X$, where X is a n.l.s. The closure of A denoted by $\overline{A}$ is the smallest closed set in X containing A. It can be shown that*

$$\overline{A} = \bigcap_{A \subset F,\, F closed} F$$

and

$$\overline{A} = \{x \in X; \textit{there is a sequence } (x_n) \subset A \textit{ such that } x_n \to x\}.$$

Examples: $X = (C[0,1], \|.\|_\infty)$. Let

$$A = \{x^n; n = 0, 1, 2, \dots\}.$$

Then $\mathrm{lin}(A)$ is the set of all polynomials on $[0,1]$:

$$\mathrm{lin}(A) = \left\{ \sum_{i=0}^{n} a_i x^i;\ a_i \in \mathbb{C} \right\}.$$

What is $\overline{\mathrm{lin}(A)}$? **Answer:** Weierstrass theorem.

$$\overline{\mathrm{lin}(A)} = C[0,1].$$

3.2 Orthogonality in inner product spaces

In this subsection X will be any inner product space. Recall first the well-known *Gram-Schmidt ortogonalization process.*

If $\{y_1, y_2, \dots, y_N\} \subset X$ are N linearly independent vectors from X, then we can find vectors $\{x_1, x_2, \dots, x_N\}$ that are orthogonal ($\langle x_i, x_j \rangle = 0$ for $i \neq j$) and

$$\mathrm{lin}(\{y_1, y_2, \dots, y_N\}) = \mathrm{lin}(\{x_1, x_2, \dots, x_N\}).$$

These vectors are defined inductively by

$$\begin{aligned}
x_1 &= y_1, \\
x_2 &= y_2 - \langle y_2, x_1 \rangle \frac{x_1}{\|x_1\|^2}, \\
\dots & \qquad \dots \\
x_N &= y_N - \sum_{i=1}^{N-1} \langle y_N, x_i \rangle \frac{x_i}{\|x_i\|^2}.
\end{aligned}$$

If required these vectors can be made orthonormal (unit length) by taking $z_i = \frac{x_i}{\|x_i\|}$, $i = 1, 2, \dots, N$.

Definition 3.6. *A family $(x_\alpha)_{\alpha\in A} \subset X$ (A is a set of indices, typically a finite set or $\mathbb{N}$) in an i.p.s. space X is called orthonormal family if*

- $\langle x_\alpha, x_\beta\rangle = 0$, *if* $\alpha \neq \beta$,
- $\langle x_\alpha, x_\alpha\rangle = 1$, *for all* $\alpha \in A$.

Examples: In ℓ^2 - the space of all square sumable sequences the family $(e_n)_{n\in\mathbb{N}}$ is othonormal. Here

$$e_n = (0, 0, \dots, 1, 0, 0, \dots), \qquad \text{where 1 appears on the } n\text{-th position.}$$

The space $(C[0,1], \|.\|_2)$ has an ON family $(e^{inx})_{n\in\mathbb{N}}$.

Lemma 3.1. *(Bessel's inequality) Let X by any i.p.s. and $(e_i)_{i=1}^N$ be a finite ON family. Then*

$$\sum_{i=1}^N |\langle x, e_i\rangle|^2 \le \|x\|^2, \qquad \textit{for all } x \in X.$$

The equality holds if and only if $x \in lin(\{e_1, e_2, \dots, e_N\})$. Then

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

Proof. Clearly:

$$\begin{aligned} 0 &\le \left\| x - \sum_{i=1}^N \langle x, e_i\rangle e_i \right\|^2 = \left\langle x - \sum_{i=1}^N \langle x, e_i\rangle e_i, x - \sum_{j=1}^N \langle x, e_j\rangle e_j \right\rangle \\ &= \|x\|^2 - \sum_{i=1}^N \langle x, e_i\rangle\overline{\langle x, e_i\rangle} - \sum_{j=1}^N \langle x, e_j\rangle\overline{\langle x, e_j\rangle} + \sum_{i,j=1}^N \langle x, e_i\rangle\overline{\langle x, e_j\rangle}\langle e_i, e_j\rangle. \\ &= \|x\|^2 - \sum_{i=1}^N |\langle x, e_i\rangle|^2, \end{aligned}$$

using the orthogonality of e_i. From this the claim follows. It also follows that the equality holds if and only if

$$\left\| x - \sum_{i=1}^N \langle x, e_i\rangle e_i \right\|^2 = 0,$$

i.e., $x - \sum_{i=1}^N \langle x, e_i\rangle e_i = 0$.

3.3 Hilbert and Banach spaces

We may ask a legitimate question whether the Bessel's inequality holds even if the ON family (e_i) is infinite. If so, what is the sum

$$\sum_{i=1}^{\infty} \langle x, e_i \rangle e_i \quad ?$$

It is fairly easy to check that the sequence (s_N) if partial sums

$$s_N = \sum_{i=1}^{N} \langle x, e_i \rangle e_i, \qquad \text{is Cauchy, i.e.,}$$

$$\lim_{n,m\to\infty} \|s_n - s_m\| = 0.$$

But, if a metric space is *not complete* it is not true that every Cauchy sequence is convergent. This leads us to define:

Definition 3.7. *Let $(X, \|.\|)$ be a normed linear space. We call $(X, \|.\|)$ a Banach space, if (X, d) with metric $d(x, y) = \|x - y\|$ is a complete metric space.*
Similarly, an inner product space that is a Banach space is called the Hilbert space.

If follows that in a Hilbert space the sum

$$\sum_{i=1}^{\infty} \langle x, e_i \rangle e_i$$

is a well defined element for any x.

Examples: 1. $\mathbb{R}^n$, $\mathbb{C}^n$ with usual inner product are Hilbert spaces

2. $\mathbb{R}^n$, $\mathbb{C}^n$ with $\|.\|_p$ norm $1 \le p \le \infty$ are Banach spaces

3. ℓ^p, $1 \le p \le \infty$ are Banach spaces, $p = 2$ is a Hilbert space

4. $(C[0, 1], \|.\|_p)$ is **not** a Banach space $1 \le p < \infty$, but when $p = \infty$ it is a Banach space.

The reason why $(C[0, 1], \|.\|_p)$ is not a Banach space $p < \infty$ is simple. Consider a sequence (f_n) if continuous functions defined by

$$\begin{aligned} f_n(x) &= 0 \quad \text{on } [0, 1/2), \qquad f_n(x) = n(x - 1/2) \quad \text{on } [1/2, 1/2 + 1/n), \\ f_n(x) &= 1, \quad \text{on } [1/2 + 1/n, 1]. \end{aligned}$$

This is a Cauchy sequence in $(C[0, 1], \|.\|_p)$, $p < \infty$ (check it!) but it limit is a discontinuous function $f(x) = 0$ on $[0, 1/2]$ and $f(x) = 1$ elsewhere. Hence $f \notin C[0, 1]$ and the space is not complete.

An important and interesting question is when a subspace of a Banach space is itself a Banach space. The complete answer provides the following lemma.

Lemma 3.2. *Let $(X, \|.\|)$ be a Banach space and $M \subset M$ its subspace. Then $(M, \|.\|)$ is a Banach space if and only if M is a closed subset of X.*

Proof. Suppose that M is closed. Let (x_n) be a Cauchy sequence in M. It follows that (x_n) is also a Cauchy sequence in X. As X is complete there is $x \in X$ such that

$$\|x_n - x\| \to 0, \qquad \text{as } n \to \infty.$$

Point x is a limit point of set M (since there is a sequence from M converging to it). But M is a closed set, i.e., it contains all its limit points - in particular it contains x. Hence $x_n \to x$ in M and so $(M, \|.\|)$ is a Banach space.

Conversely, if $(M, \|.\|)$ is a Banach space and $x_n \to x$ be a convergent sequence of points $x_n \in M$ with limit $x \in X$, then (x_n) is a Cauchy sequence in M. Hence by the assumption, there is $y \in M$ such that $x_n \to y$ in M. Clearly, $y = x$, hence $x \in M$ and so the set M is closed.

3.4 Completion of a normed linear space

As we have seen, a very important space $(C[0,1], \|.\|_p)$, $p < \infty$ we used in examples is not a Banach space. Is is therefore a legitimate question to ask whether this can be "fixed" somehow. To formulate this question mathematically, we ask whether for any n.l.s $(X, \|.\|)$ one can find a Banach space $(\widetilde{X}, \|.\|_*)$ such that $X \subset \widetilde{X}$ is a (dense) subspace and

$$\|x\| = \|x\|_* \qquad \text{for all } x \in X.$$

This is indeed the case, the abstract construction if such completion can be found below.

Theorem 3.3. *Let $(X, \|.\|)$ be a normed linear space. Then there exists a Banach space $(\widetilde{X}, \|.\|_*)$ (called completion of X) such that*

- $X \subset \widetilde{X}$
- $\overline{X} = \widetilde{X}$*, where is closure is taken w.r.t* $\|.\|_*$ *norm in* $\widetilde{X}$
- $\|x\| = \|x\|_* \qquad$ *for all* $x \in X$.

Proof. The construction of the completion is somewhat standard and you may have seen it before in different context. The idea of constructing $\widetilde{X}$ is the following:

Consider first a set $\mathcal{U}$ of all Cauchy sequences in X:

$$\mathcal{U} = \{v = (x_n)_{n=1}^{\infty};\ x_n \in X \text{ and the sequence } (x_n) \text{ is Cachy in } X.\}$$

We say two sequences $v = (x_n)_{n=1}^{\infty}$ and $w = (y_n)_{n=1}^{\infty}$ from $\mathcal{U}$ are equivalent $v \sim w$ if and only if

$$\lim_{n\to\infty} x_n - y_n = 0.$$

Clearly, this is an equivalence relation and it allows us to define

$$\widetilde{X} = \mathcal{U}/\sim = \{\text{set of equivalence classes in } \mathcal{U} \text{ under } \sim\}.$$

We denote the equivalence class of a sequence $v = (x_n)_{n=1}^{\infty}$ by $[v]$. To make $\widetilde{X}$ a vector space we define

$$\begin{aligned} [v] + [w] &= [(x_n + y_n)], && \text{for } v = (x_n)_{n=1}^{\infty} \text{ and} w = (y_n)_{n=1}^{\infty} \\ \alpha[v] &= [(\alpha x_n)], && \text{for a scalar } \alpha. \end{aligned} \tag{3.3}$$

One needs to verify that this definition is independent of particular choice of representatives u and v.

To make $\widetilde{X}$ a normed linear space we put

$$\|[v]\|_* = \lim_{n\to\infty} \|x_n\|, \qquad \text{where } v = (x_n)_{n=1}^{\infty}.$$

Again this definition is independent of particular choice of v from the equivalence class $[v]$.

We want to this about X as being a subset of $\widetilde{X}$. We can make it so, by identifying an element $x \in X$ with and equivalence class $[(x, x, x, \dots)] \in \widetilde{X}$. Clearly,

$$\|x\| = \|[(x, x, x, \dots)]\|_*.$$

Note also that X is a dense set in $\widetilde{X}$. To see this consider any $[v] \in \widetilde{X}$. Let $v = (x_n)_{n=1}^{\infty}$. Then one can see that

$$\|x_n - [v]\|_* \to 0 \qquad \text{as } n \to \infty.$$

So $[v]$ can be approximated by sequence from the subset X.

Finally we note that $(\widetilde{X}, \|.\|_*)$ is a Banach space. Indeed, let $([v_n])_{n=1}^{\infty}$ be a Cauchy sequence in $\widetilde{X}$, i.e.,

$$\lim_{n,m\to\infty} \|[v_n] - [v_m]\|_* = 0.$$

Fix an integer $n \in \mathbb{N}$. We have that $v_n = (x_i^n)_{i=1}^{\infty}$ is a Cauchy sequence in X. It is therefore possible to find an index $k = k(n)$ such that

$$\|x_m - x_{k(n)}\| < \frac{1}{n}, \qquad \text{for all } m \geq k(n).$$

Consider now a sequence $y_n = x_{k(n)}$, $n = 1, 2, 3, \dots$. We claim that this is a Cauchy sequence in X, hence $[(y_n)] \in \widetilde{X}$. Moreover one can check that

$$\|[v_n] - [(y_n)]\|_* \to 0, \qquad \text{as } n \to \infty.$$

It follows that $[(y_n)]$ is the limit point of the Cauchy sequence $([v_n])_{n=1}^{\infty}$.

Definition 3.8. *For any $1 \leq p < \infty$ we will denote by $L^p(0,1)$ the Banach space that is the completion of the normed linear space*

$$(C[0,1], \|.\|_p).$$

Note: (for students who had measure theory). Equivalently one can define the space $L^p(0,1)$ as

$$L^p(0,1) = \{f : [0,1] \to \mathbb{C};\ \text{function } f \text{ is measurable and } \int_0^1 |f(x)|^p\, dx < \infty\},$$

with norm

$$\|f\|_p = \left(\int_0^1 |f(x)|^p\, dx\right)^{1/p}.$$

Here we use the Lebesgue integral (which is a generalization of the Riemann integral). In this definition we also consider two functions $f, g : [0,1] \to \mathbb{C}$ to be the same if $\int_0^1 |f - g|^p\, dx = 0$.

The space $L^\infty(0,1)$ is defined as

$$L^\infty(0,1) = \{f : [0,1] \to \mathbb{C};\ \text{function } f \text{ is measurable and essentially bounded}\},$$

with norm

$$\|f\|_\infty = \text{ess sup}_{x\in[0,1]}|f(x)|.$$

(If you did not have the measure theory, imagine here f to be just bounded and its norm defined as the supremum of $|f(x)|$).

As an observation we note that the space $C[0,1]$ is a proper closed subspace of $L^\infty(0,1)$ w.r.t the norm $\|.\|_\infty$.

3.5 Equivalence of norms

As we have seen before, in many cases such as on $\mathbb{R}^n$ or $\mathbb{C}^n$ we had a choice of which norm to use. It did not mattered much however, as it had no effect on convergence of sequences, sets being open or closed, etc. On the other hand we know that if we consider the space $(C[0,1], \|.\|_p)$ for different p the spaces are substantially different, $p = \infty$ is a Banach space, $p < \infty$ is not; also certain sequences convergent for some value of p do not converge for the other values.

It turn out that this is connected with the notion of equivalence of norms.

Definition 3.9. *Let X be a vector space with two norms $\|.\|_1$ and $\|.\|_2$. We say these two norms are equivalent if there is number $C > 0$ such that*

$$C^{-1}\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1, \qquad \textit{for all } x \in X.$$

Lemma 3.4. *If the norms $\|.\|_1$ and $\|.\|_2$ are equivalent on X, then*

$$x_n \to x \ w.r.t\ \|.\|_1 \qquad \text{if and only if} \qquad x_n \to x \ w.r.t\ \|.\|_2.$$

Corollary 3.5. *If two norms $\|.\|_1$ and $\|.\|_2$ are equivalent on X, they generate the same* topology *on X, that is set that are open(closed) w.r.t one norm are also open(closed) w.r.t the other norm. Also,*

$$(X, \|.\|_1) \text{ is a Banach space if and only if } (X, \|.\|_2) \text{ is a Banach space.}$$

Examples: 1. All norms on $\mathbb{R}^n$ or $\mathbb{C}^n$ are equivalent (see below).

2. The norms $\|.\|_p$ on $C[0, 1]$ are not equivalent for different values of p (prove it!).

Theorem 3.6. *Let X be finite dimensional vector space. Then all norms on X are equivalent.*

Proof. Let dim $X = n$ and find a basis $(e_1, e_2, \dots, e_n)$ for X. Let us consider the norm

$$\|x\|_\infty = \max_{1\le i\le n} |x^i|, \qquad \text{for } x = \sum_{i=1}^{n} x^i e_i.$$

Let $\|.\|$ be ANY other norm on X. We claim that $\|x\| \le C_0\|x\|_\infty$, for $C_0 = n\left(\max_{1\le i\le n} \|e_i\|\right)$. Indeed, by triangle inequality (for $x = \sum_{i=1}^{n} x^i e_i$):

$$\|x\| \le \sum_{i=1}^{n} |x^i|\|e_i\| \le \left(\max_{1\le i\le n} \|e_i\|\right) \sum_{i=1}^{n} |x^i| \le n \left(\max_{1\le i\le n} \|e_i\|\right) \|x\|_\infty.$$

Furthermore we claim that there is $C > 0$ such that

$$\|x\|_\infty \le C\|x\| \qquad \text{for all } x \in X. \tag{3.4}$$

If (3.4) is true, then indeed $\|.\|$ and $\|.\|_\infty$ are equivalent. From this it follows that all norms on X must be equivalent. To see (3.4) consider the set

$$\mathcal{S} = \{x \in X;\ \|x\|_\infty = 1\}.$$

This set is compact, as it is closed and it is a subset of the set

$$\overline{B} = \{x \in X;\ \|x\|_\infty \le 1\}$$

which is compact by to following argument:

Set $\overline{B}$ is a continuous image of the compact set

$$\{(x_1, x_2, \dots, x_n) \in \mathbb{C}^n;\ |x_i| \le 1\}$$

by a map $(x_1, x_2, \dots, x_n) \mapsto \sum_{i=1}^n x_i e_i$.

Having this we see that

$$\|x\| = \left\| \|x\|_\infty \frac{x}{\|x\|_\infty} \right\| = \|x\|_\infty \left\| \frac{x}{\|x\|_\infty} \right\| \geq \left(\inf_{y \in \mathcal{S}} \|y\| \right) \|x\|_\infty,$$

since $\frac{x}{\|x\|_\infty} \in \mathcal{S}$. If we prove that $\alpha = \inf_{y \in \mathcal{S}} \|y\| > 0$ we are done as (3.4) follows.

We prove it by contradiction. Assume that $\alpha = 0$. It follows that there is a sequence $(y_n) \subset \mathcal{S}$ such that

$$\|y_n\|_\infty = 1, \qquad \text{and} \qquad \|y_n\| \to 0.$$

As the set $\mathcal{S}$ is compact (see above), there is a subsequence (y_{n_k}) such that

$$y_{n_k} \to y \in \mathcal{S} \quad \text{w.r.t } \|.\|_\infty, \qquad \text{i.e., } \lim_{k \to \infty} \|y_{n_k} - y\|_\infty = 0.$$

But, we already now that $\|y_{n_k} - y\| \leq C_0 \|y_{n_k} - y\|_\infty$, hence $y_{n_k} \to y \quad$ w.r.t $\|.\|$ as well. It follows that

$$\|y\| = \lim_{k \to \infty} \|y_{n_k}\| = 0, \qquad \text{i.e.,} \qquad y = 0.$$

But this is impossible, as $y = 0 \notin \mathcal{S}$! So we have a contradiction.

Corollary 3.7.

- *Any two norms on finite dimensional vector space are equivalent*
- *Any normed linear space* $(X, \|.\|)$ *such that dim* $X < \infty$ *is a Banach space*
- *Any subspace* $Y \subset X$ *of a n.l.s* X *such that dim* $Y < \infty$ *is closed.*

4 Geometry of Hilbert spaces

4.1 Orthogonal projection

Let H be a Hilbert space and $M \subset H$ its closed subspace. We are looking for a function $P : H \to M$ with the following properties (we set $Q = I - P$):

- $x = Px + Qx$
- $Px \perp Qx, \qquad$ i.e. $\langle Px, Qx \rangle = 0$
- $\|x\|^2 = \|Px\|^2 + \|Qx\|^2$
- $\|Qx\| = \text{dist}(x, M) = \inf_{y \in M} \|x - y\|$

Such function (if exists) is called an orthogonal projection onto M.

Case 1: $\dim M < \infty$.

In this case we can construct P directly. Take any ON basis $\{e_1, e_2, \dots, e_n\}$ for M (use G-S to achieve this). We claim that the function

$$Px = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

does the job! To see this we first establish the following:

Claim: $\|x - Px\| = \|x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i\| = \inf_{y \in M} \|x - y\|$.

Indeed:

$$\inf_{y \in M} \|x - y\| = \inf_{\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}} \left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\|$$

$$\begin{aligned} \left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\|^2 &= \left\langle x - \sum_{i=1}^{n} \alpha_i e_i, x - \sum_{i=1}^{n} \alpha_i e_i \right\rangle = \\ &= \langle x, x \rangle - \sum_{i=1}^{n} \overline{\alpha_i} \langle x, e_i \rangle - \sum_{i=1}^{n} \alpha_i \overline{\langle x, e_i \rangle} + \sum_{i,j=1}^{n} \alpha_i \overline{\alpha_j} \langle e_i, e_j \rangle \\ &= \|x\|^2 - 2 \operatorname{Re} \sum_{i=1}^{n} \overline{\alpha_i} \langle x, e_i \rangle + \sum_{i=1}^{n} |\alpha_i|^2. \end{aligned} \tag{4.5}$$

We note that

$$|\alpha_i - \langle x, e_i \rangle|^2 = |\alpha_i|^2 - 2 \operatorname{Re} \overline{\alpha_i} \langle x, e_i \rangle + |\langle x, e_i \rangle|^2,$$

hence

$$\left\| x - \sum_{i=1}^{n} \alpha_i e_i \right\|^2 = \|x\|^2 + \sum_{i=1}^{n} |\alpha_i - \langle x, e_i \rangle|^2 - \sum_{i=1}^{n} |\langle x, e_i \rangle|^2. \tag{4.6}$$

It follows that (4.6) is minimized when $\alpha_i = \langle x, e_i \rangle$, for all $i = 1, 2, \dots, n$.

The fact that P satisfies the other properties can be easily checked. For example:

$$\langle Qx, e_j \rangle = \langle x - Px, e_j \rangle = \left\langle x - \sum_{i=1}^{n} \langle x, e_i \rangle e_i, e_j \right\rangle = \langle x, e_j \rangle - \sum_{i,j=1}^{n} \langle x, e_j \rangle \langle e_i, e_j \rangle = 0$$

hence $Qx \perp e_j$ for all $j = 1, 2, \dots, n$. From this $Qx \perp Px$ as Px is a linear combination of vectors e_j.

Exercise: Find

$$\min_{a,b,c \in \mathbb{R}} \int_{-1}^{1} |x^3 - a - bx - cx^2| \, dx.$$

Hint of a solution: Denote by $M = \{\text{all polynomials of degree } \leq 2\}$. Then

$$\min_{a,b,c\in\mathbb{R}} \int_{-1}^{1} |x^3 - a - bx - cx^2| \, dx = \inf_{g\in M} \|x^3 - g\|$$

where the norm $\|.\|$ is defined using the standard inner product:

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\overline{g(x)} \, dx.$$

Hence the minimum is equal to $\|x^3 - P(x^3)\|$, where P is the orthogonal projection onto M. It remains to find and orthonomal basis of M and the problem is solved.

Case 2: dim M is arbitrary (finite or infinite).

Definition 4.1. *For any $x \in H$ we define*

$$x^\perp = \{y \in H; \langle x, y \rangle = 0\}$$

and for any subset $A \subset H$ we define

$$A^\perp = \{y \in H; \langle x, y \rangle = 0 \textit{ for all } x \in A\}.$$

Note that $A \cap A^\perp = \{0\}$. We claim that

Proposition 4.1. *$A^\perp$ is a closed subspace of H.*

Proof. Clearly,

$$A^\perp = \bigcap_{x\in A} x^\perp.$$

Hence if we prove that $x^\perp$ is a closed subspace of H we are done, as any intersection of subspaces is a subspace and any intersection of closed sets is again a closed set. We leave it to the reader that $x^\perp$ is a subspace. To see it is closed we consider a convergent sequence $(y_n)_{n\in\mathbb{N}} \subset x^\perp$ such that $y_n \to y$. We want to see that $y \in x^\perp$. We know that $\langle y_n, x \rangle = 0$, hence

$$0 = \lim_{n\to\infty} \langle y_n, x \rangle = \langle \lim_{n\to\infty} y_n, x \rangle = \langle y, x \rangle.$$

From this $y \in x^\perp$.

Remark. Notice that the function Q defined is a map $Q : H \to M^\perp$.

Definition 4.2. *Let X be a vector space and $C \subset X$. We say the set C is convex if for any two points $x, y \in C$ we have that*

$$tx + (1 - t)y \in C \qquad \textit{for all } 0 \leq t \leq 1.$$

Examples:

- Subspaces are convex sets
- In a n.l.s $B_1(0) = \{x; \|x\| < 1\}$ and $\overline{B_1(0)} = \{x; \|x\| \le 1\}$ are convex.
- If C is convex then also the set $C - x = \{y - x; y \in C\}$ is convex.

Theorem 4.2. *Let C be a nonempty close and convex set in a Hilbert space H. Then C contains a unique element of smallest norm, that is there exists unique $x_0 \in C$ such that*

$$\|x_0\| \le \|x\|, \qquad \text{for all } x \in C.$$

Proof. For any $x, y \in C$ using parallelogram law we know that

$$\left\|\frac{x}{2} - \frac{y}{2}\right\|^2 + \left\|\frac{x}{2} + \frac{y}{2}\right\|^2 = 2\left[\left\|\frac{x}{2}\right\|^2 + \left\|\frac{y}{2}\right\|^2\right].$$

As C is convex, $\frac{x}{2} + \frac{y}{2} \in C$, hence

$$\begin{aligned} \|x - y\|^2 &= 2[\|x\|^2 + \|y\|^2] - 4\left\|\frac{x}{2} + \frac{y}{2}\right\|^2 \\ &\le 2[\|x\|^2 + \|y\|^2] - 4\delta^2, \end{aligned} \tag{4.7}$$

where by δ we denote

$$\delta = \inf_{x \in C} \|x\|.$$

It follows that if there are two points $x, y \in C$ such that $\delta = \|x\| = \|y\|$, then necessary by (4.7):

$$\|x - y\|^2 \le 2[\delta^2 + \delta^2] - 4\delta^2 = 0.$$

From this $x = y$. This show uniqueness.

For existence, consider a sequence of points $(y_n)_{n\in\mathbb{N}} \subset C$ such that $\|y_n\| \to \delta$ as $n \to \infty$. We claim this sequence is Cauchy. Indeed, by (4.7) we see that

$$\|y_n - y_m\|^2 \le 2[\|y_n\|^2 + \|y_m\|^2] - 4\delta^2 \to 0 \qquad \text{as } m, n \to \infty.$$

Hence there is $y \in C$ such that $y = \lim_{n\to\infty} \|y_n\|$. It follows that $\|y\| = \lim_{n\to\infty} \|y_n\| = \delta$.

Corollary 4.3. *Let E be a nonempty closed convex set in a Hilbert space H and let $y_0 \notin E$. Then there exists unique $x_0 \in E$ such that*

$$\|x_0 - y_0\| = \inf_{x \in E} \|x - y_0\|.$$

Proof. Apply the previous theorem to the set $C = E - y_0$. Note that C is a nonempty closed convex set.

Theorem 4.4. *(On orthogonal projection) Let M be a closed subspace of a Hilbert space H. Then the following holds:*

(i) Every $x \in H$ has a unique decomposition $x = Px + Qx$ where $Px \in M$ and $Qx \in M^{\perp}$.

(ii) Px is the nearest point to x on M and Qx is the nearest point to x on $M^{\perp}$.

(iii) The mappings $P : H \to M$ and $Q : H \to M^{\perp}$ are linear.

(iv) $\|x\|^2 = \|Px\|^2 + \|Qx\|^2$.

The mappings P and Q are called orthogonal projections onto M ($M^{\perp}$), respectively.

Proof. Uniqueness: Suppose that $x = x' + y' = x'' + y''$, where $x', x'' \in M$ and $y', y'' \in M^{\perp}$. Then

$$M \ni x' - x'' = y'' - y' \in M^{\perp}.$$

But $M \cap M^{\perp} = \{0\}$, hence $x' - x'' = y'' - y' = 0$. From this $x' = x''$ and $y' = y''$ so uniqueness follows.

Existence: Choose any $x \in H$. Note that M is a closed nonempty convex set, hence by Corollary 4.3 there exists a unique point x_0 such that

$$\|x - x_0\| = \operatorname{dist}(x, M) = \inf_{y \in M} \|x - y\|.$$

We denote by Px this point x_0, i.e., we set $Px = x_0$. Clearly, $P : H \to M$. Let $Qx = x - Px$. We claim that

$$Qx \in M^{\perp}, \qquad \text{i.e.,} \qquad \langle Qx, y \rangle = 0 \quad \text{for all } y \in M.$$

Indeed, consider any $y \in M$ such that $\|y\| = 1$. Then for any $\alpha \in \mathbb{C}$:

$$\begin{aligned} \|Qx\|^2 &= \|x - Px\|^2 \le \|x - Px - \alpha y\|^2 = \langle Qx - \alpha y, Qx - \alpha y \rangle \\ &= \|Qx\|^2 - \overline{\alpha}\langle Qx, y \rangle - \alpha \langle y, Qx \rangle + |\alpha|^2. \quad \text{Hence:} \\ 0 &\le -\overline{\alpha}\langle Qx, y \rangle - \alpha \langle y, Qx \rangle + |\alpha|^2. \end{aligned} \tag{4.8}$$

Take $\alpha = \langle Qx, y \rangle$. We get that

$$0 \le -|\langle Qx, y \rangle|^2.$$

Hence $\langle Qx, y \rangle = 0$, or $Qx \in M^{\perp}$.

To see that P, Q are linear we consider any $x, y \in H$ and $\alpha, \beta \in \mathbb{C}$. We know that $x = Px + Qx$, $y = Py + Qy$ and $\alpha x + \beta y = P(\alpha x + \beta y) + Q(\alpha x + \beta y)$. This gives

$$\alpha(Px + Qx) + \beta(Py + Qy) = P(\alpha x + \beta y) + Q(\alpha x + \beta y),$$

or

$$M \ni P(\alpha x + \beta y) - \alpha Px - \beta Py = \alpha Qx + \beta Qy - Q(\alpha x + \beta y) \in M^{\perp}.$$

Again, using the fact that $M \cap M^{\perp} = \{0\}$ we see that $P(\alpha x + \beta y) - \alpha Px - \beta Py = \alpha Qx + \beta Qy - Q(\alpha x + \beta y) = 0$ from which linearity follows.

The remaining parts of the theorem are easy to show and are left to the reader.

Corollary 4.5. *If M is a closed subspace and $M \neq H$, then there is $y \neq 0$ such that*

$$\langle x, y \rangle = 0, \qquad \text{for all } x \in M.$$

Hence the subspace $M^{\perp}$ contains at least one nonzero vector.

Proof. Since $M \neq H$ there is $x \in H \setminus M$. Consider $Qx \in M^{\perp}$. As $Qx = x - Px$ we see that $Qx \neq 0$, since $x \notin M$ but $Px \in M$. So $y = Qx$ is the nonzero vector we were looking for.

4.2 Orthogonal projection on an infinite dimensional subspace

In this section we consider the following specific situation. Let $(e_n)_{n\in\mathbb{N}}$ be an infinite ON family from a Hilbert space H and let

$$M = \overline{\mathrm{lin}(\{e_n\}_{n=1}^{\infty})}.$$

By previous section we know that there exists orthogonal projection $P : H \to M$. We ask the following question:

$$\textit{Is it true that } Px = \textstyle\sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \ ?$$

To answer this question we need to deal with several issues. First we look at infinite sums in a Hilbert (or Banach) spaces.

Definition 4.3. *An infinite sequence $(x_n)_{n\in\mathbb{N}} \subset H$ in a Hilbert space is said to be summable if the sequence of partial sums*

$$s_N = \sum_{n=1}^{N} x_n$$

converges to an element $x \in H$. In this case we write:

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

Remark. Sequence $(x_n)_{n\in\mathbb{N}} \subset H$ is summable if and only if the sequence $(s_N)_{N\in\mathbb{N}}$ is Cauchy.

Lemma 4.6. *(Extension of Pythagoras theorem) Let $(u_n)_{n\in\mathbb{N}}$ be an orthogonal sequence in a Hilbert space H ($\langle u_n, u_m \rangle = 0$ for $n \neq m$). Then sequence $(u_n)_{n\in\mathbb{N}}$ is summable if and only if*

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

In this case

$$\left\| \sum_{n=1}^{\infty} u_n \right\|^2 = \sum_{n=1}^{\infty} \|u_n\|^2.$$

Proof. If $s_N = \sum_{n=1}^{N} u_n$, then

$$\|s_N - s_M\|^2 = \left\| \sum_{n=M+1}^{N} u_n \right\|^2 = \left\langle \sum_{n=M+1}^{N} u_n, \sum_{m=M+1}^{N} u_m \right\rangle = \sum_{n,m=M+1}^{N} \langle u_n, u_m \rangle$$

$$= \sum_{n=M+1}^{N} \|u_n\|^2.$$

It follows that (s_N) is a Cauchy sequence if and only if the sum $\sum_{n=1}^{\infty} \|u_n\|^2$ is finite. In that case, let $x = \sum_{n=1}^{\infty} u_n$. Since $\|s_N - x\| \to 0$ we see that

$$\|x\|^2 = \lim_{N\to\infty} \|s_N\|^2 = \lim_{N\to\infty} \sum_{n=1}^{N} \|u_n\|^2.$$

Here we used previously derived formula for $\|s_N - s_M\|^2$ with $M = 0$.

Now we are ready to answer the question posed at the beginning of this section.

Lemma 4.7. *(Extension of Bessel's inequality) Let $(e_n)_{n\in\mathbb{N}}$ be an ON family in a Hilbert space H and set*

$$M = \overline{lin(\{e_n\}_{n=1}^{\infty})}.$$

Then for a any $x \in H$

(i) $\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 \le \|x\|^2.$

(ii) Sequence $(\langle x, e_n\rangle e_n)_{n=1}^{\infty}$ is summable and its sum is

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

where $P : H \to M$ is the orthogonal projection of H onto M.

(iii) Equality holds in (i) if and only if $x \in M$. In this case

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

Proof. (i) For any integer N, Bessel's inequality (Lemma 3.1) implies that

$$\sum_{n=1}^{N} |\langle x, e_n\rangle|^2 \le \|x\|^2, \qquad \text{for all } x \in H.$$

Taking the limit $N \to \infty$ we obtain

$$\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leq \|x\|^2, \qquad \text{for all } x \in H.$$

(ii) By Lemma 4.6 applied to $u_n = \langle x, e_n \rangle e_n$ we know that $(\langle x, e_n \rangle e_n)_{n=1}^{\infty}$ is summable if and only if

$$\sum_{n=1}^{\infty} \|u_n\|^2 = \sum_{n=1}^{\infty} \|\langle x, e_n \rangle e_n\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 < \infty.$$

But we know that this holds by (i). Hence

$$\lim_{N \to \infty} \sum_{n=1}^{N} \langle x, e_n \rangle e_n = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$$

exists and belongs to M as M is closed. Let

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

and note that for all e_m we have:

$$\begin{aligned} \langle y, e_m \rangle &= \langle x, e_m \rangle - \left\langle \lim_{N \to \infty} \sum_{n=1}^{N} \langle x, e_n \rangle e_n, e_m \right\rangle \\ &= \langle x, e_m \rangle - \lim_{N \to \infty} \sum_{n=1}^{N} \langle x, e_n \rangle \langle e_n, e_m \rangle = \langle x, e_m \rangle - \langle x, e_m \rangle = 0. \end{aligned} \tag{4.9}$$

It follows that

$$\langle y, z \rangle = 0, \qquad \text{for all } z \in \operatorname{lin}(\{e_n\}_{n=1}^{\infty}),$$

hence also

$$\langle y, z \rangle = 0, \qquad \text{for all } z \in M = \overline{\operatorname{lin}(\{e_n\}_{n=1}^{\infty})}.$$

This implies that $y \in M^{\perp}$. So we can write

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n + y, \qquad \text{where: } \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \in M \text{ and } y \in M^{\perp}.$$

Hence by part (i) of Theorem 4.4 we see that $Qx = y$ and

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

(iii) The second part of Lemma 4.6 implies that

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

so if equality holds in (i) then $\|Px\|^2 = \|x\|^2$ from which $\|Qx\|^2 = \|x\|^2 - \|Px\|^2 = 0$. It follows that $Qx = 0$, hence $x = Px$. But $Px \in M$, hence $x \in M$.

Definition 4.4. *In a Hilbert space H if there is an ON sequence $(e_n)_{n\in\mathbb{N}}$ such that*

$$H = \overline{lin(\{e_n\}_{n=1}^{\infty})},$$

then the sequence $(e_n)_{n\in\mathbb{N}}$ is called an orthonormal (ON) basis of H.

Note: In this case

$$x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n, \qquad \text{for all } x \in H,$$

by the previous lemma.

Definition 4.5. *In a Hilbert space H an ON family $(u_\alpha)_{\alpha\in A} \subset H$ is called a* maximal ON family *if*

$$\langle x, u_\alpha\rangle = 0, \quad \text{for all } \alpha \in A \quad \Longrightarrow \quad x = 0.$$

Theorem 4.8. *Let $(e_n)_{n\in\mathbb{N}}$ be an ON sequence of vectors in a Hilbert space H. Then the following statements are equivalent:*

(i) $(e_n)_{n\in\mathbb{N}}$ *is maximal ON family.*

(ii) $H = \overline{lin(\{e_n\}_{n=1}^{\infty})}$.

(iii) (Parseval's equality) $\sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 = \|x\|^2$ *for all* $x \in H$.

(iv) $\sum_{n=1}^{\infty} \langle x, e_n\rangle \overline{\langle y, e_n\rangle} = \langle x, y\rangle$ *for all* $x, y \in H$.

Proof. *(i)* $\Longrightarrow$ *(ii)* Suppose that $H \neq \overline{\text{lin}(\{e_n\}_{n=1}^{\infty})}$. Then there is $u \in M^{\perp}$, $\|u\| = 1$. Hence we have that

$$\langle u, e_n\rangle = 0, \quad \text{for all } n \in \mathbb{N},$$

but $u \neq 0$, hence the family $(e_n)_{n\in\mathbb{N}}$ is not maximal - contradiction.

(ii) $\Longrightarrow$ *(iii)* Lemma 4.7.

(iii) $\Longrightarrow$ *(iv)* Polarization identity.

(iv) $\Longrightarrow$ *(iii)* Trivial, set $x = y$.

(iii) $\Longrightarrow$ *(i)* If $(e_n)_{n\in\mathbb{N}}$ is not maximal, then there is $u \neq 0$ such that $\langle u, e_n\rangle = 0$ for all n. But by (iii)

$$\|u\|^2 = \sum_{n=1}^{\infty} |\langle u, e_n\rangle|^2 = 0,$$

hence $u = 0$ which is a contradiction.

Example 1: $L^2(0,2\pi) = \overline{C[0,2\pi]}$ with inner product $\langle f, g\rangle = \int_0^{2\pi} f\overline{g}\, dx$. Then

$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos nx}{\sqrt{\pi}}\right\}_{n=1}^{\infty}$$

is an ON sequence.

Theorem 4.9. *This family is maximal.*

Theorem 4.10. *(Stone-Weierstrass) Let $X \subset \mathbb{R}^n$ be compact and $\mathcal{A} \subset C(X,\mathbb{R})$ (or $C(X,\mathbb{C})$) be an algebra of functions, that is*

$$f, g \in \mathcal{A} \quad \Longrightarrow \quad \alpha f + \beta g \in \mathcal{A}, \qquad f.g \in \mathcal{A}.$$

If $\mathcal{A}$ contains constant functions and separates points, that is

$$\forall\, x, y \in X, \quad x \neq y \quad \exists f \in \mathcal{A}: \quad f(x) \neq f(y),$$

then $\overline{\mathcal{A}} = C(X,\mathbb{R})$ (or $\overline{\mathcal{A}} = C(X,\mathbb{C})$).

Proof of Theorem 4.9. The linear span of ON family is the set of all trigonometric polynomials:

$$\mathcal{A} = \{a_0 + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx); a_n, b_n \text{ are scalars and } N \text{ an integer}\}.$$

Using Stone-Weierstrass we get that

$$\overline{\mathcal{A}} = C[0,2\pi], \qquad \text{w.r.t. } \|.\|_\infty \text{ norm.}$$

Since

$$\overline{C[0,2\pi]} = L^2(0,2\pi) \qquad \text{w.r.t. } \|.\|_2 \text{ norm}$$

and $\|.\|_2 \leq C\|.\|_\infty$ we get that

$$\overline{\mathcal{A}} = L^2(0,2\pi) \qquad \text{w.r.t. } \|.\|_2 \text{ norm.}$$

This shows maximality of the ON family.

Example 2: ℓ^2 with ON basis $(e_n)_{n\in\mathbb{N}}$, where $e_n = (0,0,\dots,1,0,\dots)$ with 1 on the n-th position.

Definition 4.6. *A Hilbert space H is called separable if there exists a countable basis of H, that is there are vectors $r_1, r_2, r_3, \dots$ from H such that*

$$H = \overline{lin\{r_1, r_2, r_3, \dots\}}.$$

Theorem 4.11. *Every separable Hilbert space is isomorphic to ℓ^2.*

Proof. Indeed, if H is a separable Hilbert space one can find (using Gram-Schmidt ON process) a countable ON basis $\{u_1, u_2, u_3, \dots\}$ of H. Then the map

$$x \in H \quad \mapsto \quad (\langle x, u_1\rangle, \langle x, u_2\rangle, \langle x, u_3\rangle, \dots) \in \ell^2$$

is the sought isomorphism.

Example 3 (Nonseparable Hilbert space): The space of almost periodic functions $AP(\mathbb{R})$ is defined as a completion of the space X:

$$X = \left\{ \sum_{n=1}^{N} c_j u_{s_j}(t);\ c_j \in \mathbb{C} \text{ and } s_j \in \mathbb{R} \right\},$$

where

$$u_s(t) = e^{ist}, \qquad \text{for } s, t \in \mathbb{R}.$$

Then $(u_s)_{s\in\mathbb{R}}$ is an ON family w.r.t the inner product

$$\langle f, g\rangle = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\overline{g(t)}\, dt, \qquad \text{for } f, g \in X. \tag{4.10}$$

Exercises: 1. Show that (4.10) defines an inner product on X.

2. Show that $(u_s)_{s\in\mathbb{R}}$ is an uncountable ON family w.r.t this inner product.

Example 4: Periodic functions $f : \mathbb{R} \to \mathbb{C}$ with period 2π ($f(t + 2\pi) = f(t)$ for all $t \in \mathbb{R}$). Alternatively, we can think about these functions as functions $\mathbb{S}^1 \to \mathbb{C}$, where $\mathbb{S}^1 = \{z \in \mathbb{C}; |z| = 1\}$. Indeed, $F : \mathbb{S}^1 \to \mathbb{C} \iff f : \mathbb{R} \to \mathbb{C}$, 2π-periodic, via

$$F(e^{it}) = f(t).$$

Pure harmonics are called the functions $e_n(t) = e^{int}$, $n \in \mathbb{Z}$ ($F_n(z) = z^n$). Let

$$X = \mathrm{lin}\{e_n; n \in \mathbb{Z}\} = \{ \sum_{n=-N}^{N} c_n e^{int}; c_n \in \mathbb{C}\}.$$

Then X is an i.p.s.(not complete !) with inner product

$$\langle f, g\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g(t)}\, dt.$$

The completion of X is called $L^2(\mathbb{S}^1)$ - the Hilbert space of square summable integrable functions. Hence $(e_n)_{n\in\mathbb{Z}}$ is an ON basis for $L^2(\mathbb{S}^1)$.

5 Linear functionals and operators.

5.1 Dual of an inner product space and Riesz representation theorem

Definition 5.1. *Let H be an inner product space. A linear map $L : H \to \mathbb{C}$ is called a* linear functional.

If in addition L is continuous, we say that L belongs to the dual of H (denoted by H^*). Hence

$$H^* = \{L : H \to \mathbb{C}; L \text{ is linear and continuous}\}.$$

Example: For any $y \in H$ consider $L_y : H \to \mathbb{C}$ to be defined by

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

Note that $L_y \in H^*$.

Theorem 5.1. *(Riesz representation theorem) Let H be a Hilbert space. For any $L \in H^*$ there exists a unique $y \in H$ such that*

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

Proof. If $L = 0$ we just take $y = 0$. If $L \neq 0$ then the set

$$M = \text{Ker } L = \{x \in H; Lx = 0\} \neq H$$

is a closed subspace of H not equal to the whole space. Hence by Corollary 4.5 there exists $z \in H$, $\|z\| = 1$ such that $z \in M^\perp$.

Consider any $x \in H$ and set

$$u = (Lx)z - (Lz)x.$$

Clearly, $Lu = 0$, hence $u \in M$. It follows that

$$0 = \langle u, z\rangle = \langle (Lx)z - (Lz)x, z\rangle = Lx\langle z, z\rangle - Lz\langle x, z\rangle.$$

But $\langle z, z\rangle = 1$, hence

$$Lx = Lz\langle x, z\rangle = \langle x, \overline{Lz}\, z\rangle.$$

Hence choosing $y = \overline{Lz}\, z$ does the job.

Uniqueness: Suppose that there are $y_1, y_2 \in H$ such that

$$Lx = \langle x, y_1\rangle = \langle x, y_2\rangle, \qquad \text{for all } x \in H.$$

Clearly this means that

$$\langle x, y_2 - y_1\rangle = 0, \qquad \text{for all } x \in H.$$

In particular, if $x = y_2 - y_1$ we get that $\|y_2 - y_1\|^2 = 0$. Hence $y_1 = y_2$.

We notice that the Riesz representation theorem implies that the dual H^* can be identified with H via the map

$$\Lambda : H \to H^* \qquad \text{defined by:} \qquad y \mapsto L_y.$$

The map Λ is 1-1 (by uniqueness part of Theorem 5.1) and onto (by the existence part of Theorem 5.1). Moreover Λ is conjugate linear, i.e.:

- $T(y_1 + y_2) = Ty_1 + Ty_2$
- $T(\alpha y) = \overline{\alpha}\, Ty$

We can also define an inner product on H^* via

$$\langle L_{y_1}, L_{y_2} \rangle = \langle y_2, y_1 \rangle.$$

Check for yourself that this makes H^* a Hilbert space. Moreover,

$$\|L_y\| = \|\Lambda y\| = \|y\|, \qquad \text{for all } y \in H,$$

hence the map Λ is an *isometry.*

Proposition 5.2. *Let H be a Hilbert space and $L_y : H \to \mathbb{C}$ the linear functional $x \mapsto \langle x, y \rangle$. Then*

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

Proof. Let $S = \sup_{\|x\| \le 1} |\langle x, y \rangle|$. Since $\|L_y\| = \|y\|$ it suffices to show that $S = \|y\|$. Indeed,

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

by Cauchy-Schwartz. On the other hand,

$$S = \sup_{\|x\| \le 1} |\langle x, y \rangle| \ge \left| \left\langle \frac{y}{\|y\|}, y \right\rangle \right| = \|y\|.$$

5.2 Bounded linear maps

The Proposition 5.2 motivates the following generalization to linear maps.

Definition 5.2. *A linear map $L : X \to Y$ between two n.l.s $(X, \|.\|_X)$ and $(Y, \|.\|_Y)$ is said to be* bounded *if the number*

$$\|L\| = \sup_{\|x\|_X \le 1} \|Lx\|_Y$$

is finite. The number $\|L\|$ is called the operator norm of L.

Remark. If we denote by $\overline{B_r(a)} = \{x : \|x - a\| \leq r\}$ the closed ball of radius r with centre at a, then L is bounded if

$$L(\overline{B_1(0)}) \subset \overline{B_R(0)}, \qquad \text{where } R = \|L\|.$$

Lemma 5.3. *Let $L : X \to Y$ be a linear map between n.l.s. Then the following statements are equivalent:*

(i) L is bounded

(ii) L is continuous

(iii) L is continuous at a point in X.

Proof. $(i) \implies (ii)$. We have:

$$\|Lx_1 - Lx_2\| = \|L(x_1 - x_2)\| = \left\| L\left(\frac{x_1 - x_2}{\|x_1 - x_2\|}\right)\right\| \|x_1 - x_2\| \leq \|L\|\|x_1 - x_2\|.$$

Hence if $x_n \to x$ in X (i.e. $\|x_n - x\| \to 0$), then

$$\|Lx_n - Lx\| \leq \|L\|\|x_n - x\| \to 0, \qquad \text{as } n \to \infty.$$

This shows that L is continuous at x for all $x \in X$.

$(ii) \implies (iii)$. Trivial.

$(iii) \implies (i)$. Suppose that L is continuous at $x_0 \in X$, i.e., for all $\varepsilon > 0$ there is $\delta > 0$ such that

$$\|x - x_0\| < \delta \quad \implies \quad \|Lx - Lx_0\| < \varepsilon.$$

Hence if $\|x\| < \delta$, then

$$\|Lx\| = \|L(x_0 + x) - Lx_0\| < \varepsilon.$$

It follows that for any $x \in X$, $\|x\| \leq 1$ we have that $\|\frac{\delta}{2}x\| < \delta$, so by the previous line

$$\tfrac{\delta}{2}\|Lx\| = \|L(\tfrac{\delta}{2}x)\| < \varepsilon.$$

Hence

$$\|Lx\| \leq \frac{2\varepsilon}{\delta}, \qquad \text{for all } \|x\| \leq 1.$$

So L is bounded and its operator norm is at most $\frac{2\varepsilon}{\delta}$.

We introduce the following notation. By $\mathcal{L}(X, Y)$ we denote the set of all bounded linear operators from n.l.s. X to Y. Hence:

$$\mathcal{L}(X, Y) = \{L : X \to Y;\ L \text{ is a bounded and linear operator from } X \to Y\}.$$

We denote by $\mathcal{L}(X) = \mathcal{L}(X, X)$. Notice that

$$X^* = \mathcal{L}(X, \mathbb{C}).$$

Proposition 5.4. *$\mathcal{L}(X, Y)$ is a normed linear space with norm*

$$\|L\| = \sup_{\|x\| \le 1} \|Lx\|, \qquad \text{for } L \in \mathcal{L}(X, Y).$$

Proof. Exercise.

Proposition 5.5. *If X is a n.l.s and Y a Banach space, then $\mathcal{L}(X, Y)$ is a Banach space.*

Proof. Exercise.

Corollary 5.6. *For any n.l.s X, the dual $X^* = \mathcal{L}(X, \mathbb{C})$ is a Banach space.*

Examples: 1. $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^m) = M_{m,n}(\mathbb{C})$ where $M_{m,n}(\mathbb{C})$ is the set of all $m \times n$ complex matrices. Indeed, every matrix $A \in M_{m,n}(\mathbb{C})$ corresponds to one linear operator $\mathbb{C}^n \to \mathbb{C}^m$ defined by

$$z \mapsto Az, \qquad \text{where } z \text{ is the column vector with } n \text{ entries.}$$

Using Cauchy-Schwartz inequality one can show that

$$\|Az\|_2 \le \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2} \|z\|_2,$$

hence

$$\|A\| \le \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}.$$

2. Shift operators on ℓ^p, $1 \le p \le \infty$. For $x = (x_1, x_2, x_3, \dots) \in \ell^p$ we define the left-shift Sx by

$$Sx = (x_2, x_3, x_4, \dots)$$

and the right-shift Rx by

$$Rx = (0, x_1, x_2, x_3, x, \dots).$$

Clearly,

$$\|Sx\|_p = \left(\sum_{n=2}^{\infty} \|x_n\|^p \right)^{1/p} \le \left(\sum_{n=1}^{\infty} \|x_n\|^p \right)^{1/p} = \|x\|_p.$$

So, $\|S\| \le 1$. Since $\|Se_2\| = \|e_1\| = 1 = \|e_2\|$ it follows that $\|S\| = 1$. Similarly, $\|R\| = 1$, in fact R is an isometry, i.e., $\|Rx\| = \|x\|$ for all $x \in \ell^p$. Also S is not 1-1 but is onto, R is 1-1 but not onto.

3. Integral operators. Let $X = (C[0,1], \|.\|_p)$ and K be a continuous function of 2 variables, i.e., $K \in C([0,1] \times [0,1])$. Consider

$$Tf(s) = \int_0^1 K(s,t) f(t)\, dt, \qquad \text{for all } s \in [0,1].$$

Clearly, $T : X \to X$ is linear. *What is its norm ?*

$p = \infty$.

$$|Tf(s)| \le \int_0^1 |K(s,t)||f(t)|\,dt \le \left(\int_0^1 |K(s,t)|\,dt\right) \|f\|_\infty.$$

Hence

$$\|Tf\| = \sup_{s\in[0,1]} |Tf(s)| \le \left(\sup_{s\in[0,1]} \int_0^1 |K(s,t)|\,dt\right) \|f\|_\infty.$$

If follows that

$$\|T\| \le \sup_{s\in[0,1]} \int_0^1 |K(s,t)|\,dt.$$

Exercise: Prove that $\|T\| = \sup_{s\in[0,1]} \int_0^1 |K(s,t)|\,dt$.

$p = 1$.

$$|Tf(s)| \le \int_0^1 |K(s,t)||f(t)|\,dt,$$

hence

$$\|Tf\|_1 \le \int_0^1 \int_0^1 |K(s,t)||f(t)|\,dt\,ds = \int_0^1 \left(\int_0^1 |K(s,t)|\,ds\right) |f(t)|\,dt$$

$$\le \left(\sup_{t\in[0,1]} \int_0^1 |K(s,t)|\,ds\right) \int_0^1 |f(t)|\,dt = \left(\sup_{t\in[0,1]} \int_0^1 |K(s,t)|\,ds\right) \|f\|_1.$$

If follows that

$$\|T\| \le \sup_{t\in[0,1]} \int_0^1 |K(s,t)|\,ds.$$

Exercise: Prove that $\|T\| = \sup_{t\in[0,1]} \int_0^1 |K(s,t)|\,ds$.

$p = 2$.

$$|Tf(s)|^2 \le \left(\int_0^1 |K(s,t)||f(t)|\,dt\right)^2 \le \int_0^1 |K(s,t)|^2 dt \times \int_0^1 |f(t)|^2 dt,$$

by the Cauchy-Schwartz inequality. Hence

$$\|Tf\|_2 = \left(\int_0^1 |Tf(s)|^2 ds\right)^{1/2} \le \left(\int_0^1 \int_0^1 |K(s,t)|^2 ds\,dt\right)^{1/2} \|f\|_2.$$

If follows that

$$\|T\| \le \left(\int_0^1 \int_0^1 |K(s,t)|^2 ds\,dt\right)^{1/2}.$$

Exercise: Find T such that $\|T\| < \left(\int_0^1 \int_0^1 |K(s,t)|^2 ds\,dt\right)^{1/2}$, showing that here we do not have always equality.

5.3 Extensions of linear operators

In the last example - integral operators we defined them on the space $X = (C[0,1], \|.\|_p)$. The problem with this space is that for $p < \infty$ it is not a Banach space. Therefore it is a natural thing to ask if we can instead consider them on the $L^p(0,1)$ space (which was defined as a completion of continuous function w.r.t $\|.\|_p$ norm). This is indeed the case

Proposition 5.7. *Let $(X, \|.\|)$ be a n.l.s and $(Y, \|.\|)$ a Banach space. Suppose that $M \subset X$ is a subspace such that $\overline{M} = X$. Let $T \in \mathcal{L}(M, Y)$. Then there exists a* **unique** *operator $\widetilde{T} \in \mathcal{L}(X, Y)$ such that*

(i) $\widetilde{T}x = Tx$ for all $x \in M$

(ii) $\|\widetilde{T}\| = \|T\|$.

The operator $\widetilde{T}$ is called the extension operator of T onto X.

Proof. Uniqueness. Suppose there are two such operators, say $T_1, T_2 \in \mathcal{L}(X, Y)$ such that

$$T_1 x = T_2 x = Tx, \qquad \text{for all } x \in M.$$

Consider $S = T_1 - T_2$. Then $Sx = 0$ for all $x \in M$. However, since S is continuous and M is a dense subset of X it follows that $Sx = 0$ for all $x \in M$. Hence $S = 0$ and $T_1 = T_2$.

Existence. We need to define $\widetilde{T}$. Take any $x \in X$. Then there exists and sequence $(x_n)_{n\in\mathbb{N}} \subset M$ such that $x_n \to x$. We define

$$\widetilde{T}x = \lim_{n\to\infty} Tx_n.$$

To see this is well defined we need to show that $(Tx_n)_{n\in\mathbb{N}} \subset M$ is a Cauchy (hence convergent - Y is a Banach space) sequence. Indeed,

$$\|Tx_n - Tx_m\| = \|T(x_n - Tx_m)\| \leq \|T\|\|x_n - x_m\| \to 0, \qquad \text{as } m, n \to \infty.$$

It is also prudent to check that $\widetilde{T}x$ does not depend on the choice of sequence $(x_n)_{n\in\mathbb{N}} \subset M$, but that trivial. We also note that

$$\|\widetilde{T}x\| = \lim_{n\to\infty} \|Tx_n\| \leq \|T\| \lim_{n\to\infty} \|x_n\| = \|T\|\|x\|.$$

This show that $\|\widetilde{T}\| \leq \|T\|$. The opposite inequality is trivial as the norm of $\widetilde{T}$ must be at least as large as that of T, since $\widetilde{T}$ is the extension of T.

Remark. Applying this theorem the the case of integral operators, we see that $M = C[0,1]$ and $X = Y = L^p(0,1)$. It follows that the integral operator

$$Tf(s) = \int_0^1 K(s,t) f(t)\, dt$$

extends to $\mathcal{L}(L^p(0,1))$, with norm estimates being same as in the Example 3 of the previous section.

What if the M is not dense in the whole space? Does there exist and extension operator? As we shall see, the answer is yes, but the operator might not be unique.

Proposition 5.8. *Let $(X, \|.\|)$ be a n.l.s and $(Y, \|.\|)$ a Banach space. Suppose that $M \subset X$ is a closed subspace of X. Let $T \in \mathcal{L}(M, Y)$. Then there exists an operator $\widetilde{T} \in \mathcal{L}(X, Y)$ such that*

(i) $\widetilde{T}x = Tx$ *for all* $x \in M$

(ii) $\|\widetilde{T}\| = \|T\|$.

The operator $\widetilde{T}$ is called the extension operator of T onto X.

Proof. (Only for the case $X = Y$ is a Hilbert space). The general proof is complicated and uses Zorn's lemma (axiom of choice). We well therefore establish only Hilbert space version of it.

Since X is a Hilbert space, it can be written as $X = M \oplus M^{\perp}$. Let $P : X \to M$ be the orthogonal projection. We define

$$\widetilde{T}x = T(Px), \qquad \text{for all } x \in X.$$

Clearly, $\widetilde{T}x = Tx$ on M as $Px = x$ on M. Also

$$\|\widetilde{T}x\| = \|T(Px)\| \leq \|T\|\|Px\| \leq \|T\|\|x\|,$$

as $\|Px\| \leq \|x\|$ for all $x \in X$. Hence $\|\widetilde{T}\| \leq \|T\|$. The opposite inequality is again trivial.

5.4 Uniform boundedness principle

Theorem 5.9. *(Banach-Steinhaus) Let $(X, \|.\|)$ be a Banach space and $(Y, \|.\|)$ a normed linear space. Assume that*

$$\{T_\alpha;\, \alpha \in A\}$$

is a collection of bounded linear operators from $\mathcal{L}(X, Y)$. Then either

(i) $\sup_{\alpha \in A} \|T_\alpha\| < \infty$, *or*

(ii) $\sup_{\alpha \in A} \|T_\alpha x\| = \infty$ *for all x belonging to a dense subset of X.*

We present this theorem without proof.

Corollary 5.10. *(Pointwise convergence of Fourier series). For any $t \in \mathbb{R}$ the Fourier series*

$$\sum_{n=-\infty}^{\infty} \langle f, e^{in\cdot} \rangle e^{int},$$

is divergent for f from a dense subset of continuous function.

Proof. For simplicity let us choose the point $t = 0$. Consider the Banach space $X = (C[-\pi,\pi], \|.\|_\infty)$. For each N we denote ny $T_N f$ the partial sum of the Fourier series

$$T_N = \sum_{n=-N}^{N} \langle f, e_n \rangle e^{in0} = \sum_{n=-N}^{N} \langle f, e_n \rangle.$$

Here $e_n(t) = e^{int}$. We see that

$$T_N f = \sum_{|n|\le N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{|n|\le N} e^{-int} \right) dt.$$

We denote by $D_N(t) = \sum_{|n|\le N} e^{-int}$. A simple calculation (it is a sum of geometric series) shows that

$$D_N(t) = \frac{\sin(N+\frac{1}{2})t}{\sin t/2}.$$

Therefore

$$\|T_n f\| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)||D_N(t)| dt \le \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| dt \right) \|f\|_\infty.$$

It follows that $T_N \in \mathcal{L}(X, \mathbb{C})$ and

$$\|T_N\| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| dt.$$

Exercise. Show that $\|T_N\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| dt$. (See the assignment sheet).

We will show that $\|T_N\| \to \infty$ as $N \to \infty$. Indeed,

$$\|T_N\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(t)| dt \ge \frac{1}{\pi} \int_0^{\pi} \frac{\sin(N+1/2)t}{t/2} dt,$$

using the fact that D_N is an even function and that $|\sin x| \le |x|$ for all x. Using the substitution $s = (N+1/2)t$ we get:

$$= \frac{2}{\pi} \int_0^{(N+1/2)\pi} \frac{|\sin s|}{s} ds \ge \frac{2}{\pi} \sum_{k=0}^{N} \int_{(k-1)\pi}^{k\pi} \frac{|\sin s|}{s} ds \ge \frac{2}{\pi^2} \sum_{k=0}^{N} \frac{1}{k} \int_{(k-1)\pi}^{k\pi} |\sin s| ds$$

$$= \frac{4}{\pi^2} \sum_{k=0}^{N} \frac{1}{k} \to \infty \qquad \text{as } N \to \infty.$$

Hence $\sup_{N\in\mathbb{N}} \|T_N\| = \infty$. Therefore, using Banach-Steinhaus theorem we see that

$$\sup_{N\in\mathbb{N}} \|T_N f\| = \sup_N \left| \sum_{n=-N}^{N} \langle f, e_n \rangle \right| = \infty,$$

for f is some dense subset of $C[-\pi,\pi]$. Hence for such f the Fourier series does not converge.

5.5 Adjoint operators.

Let X, Y be any n.l.s. and $T \in \mathcal{L}(X,Y)$. Then for any $L \in Y^*$ the composition L^oT belongs to X^*. We call this mapping an adjoint operator.

Definition 5.3. *For $T \in \mathcal{L}(X,Y)$ the linear map*

$$T^* : Y^* \to X^*$$

defined by

$$T^*L = L^oT$$

is called the adjoint of T.

Lemma 5.11. $T^* \in \mathcal{L}(Y^*, X^*)$ *and* $\|T\| = \|T\|^*$.

Proof. The fact that T^* is linear is trivial. We check boundedness of T^*.

$$\|T^*L\| = \sup_{\|x\|\leq 1} \|T^*Lx\| = \sup_{\|x\|\leq 1} \|L(Tx)\| \leq \|L\| \sup_{\|x\|\leq 1} \|Tx\| = \|L\|\|T\|,$$

for all $L \in Y^*$. Hence $\|T^*\| \leq \|T\|$.

The reverse direction relies on the following fact:

$$\|x\| = \sup_{L\in X^*,\, \|L\|\leq 1} |Lx|, \qquad \text{for all } x \in X.$$

Showing this in a Hilbert space is trivial, the Banach case is considerably harder. Taking this for granted we see that

$$\|Tx\| = \sup_{L\in Y^*,\, \|L\|\leq 1} |LTx| = \sup_{L\in Y^*,\, \|L\|\leq 1} \|T^*Lx\| \leq \sup_{\|L\|\leq 1} \|T^*L\|\|x\| = \|T^*\|\|x\|,$$

for all $x \in X$. This implies that $\|T\| \leq \|T^*\|$.

Examples. 1. Hilbert space adjoints (most important example). If $T \in \mathcal{L}(H)$, where H is a Hilbert space, then apriori adjoint is an operator on $\mathcal{L}(H^*)$. However, by Riesz representation theorem we know that H and H^* are isomorphic via the map

$$\Lambda : H \to H^*, \qquad y \mapsto L_y,$$

where $(L_y)x = \langle x, y\rangle$ for all $x \in X$. In this case it is therefore possible to this about T^* as an operator on $\mathcal{L}(H)$, not just $\mathcal{L}(H^*)$. Let us work out what is this operator. Choose any $y \in H$. Then $L_y \in H^*$, and T^*L_Y is a linear operator

$$x \mapsto \langle Tx, y\rangle.$$

By Riesz theorem there exists a unique $\in H$ such that $\langle Tx, y\rangle = \langle x, z\rangle$. This element z represents the linear map T^*L_Y and so we put $T^*z = y$. So $T^* \in \mathcal{L}(H)$ is defined by the property:

$$\langle Tx, y\rangle = \langle x, T^*y\rangle, \qquad \text{for all } x, y \in H. \tag{5.11}$$

2. $H = \mathbb{C}^n$. In this case $\mathcal{L}(\mathbb{C}^n) = M_n(\mathbb{C})$ - the set of all $n \times n$ complex matrices. Given a matrix $A \in M_n(\mathbb{C})$ let us compute its adjoint matrix defined by (5.11). We are looking for a matrix A^* such that

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

We observe that $\langle x, y\rangle = \overline{y}^T x$, hence

$$\langle Ax, y\rangle = \overline{y}^T Ax = (\overline{y}^T Ax)^T = x^T A^T \overline{y} = \overline{\overline{x}^T \overline{A}^T y} = \overline{\langle \overline{A}^T y, x\rangle} = \langle x, \overline{A}^T y\rangle.$$

We see that

$$A^* = \overline{A}^T.$$

3. Shift operators on ℓ^2. Recall that for $x = (x_1, x_2, x_3, \dots) \in \ell^2$ we defined

$$Sx = (x_2, x_3, x_4, \dots), \qquad \text{to the the left shift}$$

and

$$Rx = (0, x_1, x_2, \dots), \qquad \text{to the the right shift.}$$

What are their adjoints? A simple computation gives us for $x, y \in \ell^2$:

$$\langle Sx, y\rangle = \sum_{n=1}^{\infty} (Sx)_n \overline{y_n} = \sum_{n=1}^{\infty} x_{n+1} \overline{y_n} = \sum_{n=2}^{\infty} x_n \overline{y_{n-1}} = \sum_{n=2}^{\infty} x_n \overline{(Ry)_n} = \langle x, Ry\rangle.$$

It follows that $S^* = R$, and similarly $R^* = S$.

4. Integral operators. Recall that these operators on $L^2(0,1)$ are defined by

$$Tf(s) = \int_0^1 K(s,t) f(t), dt,$$

for a reasonable function K. What is T^* ?

$$\langle Tf, g\rangle = \int_0^1 Tf(s)\overline{g(s)} ds = \int_0^1 \int_0^1 K(s,t) f(t) \overline{g(s)} dt\, ds =$$

$$= \int_0^1 f(t) \left(\int_0^1 K(s,t) \overline{g(s)} ds \right) dt = \int_0^1 f(t) \overline{\int_0^1 \overline{K(s,t)} g(s) ds} dt = \langle f, T^* g\rangle,$$

where:

$$T^* g(s) = \int_0^1 \overline{K(t,s)} g(t)\, ds.$$

We see that T^* is also an integral operator, with kernel:

$$K^*(s,t) = \overline{K(t,s)}.$$

Definition 5.4. *Let H be a Hilbert space and $T \in \mathcal{L}(H)$. We say that T is self-adjoint if $T^* = T$.*

5.6 Compact, finite rank and Hilbert-Schmidt operators

A well-known theorem from linear algebra is saying that for a complex $n \times n$ matrix A such that $A^* = \overline{A}^T$ there exists an ON basis of $\mathbb{C}^n$ consisting of eigenvectors of A. We are interested in generalizing this theorem to infinite dimensional Hilbert spaces.

Let H be a Hilbert space and $T \in \mathcal{L}(H)$ a self-adjoint operator. Following the argument in the previous paragraph, we ask whether it is true that there exists an ON basis of H consisting of eigenvectors (or eigenfunctions) of T? A simple example shows that this in such generality is false.

Example. Let $H = L^2(0,1)$ and $Tf(x) = xf(x)$. Clearly, T is linear and bounded, since

$$\|Tf\|_2^2 = \int_0^1 x^2|f(x)|^2dx \leq \int_0^1 |f(x)|^2dx = \|x\|_2^2.$$

So $\|T\| \leq 1$. Also T is self-adjoint. Indeed,

$$\langle Tf, g\rangle = \int_0^1 Tf(x)\overline{g(x)}dx = \int_0^1 xf(x)\overline{g(x)}dx = \int_0^1 f(x)\overline{xg(x)}dx = \langle f, Tg\rangle.$$

Finally, we us find eigenfunctions of T. We need,

$$Tf(x) = xf(x) = \lambda f(x), \qquad \text{for some } \lambda \in \mathbb{C}.$$

But $(x-\lambda)f(x) = 0$ meas that $f = 0$ everywhere (except possibly the point $x = \lambda$). Hence $\|f\|_2 = 0$, so f is not an eigenfunction. This proves that T has no eigenfunction, in particular it follows that H does not have ON basis consisting of eigenfunctions of T.

It turns out that the missing ingredient for the statement to be true is that T is self-adjoint and *compact*. In what follows we define what this means.

Definition 5.5. *Let X, Y be n.l.s and $T \in \mathcal{L}(X,Y)$. We say that the operator T is* compact *if the set*

$$\overline{T(B(0,1))} = \overline{\{Tx;\ \|x\| \leq 1\}}$$

is compact in Y. We denote by $K(X,Y)$ the set of all comact operators from X to Y.

Examples: 1. Finite rank operators. We say that $T \in \mathcal{L}(X,Y)$ is a finite rank operator if

$$\dim\, T(X) = \dim(\{Tx;\ x \in X\}) < \infty.$$

Indeed, if T is a finite rank operator, then $\overline{T(B(0,1))}$ is a closed bounded set in a finite dimensional vector space $T(H)$. Hence this set is compact.

Corollary 5.12. $\mathcal{L}(X,Y) = K(X,Y)$, *provided either X or Y are finite dimensional, as in such case any bounded linear operator is of finite rank.*

2. Identity. Consider $Ix = x$ on $\mathcal{L}(X)$. Then I is compact if and only if the set

$$\overline{I(B(0,1))} = \overline{\{x;\ \|x\| \leq 1\}} = \overline{B(0,1)} \qquad \text{is compact.}$$

However, the closed unit ball is compact if and only if dim $X < \infty$.

3. Integral operators on $L^2(0,1)$. This will be shown a bit later.

Proposition 5.13. *Let X, Y be a n.l.s and $T \in \mathcal{L}(X,Y)$. Then*

- *(i) T is compact if and only if for every bounded sequence $(x_n) \subset X$ the sequence $(Tx_n) \subset Y$ has a convergent subsequence.*
- *(ii) $K(X,Y)$ is a subspace of $\mathcal{L}(X,Y)$.*
- *(iii) If also $S \in \mathcal{L}(Y,Z)$ and either T or S is compact, then T^oS is in $K(X,Z)$.*

Proof. Exercise.

If $X = Y$ is a Hilbert space another class of operators can be defined.

Definition 5.6. *Let H be Hilbert space and $T \in \mathcal{L}(H)$. We say that the operator T is a Hilbert-Schmidt operator (HS) if for some ON basis (e_n) of H the number*

$$\sum_n \|Te_n\|^2 < \infty.$$

If that happens, we denote by $\|T\|_{HS}$ the Hilbert-Schmidt norm of T defined by

$$\|T\|_{HS} = \sqrt{\sum_n \|Te_n\|^2}.$$

Note that this definition leaves open the question whether the number $\|T\|_{HS}$ depends on the choice of ON basis for H. We settle this question now.

Lemma 5.14. *1. T is HS operator if and only if T^* is an HS operator and*

$$\|T\|_{HS} = \|T^*\|_{HS}.$$

2. If (e_n) and (u_m) are two ON basis of H, then

$$\sum_n \|Te_n\|^2 = \sum_m \|Tu_m\|^2.$$

Proof. We do the following calculation. Each Te_n can be expanded in the ON basis (u_m):

$$Te_n = \sum_m \langle Te_n, u_m\rangle u_m.$$

If follows that

$$\|Te_n\|^2 = \sum_m |\langle Te_n, u_m\rangle|^2.$$

Hence

$$\sum_n \|Te_n\|^2 = \sum_n \sum_m |\langle Te_n, u_m\rangle|^2 = \sum_n \sum_m |\langle e_n, T^* u_m\rangle|^2 = \sum_m \sum_n |\langle T^* u_m, e_n\rangle|^2$$

Now we recognize that $\sum_n |\langle T^* u_m, e_n\rangle|^2$ is just $\|T^* u_m\|^2$ by Parseval's equality (Theorem 4.8). Hence

$$\sum_n \|Te_n\|^2 = \sum_m \|T^* u_m\|^2.$$

From this (i) follows by choosing $e_n = u_m$, proving that $\sum_n \|Te_n\|^2 = \sum_n \|T^* e_n\|^2$. For (ii) we see that

$$\sum_n \|Te_n\|^2 = \sum_m \|T^* u_m\|^2 = \sum_m \|Tu_m\|^2.$$

Examples: 1. Integral operators on $L^2(0,1)$. Consider

$$Tf(s) = \int_0^1 K(s,t) f(t)\, dt.$$

We have liberty to choose ON basis of $L^2(0,1)$ hence let us take $(e_n)_{n\in\mathbb{Z}}$, where $e_n(t) = e^{2\pi i t}$. Then

$$Te_n(s) = \int_0^1 K(s,t) e^{2\pi i n t} dt = \langle K_s, e_{-n}\rangle, \qquad \text{where: } K_s(t) = K(s,t).$$

It follows that

$$\|Te_n\|^2 = \int_0^1 |\langle K_s, e_{-n}\rangle|^2 ds,$$

hence

$$\|T\|_{HS}^2 = \sum_{n\in\mathbb{Z}} \|Te_n\|^2 = \int_0^1 \left(\sum_{n\in Z} |\langle K_s, e_{-n}\rangle|^2 \right) ds,$$

by exchanging the sum and integral. However again by Parseval's equality we recognize that $\sum_{n\in Z} |\langle K_s, e_{-n}\rangle|^2 = \|K_s\|^2$. This gives

$$\|T\|_{HS}^2 = \int_0^1 \|K_s\|^2 ds = \int_0^1 \int_0^1 |K(s,t)|^2 dt ds.$$

Hence
$$\|T\|_{HS} = \sqrt{\int_0^1 \int_0^1 |K(s,t)|^2 dtds}.$$

2. Operators of finite rank. Let $T \in \mathcal{L}(H)$ be of finite rank. Then T is a HS operator.

Indeed, We claim that since $T(H)$ is finite dimensional it is closed and hence we can write $H = \text{Ker } T^* \oplus T(H)$. This can be seen by proving that Ker $T^* = T(H)^\perp$. To see this consider any $x \in \text{Ker } T^*$. Then $T^*x = 0$, hence for any $y \in H$:

$$0 = \langle T^*x, y\rangle = \langle x, Ty\rangle.$$

So $x \perp Ty$ for all y, hence $x \perp T(H)$ from which $x \in T(H)^\perp$ follows. Conversely, if $x \in T(H)^\perp$, then
$$0 = \langle x, Ty\rangle = \langle T^*x, y\rangle,$$
for all $y \in H$. So $T^*x = 0$.

Having this, we may pick ON basis of H as follows. Since dim $T(H) = k < \infty$ we find any ON basis $\{e_1, e_2, \ldots, e_k\}$ of $T(H)$ and a basis $\{f_n\}_{n\in A}$ of Ker T^*. Putting these two together gives us an ON basis of H. It follows that

$$\|T\|_{HS}^2 = \|T^*\|_{HS}^2 = \sum_{i=1}^{k} \|T^*e_i\|^2 + \sum_{n\in A} \|T^* f_n\|^2 = \sum_{i=1}^{k} \|T^*e_i\|^2,$$

as $T^* f_n = 0$ for all n. However we now see that the sum of the right-hand side is a finite sum of real numbers, so it is itself a real number (finite). This gives that the HS norms of T and T^* operators are finite.

Lemma 5.15. *Let H be a Hilbert space and $T \in \mathcal{L}(H)$ a HS operator. Then*

$$\|T\| \leq \|T\|_{HS}.$$

Proof. We can write Tx as

$$Tx = \sum_n \langle Tx, e_n\rangle e_n, \quad \text{so:}$$

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

If follows that

$$\|Tx\| \leq \left(\sum_n \|T^*e_n\|^2\right)^{1/2} \|x\| = \|T^*\|_{HS}\|x\| = \|T\|_{HS}\|x\|.$$

From this $\|T\| \leq \|T\|_{HS}$.

Proposition 5.16. *Let X be a n.l.s and Y a Banach space. Then $K(X,Y)$ is a closed subspace of $\mathcal{L}(X,Y)$.*

Hence, for any sequence of operators $(T_n)_{n\in\mathbb{N}} \subset K(X,Y)$ such that $T_n \to T \in \mathcal{L}(X,Y)$ (i.e., $\|T_n - T\| \to 0$) we have that $T \in K(X,Y)$.

Proof. The proof is a classical diagonalization argument. We have to show that for any bounded sequence $(x_n) \subset X$, $\|x_n\| \le 1$ the sequence $(Tx_n) \subset Y$ has a convergent subsequence.

Let us use first the fact that T_1 is compact. It follows that one can find indices $n_1^1 < n_2^1 < n_3^1 < \dots$ such that $(T_1 x_{n_i^1})_{i\in\mathbb{N}}$ is convergent. Now we use compactness of T_2. Let $n_1^2 < n_2^2 < n_3^2 < \dots$ be a subset of $\{n_1^1, n_2^1, n_3^1, \dots\}$ such that $(T_2 x_{n_i^2})_{i\in\mathbb{N}}$ is convergent. We proceed inductively and define indices n_i^j, $i,j \in \mathbb{N}$. Now we set

$$m_i = n_i^i, \qquad i = 1,2,3,\dots.$$

Obviously, for each $j = 1,2,3,\dots$ the sequence

$$T_j x_{m_1}, T_j x_{m_2}, T_j x_{m_3}, T_j x_{m_4}, \dots \qquad \text{is convergent,}$$

as $m_1 < m_2 < m_3 < \dots$ (with exception of first few terms) is a subsequence of $n_1^j < n_2^j < n_3^j < \dots$ for which $(T_j x_{n_i^j})_{i\in\mathbb{N}}$ converges. Consider now the sequence

$$Tx_{m_1}, Tx_{m_2}, Tx_{m_3}, Tx_{m_4}, \dots.$$

We want to prove this sequence is Cauchy. Choose any $\varepsilon > 0$. We find an index j such that $\|T_j - T\| < \varepsilon/3$. Using the convergence of $(T_j x_{m_i})_{i\in\mathbb{N}}$ we see that there is $N > 0$ such that for all $k, l \ge N$

$$\|T_j x_{m_l} - T_j x_{m_k}\| < \varepsilon/3.$$

It follows that for all $k, l \ge N$

$$\|Tx_{m_l} - Tx_{m_k}\| \le \|(T - T_j)x_{m_l}\| + \|T_j x_{m_l} - T_j x_{m_k}\| + \|(T_j - T)x_{m_k}\|$$

$$\le \|T - T_j\| + \|T_j x_{m_l} - T_j x_{m_k}\| + \|T_j - T\| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

So, the sequence is Cauchy and since Y is a Banach space it converges.

Proposition 5.17. *Let H be Hilbert space and $T \in \mathcal{L}(H)$ a HS operator. Then T is compact.*

Proof. We know that

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

Consider

$$T^k x = \sum_{n=1}^{k} \langle Tx, e_n \rangle e_n, \qquad \text{for } k = 1,2,3,\dots.$$

Clearly each T^k is a finite rank (hence compact) operator, since

$$T^k(H) \subset \text{lin}\{e_1, e_2, \dots, e_k\}.$$

If we prove that $\|T - T^k\| \to 0$ we are done, as by Proposition 5.16 T is compact. Indeed, for any $x \in H$

$$\|(T - T^k)x\|^2 = \|\sum_{n=k+1}^{\infty} \langle Tx, e_n\rangle e_n\|^2 = \sum_{n=k+1}^{\infty} \|\langle Tx, e_n\rangle e_n\|^2 = \sum_{n=k+1}^{\infty} |\langle x, T^* e_n\rangle|^2 \leq$$

$$\left(\sum_{k=n+1}^{\infty} \|T^* e_n\|^2\right) \|x\|^2.$$

Hence

$$\|T - T^k\| \leq \left(\sum_{k=n+1}^{\infty} \|T^* e_n\|^2\right)^{1/2} \to 0 \qquad \text{as } k \to \infty.$$

We now look at a special class of operators called diagonal operators. These are the operators whose eigenvectors form an ON basis of appropriate Hilbert space.

Definition 5.7. *Let H be a Hilbert space and $T \in \mathcal{L}(H)$ a bounded operator. Let (e_n) be an ON basis of H. We say that T is diagonal with respect to the ON basis (e_n) if*

$$Te_n = \lambda_n e_n, \qquad \text{for all } n \in \mathbb{N},$$

and $(\lambda_n) \subset \mathbb{C}$.

Proposition 5.18. *Let H be Hilbert space, (e_n) an ON basis of H and $T : H \to H$ a linear operator such that*

$$Te_n = \lambda_n e_n, \qquad \text{for all } n \in \mathbb{N},$$

and $(\lambda_n) \subset \mathbb{C}$. Then

- *(i) T is bounded if and only if $\sup_n |\lambda_n| < \infty$. Also $\|T\| = \sup_n |\lambda_n|$. (see the last exercise sheet)*
- *(ii) T is compact if and only if $|\lambda_n| \to 0$. (see Proposition 5.19)*
- *(iii) T is HS if $\|T\|_{HS} = \sqrt{\sum_n |\lambda_n|^2} < \infty$.*
- *(iv) T is of finite rank if and only if only finitely many numbers λ_n are nonzero.*

Proposition 5.19. *Let H be Hilbert space and $T \in \mathcal{L}(H)$ be a diagonal operator w.r.t ON basis (e_n) ($Te_n = \lambda_n e_n$). Then T is compact of and only if*

$$\lim_{n\to\infty} |\lambda_n| = 0.$$

Proof. Assume first that $|\lambda_n| \to 0$. We prove that T is compact. Since

$$x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, \qquad \text{it follows that:} \qquad Tx = \sum_{n=1}^{\infty} \langle x, e_n \rangle T e_n = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n.$$

Consider

$$T^k x = \sum_{n=1}^{k} \lambda_n \langle x, e_n \rangle e_n.$$

Each of these operators is of finite rank (hence compact) as $T^k(H) \subset \operatorname{lin}\{e_1, e_2, \ldots, e_k\}$. It remains to show that $\|T - T^k\| \to 0$. Clearly,

$$\|(T - T^k)x\|^2 \leq \left\| \sum_{n=k+1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \right\|^2 = \sum_{k+1}^{\infty} \|\lambda_n \langle x, e_n \rangle e_n\|^2 = \sum_{n=k+1}^{\infty} |\lambda_n|^2 |\langle x, e_n \rangle|^2 \leq$$

$$\left(\sup_{n \geq k} |\lambda_n| \right)^2 \sum_{n=k+1}^{\infty} |\langle x, e_n \rangle|^2 \leq \left(\sup_{n \geq k} |\lambda_n| \right)^2 \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 = \left(\sup_{n \geq k} |\lambda_n| \right)^2 \|x\|^2.$$

Hence

$$\|T - T^k\| \leq \left(\sup_{n \geq k} |\lambda_n| \right) \to 0, \qquad \text{as } k \to \infty.$$

Conversely, assume that T is compact, but there is a subsequence $\|Te_{n_k}\| = |\lambda_{n_k}| \geq \varepsilon > 0$. Then using compactness of T we see that there is a convergent subsequence of the sequence $(Te_{n_k})_{k \in \mathbb{N}}$ which we denote by $(Te_{n_{k_l}})_{l \in \mathbb{N}}$ (notice the double index). Let $y = \lim_{l \to \infty} Te_{n_{k_l}}$. If follows that

$$\|y\| = \lim \|Te_{n_{k_l}}\| \geq \varepsilon.$$

On the other hand, let us expand $T^* y$ into a Fourier series. We see by Parseval's equality that

$$\|T^* y\| = \sum_{n=1}^{\infty} |\langle T^* y, e_n \rangle|^2 = \sum_{n=1}^{\infty} |\langle y, Te_n \rangle|^2,$$

hence $\langle y, Te_n \rangle \to 0$. In particular $\langle y, Te_{n_{k_l}} \rangle \to 0$ as $l \to \infty$. But

$$\|y\|^2 = \langle y, y \rangle = \lim_{l \to \infty} \langle y, Te_{n_{k_l}} \rangle = 0,$$

which is a contradiction as $\|y\| \geq \varepsilon$.

5.7 Spectral theorem for compact self-adjoint operators.

In this final section we show that compact self-adjoint operators are diagonal w.r.t some ON basis (consisting of eigenvectors of the operator). Let us note here, that there are diagonal operators that are neither compact nor self-adjoint, hence being compact and self-adjoint is a sufficient condition for being diagonal, but not necessary.

Theorem 5.20. *(Spectral theorem) Let H be a Hilbert space and $T \in \mathcal{L}(H)$ a compact, self-adjoint operator on H. Then there exist an ON basis on H, which we denote by (e_n) such that*

$$Te_n = \lambda_n e_n, \qquad \text{for all } n \text{ and some } \lambda_n \in \mathbb{R}.$$

It follows that T is a diagonal operator w.r.t (e_n). Moreover, for all $x \in H$:

$$Tx = \sum_n \lambda_n \langle x, e_n \rangle e_n.$$

We split the proof into small steps, each formulated as a lemma.

Lemma 5.21. *Let $T \in \mathcal{L}(H)$ a self-adjoint operators and $\lambda \neq \mu$ its two nonzero eigenvalues, i.e., there are $f, g \in H$, $f \neq 0$, $g \neq 0$ such that*

$$Tf = \lambda f, \qquad Tg = \mu g.$$

Then

(i) λ, μ are real numbers

(ii) f, g are orthogonal, i.e. $\langle f, g \rangle = 0$.

Proof. We have the following:

$$\lambda \langle f, g \rangle = \langle \lambda f, g \rangle = \langle Tf, g \rangle = \langle f, Tg \rangle = \langle f, \mu g \rangle = \overline{\mu} \langle f, g \rangle.$$

Hence

$$(\lambda - \overline{\mu}) \langle f, g \rangle = 0.$$

Let us choose first $\lambda = \mu$ and $f = g$. Then this gives us that $(\lambda - \overline{\lambda}) \langle f, f \rangle = 0$, from which we get that $\lambda = \overline{\lambda}$, as $f \neq 0$. So λ is real (same argument works for μ). Having this in mind, if λ, μ are different we see that

$$(\lambda - \overline{\mu}) \langle f, g \rangle = (\lambda - \mu) \langle f, g \rangle = 0,$$

from which $\langle f, g \rangle = 0$ follows as $\lambda \neq \mu$.

Lemma 5.22. *If T is self adjoint ($T = T^*$) then*

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x \rangle|. \tag{5.12}$$

Proof. We denote by m the sup

$$m = \sup_{\|x\|=1} |\langle Tx, x \rangle|. \tag{5.13}$$

We expand $\langle T(x \pm y), x \pm y \rangle$.

$$\begin{aligned} \langle T(x \pm y), x \pm y \rangle &= \langle Tx, x \rangle + \langle Ty, y \rangle \pm \langle Tx, y \rangle \pm \langle Ty, x \rangle = \\ &= \langle Tx, x \rangle + \langle Ty, y \rangle \pm 2\text{Re } \langle Tx, y \rangle. \end{aligned} \tag{5.14}$$

After we take the difference we obtain:

$$4\text{Re } \langle Tx, y\rangle = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle.$$

By (5.13): $\|\langle T(x \pm y), x \pm y\rangle\| \leq m\|x \pm y\|^2$. Hence:

$$4|\text{Re } \langle Tx, y\rangle| \leq m(\|x+y\|^2 + \|x-y\|^2) = 2m(\|x\|^2 + \|y\|)^2,$$

the last equality follows from the parallelogram law. We claim this must also hold without "Re". Indeed, $\langle Tx, y\rangle$ can be written as $e^{i\theta}|\langle Tx, y\rangle|$. If follows that

$$|\langle Tx, y\rangle| = e^{-i\theta}\langle Tx, y\rangle = \langle T(e^{-i\theta}x), y\rangle$$

Now, since $\langle T(e^{-i\theta}x), y\rangle$ is real we can use the previous estimate to get

$$|\langle Tx, y\rangle| = \langle T(e^{-i\theta}x), y\rangle \leq \frac{m}{2}(\|e^{-i\theta}x\|^2 + \|y\|^2) = \frac{m}{2}(\|x\|^2 + \|y\|^2).$$

We apply this to $y = \frac{\|x\|}{\|Tx\|}Tx$. We get

$$\|Tx\|\|x\| \leq \frac{m}{2}(\|x\|^2 + \|x\|^2) = m\|x\|^2.$$

This yields $\|Tx\| \leq m\|x\|$, or $\|T\| \leq m$. On the other hand proving that $m \leq \|T\|$ is trivial (Cauchy-Schwartz inequality).

Lemma 5.23. *If T is self adjoint ($T = T^*$) and compact then either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. It follows from (5.12) that there is a sequence of unit vectors so that

$$|\langle Tx_n, x_n\rangle| \to \|T\|.$$

Since $\langle Tx_n, x_n\rangle$ is real (as T is self-adjoint) we can remove the absolute value and assume that for a subsequence

$$\langle Tx_n, x_n\rangle \to L, \qquad \text{where} \qquad L = \pm\|T\|.$$

Now,

$$\begin{aligned} \|Tx_n - Lx_n\|^2 &= \langle Tx_n - Lx_n, Tx_n - Lx_n\rangle = \|Tx_n\|^2 - 2L\langle Tx_n, x_n\rangle + L^2\|x_n\|^2 \\ &\leq 2L^2 - 2L\langle Tx_n, x_n\rangle \to 0 \quad \text{as } n \to \infty. \end{aligned} \tag{5.15}$$

Since T is compact, there is a subsequence (which for convenience we will denote again (x_n)) such that $Tx_n \to y$. By (5.15) this implies that $Lx_n \to y$. So we see that $LTx_n \to Ty$ but also $LTx_n \to Ly$, hence $Ty = Ly$. This means y is an eigenvector, provided $y \neq 0$. Note however, that

$$\|y\| = \lim_{n\to\infty} \|Lx_n\| = |L| > 0.$$

Lemma 5.24. *If M is a closed subspace and $T(M) \subset M$, then $T^*(M^\perp) \subset M^\perp$.*

Proof. Indeed, it is enough to show that $\langle T^*x, y\rangle = 0$ for all $y \in M$, provided $x \in M^\perp$. But

$$\langle T^*x, y\rangle = \langle x, Ty\rangle$$

Here $Ty \in T(M) \subset M$ and $x \in M^\perp$, hence this inner product is indeed zero.

Lemma 5.25. *If T is self adjoint ($T = T^*$) and compact then the set of nonzero eigenvalues is either finite, or an infinite sequence with limit zero.*

Proof. If the set of eigenvalues is infinite and not tending to zero, then there is $\delta > 0$ and many distinct eigenvalues λ_n so that $|\lambda_n| \geq \delta$. If (e_n) is the sequence of associated unit vectors, then

$$\|Te_n - Te_m\|^2 = \|\lambda_n e_n - \lambda_m e_m\|^2 = |\lambda_n|^2 + |\lambda_m|^2 \geq 2\delta^2.$$

This contradicts the assumption that T is compact.

Proposition 5.26. *Let T be a compact, self-adjoint operator on a Hilbert space H. Then there is a finite or infinite sequence of eigenvectors of T with corresponding eigenvalues $\lambda_n \neq 0$ such that*

$$Tx = \sum_n \lambda_n \langle x, e_n\rangle e_n, \qquad \textit{for all } x \in H.$$

If follows that the Hilbert space H has an ON basis consisting of eigenvectors of T.

Proof. By Lemma 5.23 the number $\lambda_1 = \pm\|T\|$ is an eigenvalue of T. If $\lambda_1 = 0$ there is nothing else to do, else let e_1 be a unit length eigenvector corresponding to λ_1. Let $\mathcal{H}_2 = \text{lin } \{e_1\}^\perp$, i.e. $\mathcal{H}_2$ is the orthogonal complement of linear span of e_1.

By Lemma 5.24 $T(\mathcal{H}_2) \subset \mathcal{H}_2$, hence $T_2 = T\big|_{\mathcal{H}_2}$ is a compact self-adjoint operator on $\mathcal{H}_2$. Let $\lambda_2 = \pm\|T_2\|$ is an eigenvalue of T_2 (hence of T) with eigenvector e_2.

Inductively, we produce a sequence $e_1, e_2, \dots$ of ON eigenvectors of T so that for each n: $\mathcal{H}_n = \text{lin}\{e_1, e_2, \dots, e_n\}^\perp$ and $T_n = T\big|_{\mathcal{H}_n}$ is a compact, self-adjoint operator on $\mathcal{H}_n$ and $\lambda_n = \pm\|T_n\|$ is an eigenvalue of T_n (and hence also of T).

This construction stops if for some n: $T_n = 0$, then for all $x \in H$:

$$x - \sum_{j=1}^{n-1} \langle x, e_j\rangle e_j \in \mathcal{H}_n,$$

which implies that

$$Tx = \sum_{j=1}^{n-1} \lambda_j \langle x, e_j\rangle e_j, \qquad \text{since } T_n = 0.$$

Or, if $T_n \neq 0$ for all $n \in \mathbb{N}$, then we denote by y_n the difference

$$y_n = x - \sum_{j=1}^{n-1} \langle x, e_j\rangle e_j \in \mathcal{H}_n.$$

By the theorem of Pythagoras

$$\|x\|^2 = \|y_n\|^2 + \sum_{j=1}^{n-1} |\langle x, e_j\rangle|^2.$$

Hence $\|y_n\| \leq \|x\|$. Also

$$\|Ty_n\| = \|T_n y_n\| \leq |\lambda_n| \|x\| \to 0,$$

since $|\lambda_n| \to 0$ by Lemma 5.25. Hence

$$T\left(x - \sum_{j=1}^{n-1} \langle x, e_j\rangle e_j\right) \to 0 \qquad \text{as } n \to \infty.$$

From this the claim follows.

To see that H has an ON basis consisting of eigenvectors we need to take all vectors (e_i) we found above. Let

$$\mathcal{T} = \overline{\text{lin}\{e_1, e_2, \dots\}}.$$

Since $Tx = \sum_n \lambda_n \langle x, e_n\rangle e_n$ it follows that $Tx \in \mathcal{T}$. It also implies that if $x \in \mathcal{T}^\perp$, then $Tx = 0$, i.e. $x \in \text{Ker } T$. This means that $\text{Ker } T = \mathcal{T}^\perp$. Let (v_m) be any ON basis of $\text{Ker } T$. All v_m vectors are eigenvectors (for eigenvalue zero). It follows that

$$\{e_1, e_2, \dots\} \cup \{v_m\}$$

is an ON basis of H consisting of eigenvectors of T.