MATH 3263 HILBERT SPACES

VLADIMIR V. KISIL

ABSTRACT. This is lecture notes for the course **MATH 3263 Hilbert Spaces** at School of Mathematics of University of Leeds. They are based on the notes of Prof. Jonathan R. Partington used in the previous years. However all misprints, omissions, and errors are only my responsibility. I am very grateful to Filipa Soares de Almeida and Eric Borgnet for pointing out some of them. Please let me know if you find more.

The notes are available also for download in PDF and PostScript formats.

The suggested textbooks are [1, 2, 3, 4]. The other nice books with many interesting problems are [5, 6].

Exercises with stars **are not** a part of mandatory material but are nevertheless worth to hear about. And they are not necessarily difficult, try to solve them!

CONTENTS

List of Figures

Notations

$\mathbb{Z}_+$, $\mathbb{R}_+$ denotes non-negative integers and reals.
$x, y, z, \ldots$ denotes vectors.
$\lambda, \mu, \nu, \ldots$ denotes scalars.
$\Re z, \Im z$ stand for real and imaginary parts of a complex number z.

Part 1. Basics of Linear Spaces

A space around us could be described as a three dimensional Euclidean space. To single out a point of that space we need a fixed *frame of references* and three real numbers, which are *coordinates* of the point. Similarly to describe a pair of points from our space we could use six coordinates; for three points—nine, end so on. This makes it reasonable to consider Euclidean (linear) spaces of an arbitrary finite dimension, which are studied in the courses of linear algebra.

The basic properties of Euclidean spaces are determined by its linear and metric structures. The *linear space* (or *vector space*) structure allows to add and subtract vectors associated to points as well as to multiply vectors by real or complex numbers (scalars). The *metric space* structure assign a *distance*—non-negative real number—to a pair of points or, equivalently, defines a *length of a vector* defined by that pair.

On the other hand we could observe that many sets admit a sort of linear *and* metric structures which are linked each other. Just few among many other examples are:

- The set of convergent sequences;
- The set of continuous functions on $[0, 1]$.

It is a very *mathematical way of thinking* to declare such sets to be *spaces* and call their elements *points*.

But shall we lose all information on a particular element (e.g. a sequence $\{1/n\}$) if we represent it by a shapeless and size-less "point" without any inner configuration? Surprisingly not: all properties of an element could be now retrieved not from its *inner configuration* but from interactions with other elements through linear and metric structures.

Another surprise is that starting from our three dimensional Euclidean space and walking far away by a road of abstraction to infinite dimensional Hilbert spaces we are arriving just to yet another picture of a surrounding us space—that time on the language of *quantum mechanics*.

1. Banach spaces (basic definitions only)

The following definition generalises the notion of *distance* known from the everyday life.

Definition 1.1. A *metric* (or *distance function*) d on a set M is a function $d : M \times M \to \mathbb{R}_+$ from the set of pairs to non-negative real numbers such that:

(i) $d(x, y) \geq 0$ for all x, $y \in M$, $d(x, y) = 0$ implies $x = y$.
(ii) $d(x, y) = d(y, x)$ for all x and y in M.
(iii) $d(x, y) + d(y, z) \geq d(x, z)$ for all x, y, and z in M (*triangle inequality*).

The following notion is a useful specialisation of metric adopted to the linear structure.

Definition 1.2. Let V be a (real or complex) vector space. A *norm* on V is a real-valued function, written $\|x\|$, such that

(i) $\|x\| \geq 0$ for all $x \in V$, and $\|x\| = 0$ implies $x = 0$.
(ii) $\|\lambda x\| = |\lambda| \, \|x\|$ for all scalar λ and vector x.
(iii) $\|x + y\| \leq \|x\| + \|y\|$ (*triangle inequality*).

A vector space with a norm is called a *normed space*.

The connection between norm and metric is as follows:

Proposition 1.3. *If $\|\cdot\|$ is a norm on V, then it gives a metric on V by $d(x, y) = \|x - y\|$.*

FIGURE 1. Triangle inequality in metric (a) and normed (b) spaces.

Proof. This is a simple exercise to derive items 1.1.i–1.1.iii of Definition 1.1 from corresponding items of Definition 1.2. For example, see the Figure 1 to derive the triangle inequality. □

An important notions known from real analysis are limit and convergence. Particularly we usually wish to have enough limiting points for all "reasonable" sequences.

Definition 1.4. A sequence $\{x_k\}$ in a metric space (M, d) is a *Cauchy sequence*, if for every $\epsilon > 0$, there exists an integer n such that $k, l > n$ implies that $d(x_k, x_l) < \epsilon$.

(M, d) is a *complete metric space* if every Cauchy sequence in M converges to a limit in M.

For example, the set of integers $\mathbb{Z}$ and reals $\mathbb{R}$ with the natural distance functions are complete spaces, but the set of rationals $\mathbb{Q}$ is not. The complete normed spaces deserve a special name.

Definition 1.5. A *Banach space* is a complete normed space.

Exercise* 1.6. A convenient way to define a norm in a Banach space is as follows. The *unit ball* Uin a normed space B is the set of x such that $\|x\| \leq 1$. Prove that:

(i) U is a *convex set*, i.e. $x, y \in U$ and $\lambda \in [0, 1]$ the point $\lambda x + (1 - \lambda)y$ is also in U.
(ii) $\|x\| = \inf\{\lambda \in \mathbb{R}_+ \mid \lambda^{-1}x \in U\}$.
(iii) U is closed if and only if the space is Banach.

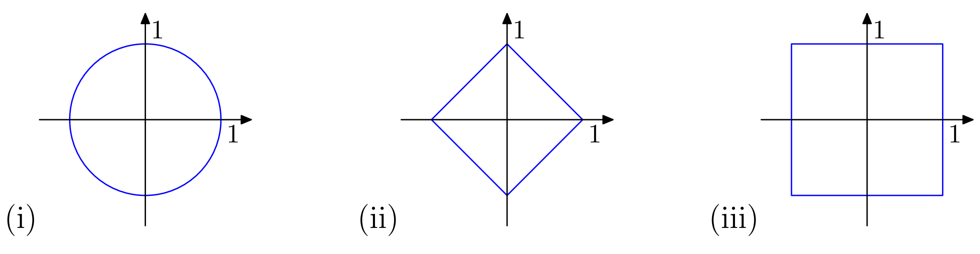

FIGURE 2. Different unit balls defining norms in $\mathbb{R}^2$ from Example 1.7.

Example 1.7. Here is some examples of normed spaces.

(i) ℓ_2^n is either $\mathbb{R}^n$ or $\mathbb{C}^n$ with norm defined by

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}.$$

(ii) ℓ_1^n is either $\mathbb{R}^n$ or $\mathbb{C}^n$ with norm defined by

$$\|(x_1, \ldots, x_n)\|_1 = |x_1| + |x_2| + \cdots + |x_n|.$$

(iii) ℓ_∞^n is either $\mathbb{R}^n$ or $\mathbb{C}^n$ with norm defined by

$$\|(x_1,\ldots,x_n)\|_\infty = \max(|x_1|,|x_2|,\cdots,|x_n|).$$

(iv) Let X be a topological space, then $C_b(X)$ is the space of continuous bounded functions $f : X \to \mathbb{C}$ with norm $\|f\|_\infty = \sup_X |f(x)|$.

(v) Let X be any set, then $\ell_\infty(X)$ is the space of all bounded (not necessarily continuous) functions $f : X \to \mathbb{C}$ with norm $\|f\|_\infty = \sup_X |f(x)|$.

All these normed spaces are also complete and thus are Banach spaces. Some more examples of both complete and incomplete spaces shall appear later.

2. Hilbert spaces

From courses of linear algebra known that the scalar product $\langle x, y\rangle = x_1y_1+\cdots+x_ny_n$ is important in a space $\mathbb{R}^n$ and defines a norm $\|x\|^2 = \langle x, x\rangle$. Here is a suitable generalisation:

Definition 2.1. A *scalar product* (or *inner product*) on a real or complex vector space V is a mapping $V \times V \to \mathbb{C}$, written $\langle x, y\rangle$, that satisfies:

(i) $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ implies $x = 0$.

(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ in complex spaces and $\langle x, y\rangle = \langle y, x\rangle$ in real ones for all x, $y \in V$.

(iii) $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$, for all x, $y \in V$ and scalar λ. (What is $\langle x, \lambda y\rangle$?).

(iv) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, for all x, y, and $z \in V$. (What is $\langle x, y + z\rangle$?).

Last two properties of the scalar product is oftenly encoded in the phrase: "it is linear in the first variable if we fix the second and anti-linear in the second if we fix the first".

Definition 2.2. An *inner product space* V is a real or complex vector space with a scalar product on it.

Example 2.3. Here is some examples of inner product spaces which demonstrate that expression $\|x\| = \sqrt{\langle x, x\rangle}$ defines a norm.

(i) The inner product for $\mathbb{R}^n$ was defined in the beginning of this section. The inner product for $\mathbb{C}^n$ is given by $\langle x, y\rangle = \sum_1^n x_j\bar{y}_j$. The norm $\|x\| = \sqrt{= \sum_1^n |x_j|^2}$ makes it ℓ_2^n from Example 1.7.i.

(ii) The extension for infinite vectors: let ℓ_2 be

$$\ell_2 = \{ \text{ sequences } \{x_j\}_1^\infty \mid \sum_1^\infty |x_j|^2 < \infty\}. \tag{2.1}$$

Let us equip this set with operations of term-wise addition and multiplication by scalars, then ℓ_2 is closed under them. Indeed it follows from the triangle inequality and properties of absolutely convergent series. From the standard Cauchy–Bunyakovskii–Schwarz inequality follows that the series $\sum_1^\infty x_j\bar{y}_j$ absolutely converges and its sum defined to be $\langle x, y\rangle$.

(iii) Let $C_b[a, b]$ be a space of continuous functions on the interval $[a, b] \in \mathbb{R}$. As we learn from Example 1.7.iv a normed space it is a normed space with the norm $\|f\|_\infty = \sup_{[a,b]} |f(x)|$. We could also define an inner product:

$$\langle f, g\rangle = \int_a^b f(x)\bar{g}(x)\, dx \quad \text{and} \quad \|f\|_2 = \left(\int_a^b |f(x)|^2\, dx\right)^{\frac{1}{2}}. \tag{2.2}$$

Now we state, probably, the most important inequality in analysis.

Theorem 2.4 (Cauchy–Schwarz–Bunyakovskii inequality). *For vectors x and y in an inner product space V let us define $\|x\| = \sqrt{\langle x, x\rangle}$ and $\|y\| = \sqrt{\langle y, y\rangle}$ then we have*

$$|\langle x, y\rangle| \leq \|x\|\, \|y\|, \tag{2.3}$$

with equality if and only if x and y are scalar multiple each other.

Proof. If $x = \lambda y$ then it is easy. Otherwise $x - \lambda y \neq 0$ for any λ, particularly for $\lambda = \langle x, y\rangle / \langle y, y\rangle$ we have:

$$0 < \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \lambda \langle y, x\rangle - \bar{\lambda}(\langle x, y\rangle - \lambda \langle y, y\rangle),$$

which yields our goal. □

Corollary 2.5. *Any inner product space is a normed space with norm $\|x\| = \sqrt{\langle x, x\rangle}$ (hence also a metric space, Prop. 1.3).*

Proof. Just to check items 1.2.i–1.2.iii from Definition 1.2. □

Again complete inner product spaces deserve a special name

Definition 2.6. A complete inner product space is *Hilbert space*.

The relations between spaces introduced so far are as follows:

$$\begin{array}{ccccc} \text{Hilbert spaces} & \Rightarrow & \text{Banach spaces} & \Rightarrow & \text{Complete metric spaces} \\ \Downarrow & & \Downarrow & & \Downarrow \\ \text{inner product spaces} & \Rightarrow & \text{normed spaces} & \Rightarrow & \text{metric spaces.} \end{array}$$

How can we tell if a given norm comes from an inner product?

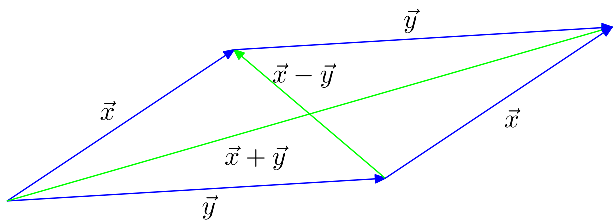

FIGURE 3. To the parallelogram identity.

Theorem 2.7 (Parallelogram identity)**.** *In an inner product space H we have for all x and $y \in H$ (see Figure 3):*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 . \tag{2.4}$$

Proof. Just by linearity of inner product:

$$\langle x + y, x + y\rangle + \langle x - y, x - y\rangle = 2\langle x, x\rangle + 2\langle y, y\rangle ,$$

because the cross terms cancel out. □

Exercise 2.8. Show that (2.4) is also a sufficient condition for a norm to arise from an inner product, namely for a norm satisfying to (2.4) the formula

$$\begin{aligned} \langle x, y\rangle &= \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2\right) \\ &= \sum_0^3 i^k \left\|x + i^k y\right\|^2 \end{aligned} \tag{2.5}$$

defines an inner product.

3. SUBSPACES

As known from the linear algebra, a *linear subspace* is a subset of a linear space is its subset, which inherits the linear structure, i.e. possibility to add vectors and multiply them by scalars. In this course we need also that subspaces inherit topological structure (coming either from a norm or an inner product) as well.

Definition 3.1. By a *subspace* of a normed space (or inner product space) we mean a linear subspace with the same norm (inner product respectively). We write $X \subset Y$ or $X \subseteq Y$.

Example 3.2. (i) $C_b(X) \subset \ell_\infty(X)$ where X is a metric space.
(ii) Any linear subspace of $\mathbb{R}^n$ or $\mathbb{C}^n$ with any norm given in Example 1.7.i–1.7.iii.
(iii) Let c_{00} be the *space of finite sequences*, i.e. all sequences (x_n) such that exist N with $x_n = 0$ for $n > N$. This is a subspace of ℓ_2 since $\sum_1^\infty |x_j|^2$ is a finite sum, so finite.

We also wish that the both inhered structures (linear and topological) should be in agreement, i.e. the subspace should be complete. Such inheritance is linked to the property be closed.

A subspace need *not* be closed—for example the sequence

$$x = (1, 1/2, 1/3, 1/4, \ldots) \in \ell_2 \qquad \text{because} \qquad \sum 1/k^2 < \infty$$

and $x_n = (1, 1/2, \ldots, 1/n, 0, 0, \ldots) \in c_{00}$ converges to x thus $x \in \overline{c_{00}} \subset \ell_2$.

Proposition 3.3. (i) *Any closed subspace of a Banach/Hilbert space is complete, hence also a Banach/Hilbert space.*
(ii) *Any complete subspace is closed.*
(iii) *The closure of subspace is again a subspace.*

Proof. (i) This is true in any metric space X: any Cauchy sequence from Y has a limit $x \in X$ belonging to $\bar{Y}$, but if Y is closed then $x \in Y$.
(ii) Let Y is complete and $x \in \bar{Y}$, then there is sequence $x_n \to x$ in Y and it is a Cauchy sequence. Then completeness of Y implies $x \in Y$.
(iii) If $x, y \in \bar{Y}$ then there are x_n and y_n in Y such that $x_n \to x$ and $y_n \to y$. From the triangle inequality:

$$\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0,$$

so $x_n + y_n \to x + y$ and $x + y \in \bar{Y}$. Similarly $x \in \bar{Y}$ implies $\lambda x \in \bar{Y}$ for any λ. □

Hence c_{00} is an *incomplete* inner product space, with inner product $\langle x, y \rangle = \sum_1^\infty x_k \bar{y}_k$ (this is a finite sum!) as it is not closed in ℓ_2.

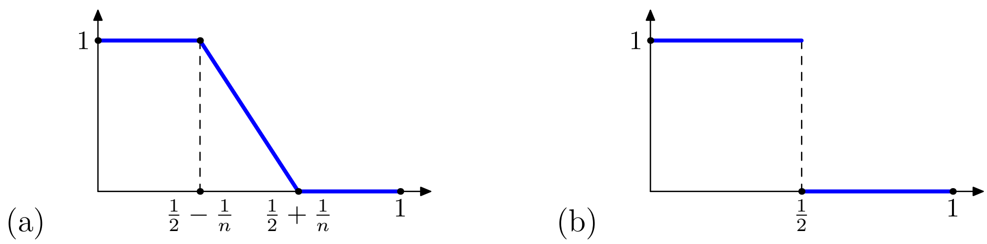

FIGURE 4. Jump function on (b) as a L_2 limit of continuous functions from (a).

Similarly $C[0,1]$ with inner product norm $\|f\| = \left(\int_0^1 |f(t)|^2\, dt\right)^{1/2}$ is incomplete—take the large space X of functions continuous on $[0,1]$ except for a possible jump at $\frac{1}{2}$ (i.e. left and right limits exists but may be unequal and $f(\frac{1}{2}) = \lim_{t \to \frac{1}{2}+} f(t)$. Then the sequence of functions defined on Figure 4(a) has the limit shown on Figure 4(b) since:

$$\|f - f_n\| = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}+\frac{1}{n}} |f - f_n|^2\, dt < \frac{2}{n} \to 0.$$

Obviously $f \in \overline{C[0,1]} \setminus C[0,1]$.

Exercise 3.4. Show alternatively that the sequence of function f_n from Figure 4(a) is a Cauchy sequence in $C[0,1]$ but has no continuous limit.

Similarly the space $C[a,b]$ is *incomplete* for any $a<b$ if equipped by the inner product and the corresponding norm:

$$\langle f,g\rangle = \int_a^b f(t)\bar{g}(t)\,dt \tag{3.1}$$

$$\|f\|_2 = \left(\int_a^b |f(t)|^2\,dt\right)^{1/2}. \tag{3.2}$$

Definition 3.5. Define a Hilbert space $L_2[a,b]$ to be the smallest complete inner product space containing space $C[a,b]$ with the restriction of inner product given by (3.1).

It is practical to realise $L_2[a,b]$ as a certain space of "functions" with the inner product defined via an integral. There are several ways to do that and we mention just two:

(i) Elements of $L_2[a,b]$ are equivalent classes of Cauchy sequences $f^{(n)}$ of functions from $C[a,b]$.

(ii) Let integration be extended from the Riemann definition to the wider *Lebesgue integration* (see MATH4011). Let L be a set of square integrable in Lebesgue sense functions on $[a,b]$ with a finite norm (3.2). Then $L_2[a,b]$ is a quotient space of L with respect to the equivalence relation $f\sim g \Leftrightarrow \|f-g\|_2=0$.

Example 3.6. Let the *Cantor function* on $[0,1]$ be defined as follows:

$$f(t)=\begin{cases}1, & t\in\mathbb{Q};\\ 0, & t\in\mathbb{R}\setminus\mathbb{Q}.\end{cases}$$

This function is *not* integrable in the Riemann sense but *does* have the Lebesgue integral. The later however is equal to 0 and as an L_2-function the Cantor function equivalent to the function identically equal to 0.

(iii) The third possibility is to map $L_2(\mathbb{R})$ onto a space of "true" functions but with an additional structure. For example, in *quantum mechanics* it is useful to work with the *Segal–Bargmann space* of analytic functions on $\mathbb{C}$ with the inner product:

$$\langle f_1,f_2\rangle = \int_{\mathbb{C}} f_1(z)\bar{f}_2(z)e^{-|z|^2}\,dz.$$

Theorem 3.7. *The sequence space ℓ_2 is complete, hence a Hilbert space.*

Proof. Take a Cauchy sequence $x^{(n)}\in\ell_2$, where $x^{(n)}=(x_1^{(n)},x_2^{(n)},x_3^{(n)},\ldots)$. Our proof will have three steps: identify the limit x; show it is in ℓ_2; show $x^{(n)}\to x$.

(i) If $x^{(n)}$ is a Cauchy sequence in ℓ_2 then $x_k^{(n)}$ is also a Cauchy sequence of numbers for any fixed k:

$$\left|x_k^{(n)}-x_k^{(m)}\right| \le \left(\sum_{k=1}^{\infty}\left|x_k^{(n)}-x_k^{(m)}\right|^2\right)^{1/2} = \left\|x^{(n)}-x^{(m)}\right\|\to 0.$$

Let x_k be the limit of $x_k^{(n)}$.

(ii) For a given $\epsilon>0$ find n_0 such that $\left\|x^{(n)}-x^{(m)}\right\|<\epsilon$ for all $n,m>n_0$. For any K and m:

$$\sum_{k=1}^{K}\left|x_k^{(n)}-x_k^{(m)}\right|^2 \le \left\|x^{(n)}-x^{(m)}\right\|^2<\epsilon^2.$$

Let $m\to\infty$ then $\sum_{k=1}^{K}\left|x_k^{(n)}-x_k\right|^2\le\epsilon^2$.

Let $K\to\infty$ then $\sum_{k=1}^{\infty}\left|x_k^{(n)}-x_k\right|^2\le\epsilon^2$. Thus $x^{(n)}-x\in\ell_2$ and because ℓ_2 is a linear space then $x=x^{(n)}-(x^{(n)}-x)$ is also in ℓ_2.

(iii) We saw above that for any $\epsilon > 0$ there is n_0 such that $\|x^{(n)} - x\| < \epsilon$ for all $n > n_0$. Thus $x^{(n)} \to x$.

Consequently ℓ_2 is complete. □

4. Linear spans

Let A be a subset (finite or infinite) of a normed space V. We may wish to upgrade it to a linear subspace.

Definition 4.1. The *linear span* of A, write $\mathrm{Lin}(A)$, is the intersection of all linear subspaces of V containing A, i.e. the smallest subspace containing A, equivalently the set of all finite linear combination of elements of A. The *closed linear span* of A write $\mathrm{CLin}(A)$ is the intersection of all *closed* linear subspaces of V containing A, i.e. the smallest *closed* subspace containing A.

Exercise* 4.2. (i) Show that if A is a subset of finite dimension space then $\mathrm{Lin}(A) = \mathrm{CLin}(A)$.
(ii) Show that for an infinite A spaces $\mathrm{Lin}(A)$ and $\mathrm{CLin}(A)$could be different. (*Hint*: use Example 3.2.iii.)

Proposition 4.3. $\overline{\mathrm{Lin}(A)} = \mathrm{CLin}(A)$.

Proof. Clearly $\overline{\mathrm{Lin}(A)}$ is a closed subspace containing A thus it should contain $\mathrm{CLin}(A)$. Also $\mathrm{Lin}(A) \subset \mathrm{CLin}(A)$ thus $\overline{\mathrm{Lin}(A)} \subset \overline{\mathrm{CLin}(A)} = \mathrm{CLin}(A)$. Therefore $\overline{\mathrm{Lin}(A)} = \mathrm{CLin}(A)$. □

Consequently $\mathrm{CLin}(A)$ is the set of all limiting points of finite linear combination of elements of A.

Example 4.4. Let $V = C[a,b]$ with the sup norm $\|\cdot\|_\infty$. Then:
$\mathrm{Lin}\{1, x, x^2, \ldots\} = \{\text{all polynomials}\}$
$\mathrm{CLin}\{1, x, x^2, \ldots\} = C[a,b]$ by the Weierstrass approximation theorem proved later.

The following simple result will be used later many times without comments.

Lemma 4.5 (about Inner Product Limit)**.** *Suppose H is an inner product space and sequences x_n and y_n have limits x and y correspondingly. Then $\langle x_n, y_n\rangle \to \langle x, y\rangle$ or equivalently:*

$$\lim_{n\to\infty} \langle x_n, y_n\rangle = \left\langle \lim_{n\to\infty} x_n, \lim_{n\to\infty} y_n \right\rangle.$$

Proof. Obviously by the Cauchy–Schwarz inequality:

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle| \\ &\leq |\langle x_n - x, y_n\rangle| + |\langle x, y_n - y\rangle| \\ &\leq \|x_n - x\|\,\|y_n\| + \|x\|\,\|y_n - y\| \to 0, \end{aligned}$$

since $\|x_n - x\| \to 0$, $\|y_n - y\| \to 0$, and $\|y_n\|$ is bounded. □

Part 2. Orthogonality

As was mentioned in the introduction the Hilbert spaces is an analog of our 3D Euclidean space and theory of Hilbert spaces similar to plane or space geometry. One of the primary result of Euclidean geometry which still survives in high school curriculum despite its continuous nasty de-geometrisation is Pythagoras' theorem based on the notion of *orthogonality*.

So far we was concerned only with distances between points. Now we would like to study angles between vectors and notably *right angles*. Pythagoras' theorem states that if the angle C in a triangle is right then $c^2 = a^2 + b^2$, see Figure 5 .

It is a very *mathematical way of thinking* to turn this *property* of right angles into their *definition*, which will work even in infinite dimensional Hilbert spaces.

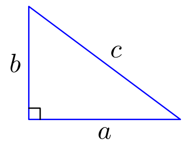

FIGURE 5. The Pythagoras' theorem $c^2 = a^2 + b^2$

5. ORTHOGONAL EXPANSIONS

In inner product spaces it is even more convenient to give a definition of orthogonality not from Pythagoras' theorem but from an equivalent property of inner product.

Definition 5.1. Two vectors x and y in an inner product space are *orthogonal* if $\langle x, y\rangle = 0$, written $x \perp y$.

An *orthogonal sequence* (or *orthogonal system*) e_n (finite or infinite) is one in which $e_n \perp e_m$ whenever $n \neq m$.

An *orthonormal sequence* (or *orthonormal system*) e_n is an orthogonal sequence with $\|e_n\| = 1$ for all n.

Exercise 5.2.
(i) Show that if $x \perp x$ then $x = 0$ and consequently $x \perp y$ for any $y \in H$.
(ii) Show that if all vectors of an orthogonal system are non-zero then they are linearly independent.

Example 5.3. These are orthonormal sequences:
(i) Basis vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ in $\mathbb{R}^3$ or $\mathbb{C}^3$.
(ii) Vectors $e_n = (0, \ldots, 0, 1, 0, \ldots)$ (with the only 1 on the nth place) in ℓ_2. (Could you see a similarity with the previous example?)
(iii) Functions $e_n(t) = 1/(\sqrt{2}\pi)e^{int}$, $n \in \mathbb{Z}$ in $C[0, 2\pi]$:

$$\langle e_n, e_m\rangle = \int_0^{2\pi} \frac{1}{2\pi} e^{int} e^{-imt} dt = \begin{cases} 1, & n = m; \\ 0, & n \neq m. \end{cases} \tag{5.1}$$

Exercise 5.4. Let A be a subset of an inner product space V and $x \perp y$ for any $y \in A$. Prove that $x \perp z$ for all $z \in \mathrm{CLin}(A)$.

Theorem 5.5 (Pythagoras'). *If $x \perp y$ then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Also if $e_1, \ldots, e_n$ is orthonormal then*

$$\left\| \sum_1^n a_k e_k \right\|^2 = \left\langle \sum_1^n a_k e_k, \sum_1^n a_k e_k \right\rangle = \sum_1^n |a_k|^2 .$$

Proof. A one-line calculation. □

The following theorem provides an important property of Hilbert spaces which will be used many times. Recall, that a subset K of a linear space V is *convex* if for all x, $y \in K$ and $\lambda \in [0, 1]$ the point $\lambda x + (1 - \lambda)y$ is also in K. Particularly any subspace is convex and any unit ball as well (see Exercise 1.6.i).

Theorem 5.6 (about the Nearest Point). *Let K be a non-empty convex closed subset of a Hilbert space H. For any point $x \in H$ there is the unique point $y \in K$ nearest to x.*

Proof. Let $d = \inf_{y\in K} d(x, y)$, where $d(x, y)$—the distance coming from the norm $\|x\| = \sqrt{\langle x, x\rangle}$ and let y_n a sequence points in K such that $\lim_{n\to\infty} d(x, y_n) = d$. Then y_n is a Cauchy sequence. Indeed from the parallelogram identity for the parallelogram generated by vectors $x - y_n$ and $x - y_m$ we have:

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

Note that $\|2x - y_n - y_m\|^2 = 4\left\|x - \frac{y_n+y_m}{2}\right\|^2 \geq 4d^2$ since $\frac{y_n+y_m}{2} \in K$ by its convexity. For sufficiently large m and n we get $\|x - y_m\|^2 \leq d + \epsilon$ and $\|x - y_n\|^2 \leq d + \epsilon$, thus $\|y_n - y_m\| \leq 4(d^2 + \epsilon) - 4d^2 = 4\epsilon$, i.e. y_n is a Cauchy sequence.

Let y be the limit of y_n, which exists by the completeness of H, then $y \in K$ since K is closed. Then $d(x, y) = \lim_{n\to\infty} d(x, y_n) = d$. This show the existence of the nearest point. Let y' be another point in K such that $d(x, y') = d$, then the parallelogram identity implies:

$$\|y - y'\|^2 = 2\|x - y\|^2 + 2\|x - y'\|^2 - \|2x - y - y'\|^2 \leq 4d^2 - 4d^2 = 0.$$

This shows the uniqueness of the nearest point. □

Exercise* 5.7. The essential rôle of the parallelogram identity in the above proof indicates that the theorem does not hold in a general Banach space.

(i) Show that in $\mathbb{R}^2$ with either norm $\|\cdot\|_1$ or $\|\cdot\|_\infty$ form Example 1.7 the nearest point could be non-unique;

(ii) Could you construct an example (in Banach space) when the nearest point does not exists?

6. Bessel's inequality

For the case then a convex subset is a subspace we could characterise the nearest point in the term of orthogonality.

Theorem 6.1 (on Perpendicular)**.** *Let M be a subspace of a Hilbert space H and a point $x \in H$ be fixed. Then $z \in M$ is the nearest point to x if and only if $x - z$ is orthogonal to any vector in M.*

Proof. Let z is the nearest point to x existing by the previous Theorem. We claim that $x - z$ orthogonal to any vector in M, otherwise there exists $y \in M$ such that $\langle x - z, y\rangle \neq 0$. Then

$$\begin{aligned} \|x - z - \epsilon y\|^2 &= \|x - z\|^2 - 2\epsilon\Re\langle x - z, y\rangle + \epsilon^2\|y\|^2 \\ &< \|x - z\|^2, \end{aligned}$$

if ϵ is chosen to be small enough and such that $\epsilon\Re\langle x - z, y\rangle$ is positive, see Figure 6(i). Therefore we get a contradiction with the statement that z is closest point to x.

On the other hand if $x - z$ is orthogonal to all vectors in H_1 then particularly $(x - z) \perp (z - y)$ for all $y \in H_1$, see Figure 6(ii). Since $x - y = (x - z) + (z - y)$ we got by the Pythagoras' theorem:

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

So $\|x - y\|^2 \geq \|x - z\|^2$ and the are equal if and only if $z = y$. □

Figure 6. (i) A smaller distance for a non-perpendicular direction; and (ii) Best approximation from a subspace

Consider now a basic case of *approximation*: let $x \in H$ be fixed and $e_1, \ldots, e_n$ be orthonormal and denote $H_1 = \mathrm{Lin}\{e_1, \ldots, e_n\}$. We could try to approximate x by a vector $y = \lambda_1 e_1 + \cdots + \lambda_n e_n \in H_1$.

Corollary 6.2. *The minimal value of $\|x - y\|$ for $y \in H_1$ is achieved when $y = \sum_1^n \langle x, e_i\rangle e_i$.*

Proof. Let $z = \sum_1^n \langle x, e_i\rangle e_i$, then $\langle x - z, e_i\rangle = \langle x, e_i\rangle - \langle z, e_i\rangle = 0$. By the previous Theorem z is the nearest point to x. □

Example 6.3. (i) In $\mathbb{R}^3$ find the best approximation to $(1,0,0)$ from the plane $V:\{x_1+x_2+x_3=0\}$. We take an orthonormal basis $e_1=(2^{-1/2},-2^{-1/2},0)$, $e_2=(6^{-1/2},6^{-1/2},-2\cdot 6^{-1/2})$ of V (Check this!). Then:

$$z=\langle x,e_1\rangle e_1+\langle x,e_2\rangle e_2=\left(\frac{1}{2},-\frac{1}{2},0\right)+\left(\frac{1}{6},\frac{1}{6},\frac{1}{3}\right)=\left(\frac{2}{3},-\frac{1}{3},-\frac{1}{3}\right).$$

(ii) In $C[0,2\pi]$ what is the best approximation to $f(t)=t$ by functions $a+be^{it}+ce^{-it}$? Let

$$e_0=\frac{1}{\sqrt{2\pi}},\qquad e_1=\frac{1}{\sqrt{2\pi}}e^{it},\qquad e_{-1}=\frac{1}{\sqrt{2\pi}}e^{-it}.$$

We find:

$$\begin{aligned}\langle f,e_0\rangle &= \int_0^{2\pi}\frac{t}{\sqrt{2\pi}}dt=\left[\frac{t^2}{2}\frac{1}{\sqrt{2\pi}}\right]_0^{2\pi}=\sqrt{2}\pi^{3/2};\\ \langle f,e_1\rangle &= \int_0^{2\pi}\frac{te^{-it}}{\sqrt{2\pi}}dt=i\sqrt{2\pi}\qquad\text{(Check this!)}\\ \langle f,e_{-1}\rangle &= \int_0^{2\pi}\frac{te^{it}}{\sqrt{2\pi}}dt=-i\sqrt{2\pi}\qquad\text{(Why we may not check this one?)}\end{aligned}$$

Then the best approximation is (see Figure 7):

$$\begin{aligned}f_0(t) &= \langle f,e_0\rangle e_0+\langle f,e_1\rangle e_1+\langle f,e_{-1}\rangle e_{-1}\\ &= \frac{\sqrt{2}\pi^{3/2}}{\sqrt{2\pi}}+ie^{it}-ie^{-it}=\pi-2\sin t.\end{aligned}$$

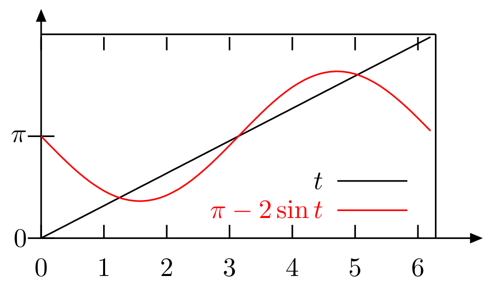

FIGURE 7. Best approximation by three trigonometric polynomials

Corollary 6.4 (Bessel's inequality). *If (e_i) is orthonormal then*

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

Proof. Let $z=\sum_1^n\langle x,e_i\rangle e_i$ then $x-z\perp e_i$ for all i therefore by Exercise 5.4 $x-z\perp z$. Hence:

$$\begin{aligned}\|x\|^2 &= \|z\|^2+\|x-z\|^2\\ &\geq \|z\|^2=\sum_{i=1}^{n}|\langle x,e_i\rangle|^2.\end{aligned}$$

□

7. The Riesz–Fischer theorem

When (e_i) is orthonormal we call $\langle x, e_n\rangle$ the nth *Fourier coefficient* of x (with respect to (e_i), naturally).

Theorem 7.1 (Riesz–Fisher). *Let $(e_n)_1^\infty$ be an orthonormal sequence in a Hilbert space H. Then $\sum_1^\infty \lambda_n e_n$ converges in H if and only if $\sum_1^\infty |\lambda_n|^2 < \infty$. In this case $\|\sum_1^\infty \lambda_n e_n\|^2 = \sum_1^\infty |\lambda_n|^2$.*

Proof. Necessity: Let $x_k = \sum_1^k \lambda_n e_n$ and $x = \lim_{k\to\infty} x_k$. So $\langle x, e_n\rangle = \lim_{k\to\infty} \langle x_k, e_n\rangle = \lambda_n$ for all n. By the Bessel's inequality for all k

$$\|x\|^2 \geq \sum_1^k |\langle x, e_n\rangle|^2 = \sum_1^k |\lambda_n|^2,$$

hence $\sum_1^k |\lambda_n|^2$ converges and the sum is at most $\|x\|^2$.

Sufficiency: Consider $\|x_k - x_m\| = \left\|\sum_m^k \lambda_n e_n\right\| = \left(\sum_m^k |\lambda_n|^2\right)^{1/2}$ for $k > m$. Since $\sum_m^k |\lambda_n|^2$ converges x_k is a Cauchy sequence in H and thus has a limit x. By the Pythagoras' theorem $\|x_k\|^2 = \sum_1^k |\lambda_n|^2$ thus for $k \to \infty$ $\|x\|^2 = \sum_1^\infty |\lambda_n|^2$ by the Lemma about inner product limit. □

Observation: the closed linear span of an orthonormal sequence in any Hilbert space looks like ℓ_2, i.e. ℓ_2 is a universal model for a Hilbert space.

By Bessel's inequality and the Riesz–Fisher theorem we know that the series $\sum_1^\infty \langle x, e_i\rangle e_i$ converges for any $x \in H$. What is its limit?

Let $y = x - \sum_1^\infty \langle x, e_i\rangle e_i$, then

$$\langle y, e_k\rangle = \langle x, e_k\rangle - \sum_1^\infty \langle x, e_i\rangle \langle e_i, e_k\rangle = \langle x, e_k\rangle - \langle x, e_k\rangle = 0 \qquad \text{for all } k. \tag{7.1}$$

Definition 7.2. An orthonormal sequence (e_i) in a Hilbert space H is *complete* if the identities $\langle y, e_k\rangle = 0$ for all k imply $y = 0$.

A complete orthonormal sequence is also called *orthonormal basis* in H.

Theorem 7.3 (on Orthonormal Basis). *Let e_i be an orthonormal basis in a Hilber space H. Then for any $x \in H$ we have*

$$x = \sum_{n=1}^\infty \langle x, e_n\rangle e_n \qquad \textit{and} \qquad \|x\|^2 = \sum_{n=1}^\infty |\langle x, e_n\rangle|^2.$$

Proof. By the Riesz–Fisher theorem, equation (7.1) and definition of orthonormal basis. □

8. Construction of Orthonormal Sequences

Natural questions are: Do orthonormal sequences always exist? Could we construct them?

Theorem 8.1 (Gram–Schmidt). *Let (x_i) be a sequence of linearly independent vectors in an inner product space V. Then there exists orthonormal sequence (e_i) such that*

$$\mathrm{Lin}\{x_1, x_2, \ldots, x_n\} = \mathrm{Lin}\{e_1, e_2, \ldots, e_n\}, \qquad \textit{for all } n.$$

Proof. We give an explicit algorithm working by induction. The *base* of induction: the first vector is $e_1 = x_1/\|x_1\|$. The *step* of induction: let e_1, e_2, …, e_n are already constructed as required. Let $y_{n+1} = x_{n+1} - \sum_{i=1}^n \langle x_{n+1}, e_i\rangle e_i$. Then by (7.1) $y_{n+1} \perp e_i$ for $i = 1, \ldots, n$. We may put $e_{n+1} = y_{n+1}/\|y_{n+1}\|$ because $y_{n+1} \neq 0$ due to linear independence of x_k's. Also

$$\begin{aligned}\mathrm{Lin}\{e_1, e_2, \ldots, e_{n+1}\} &= \mathrm{Lin}\{e_1, e_2, \ldots, y_{n+1}\}\\ &= \mathrm{Lin}\{e_1, e_2, \ldots, x_{n+1}\}\\ &= \mathrm{Lin}\{x_1, x_2, \ldots, x_{n+1}\}.\end{aligned}$$

So (e_i) are orthonormal sequence. □

Example 8.2. Consider $C[0,1]$ with the usual inner product (3.1) and apply orthogonalisation to the sequence 1, x, x^2, ... Because $\|1\| = 1$ then $e_1(x) = 1$. The continuation could be presented by the table:

$$e_1(x) = 1$$

$$y_2(x) = x - \langle x, 1\rangle\, 1 = x - \frac{1}{2}, \quad \|y_2\|^2 = \int_0^1 (x - \frac{1}{2})^2\, dx = \frac{1}{12}, \quad e_2(x) = \sqrt{12}(x - \frac{1}{2})$$

$$y_3(x) = x^2 - \left\langle x^2, 1\right\rangle 1 - \left\langle x^2, x - \frac{1}{2}\right\rangle (x - \frac{1}{2}) \cdot 12, \quad \ldots, \quad e_3 = \frac{y_3}{\|y_3\|}$$

$$\cdots\cdots\cdots$$

Example 8.3. Many famous sequences of orthogonal polynomials, e.g. Chebyshev, Legendre, Laguerre, Hermite, can be obtained by orthogonalisation of 1, x, x^2, ... with various inner products.

(i) *Legendre polynomials* in $C[-1,1]$ with inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(t)\overline{g(t)}\, dt. \tag{8.1}$$

(ii) *Chebyshev polynomials* in $C[-1,1]$ with inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(t)\overline{g(t)} \frac{dt}{\sqrt{1-t^2}} \tag{8.2}$$

(iii) *Laguerre polynomials* in the space of polynomials $P[0,\infty)$ with inner product

$$\langle f, g\rangle = \int_{0}^{\infty} f(t)\overline{g(t)} e^{-t}\, dt.$$

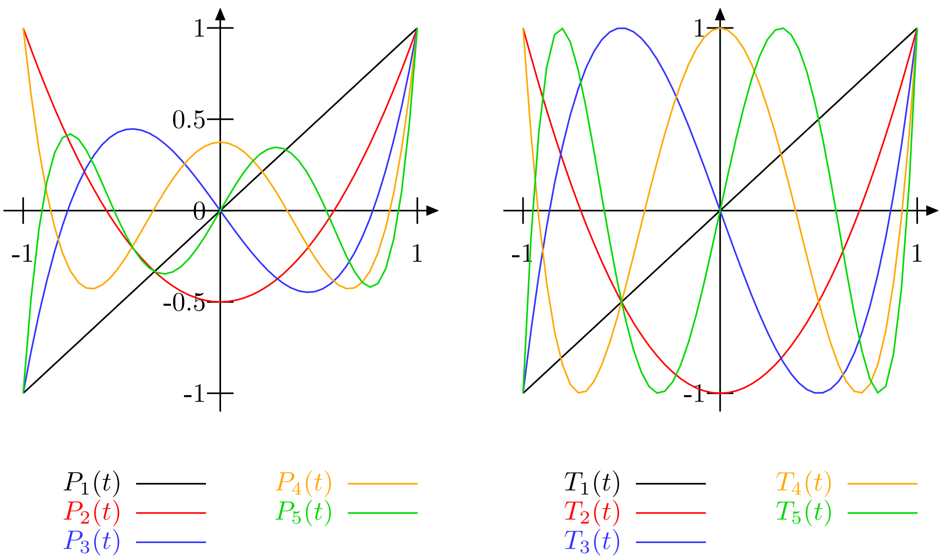

FIGURE 8. Five first Legendre P_i and Chebyshev T_i polynomials

See Figure 8 for the five first Legendre and Chebyshev polynomials. Observe the difference caused by the different inner products (8.1) and (8.2). On the other hand note the similarity in oscillating behaviour with different "frequencies".

Another natural question is: When is an orthonormal sequence complete?

Proposition 8.4. *Let (e_n) be an orthonormal sequence in a Hilbert space H. The following are equivalent:*

(i) *(e_n) is an orthonormal basis.*
(ii) *$\mathrm{CLin}((e_n)) = H$.*
(iii) *$\|x\|^2 = \sum_1^\infty |\langle x, e_n\rangle|^2$ for all $x \in H$.*

Proof. Clearly 8.4.i implies 8.4.ii because $x = \sum_1^\infty \langle x, e_n\rangle e_n$ in $\mathrm{CLin}((e_n))$ and $\|x\|^2 = \sum_1^\infty \langle x, e_n\rangle e_n$ by Theorem 7.3.

If (e_n) is *not* complete then there exists $x \in H$ such that $x \neq 0$ and $\langle x, e_k\rangle$ for all k, so 8.4.iii fails, consequently 8.4.iii implies 8.4.i.

Finally if $\langle x, e_k\rangle = 0$ for all k then $\langle x, y\rangle = 0$ for all $y \in \mathrm{Lin}((e_n))$ and moreover for all $y \in \mathrm{CLin}((e_n))$, by the Lemma on continuity of the inner product. But then $x \notin \mathrm{CLin}((e_n))$ and 8.4.ii also fails because $\langle x, x\rangle = 0$ is not possible. Thus 8.4.ii implies 8.4.i. □

Corollary 8.5. *A* separable Hilbert space *(i.e. one with a countable dense set) can be identified with either ℓ_2^n or ℓ_2, in other words it has an orthonormal basis (e_n) (finite or infinite) such that*

$$x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n \qquad \text{and} \qquad \|x\|^2 = \sum_{n=1}^{\infty} |\langle x, e_n\rangle|^2 .$$

Proof. Take a countable dense set (x_k), then $H = \mathrm{CLin}((x_k))$, delete all vectors which are a linear combinations of preceding vectors, make orthonormalisation by Gram–Schmidt the remaining set and apply the previous proposition. □

9. Orthogonal complements

Definition 9.1. Let M be a subspace of an inner product space V. The *orthogonal complement*, written $M^\perp$, of M is

$$M^\perp = \{x \in V : \langle x, m\rangle = 0 \ \forall m \in M\}.$$

Theorem 9.2. *If M is a closed subspace of a Hilbert space H then $M^\perp$ is a closed subspace too (hence a Hilbert space too).*

Proof. Clearly $M^\perp$ is a subspace of H because $x, y \in M^\perp$ implies $ax + by \in M^\perp$:

$$\langle ax + by, m\rangle = a\langle x, m\rangle + b\langle y, m\rangle = 0.$$

Also if all $x_n \in M^\perp$ and $x_n \to x$ then $x \in M^\perp$ due to inner product limit Lemma. □

Theorem 9.3. *Let M be a closed subspace of a Hilber space H. Then for any $x \in H$ there exists the unique decomposition $x = m + n$ with $m \in M$, $n \in M^\perp$ and $\|x\|^2 = \|m\|^2 + \|n\|^2$. Thus $H = M \oplus M^\perp$ and $(M^\perp)^\perp = M$.*

Proof. For a given x there exists the unique closest point m in M by the Theorem on nearest point and by the Theorem on perpendicular $(x - m) \perp y$ for all $y \in M$.

So $x = m + (x - m) = m + n$ with $m \in M$ and $n \in M^\perp$. The identity $\|x\|^2 = \|m\|^2 + \|n\|^2$ is just Pythagoras' theorem and $M \cap M^\perp = \{0\}$ because null vector is the only vector orthogonal to itself.

Finally $(M^\perp)^\perp = M$. We have $H = M \oplus M^\perp = (M^\perp)^\perp \oplus M^\perp$, for any $x \in (M^\perp)^\perp$ there is a decomposition $x = m + n$ with $m \in M$ and $n \in M^\perp$, but then n is orthogonal to itself and therefore is zero. □

Corollary 9.4 (about Orthoprojection)**.** *There is a linear map P_M from H onto M (the* orthogonal projection *or* orthoprojection*) such that*

$$P_M^2 = P_M, \qquad \ker P_M = M^\perp, \qquad P_{M^\perp} = I - P_M. \tag{9.1}$$

Proof. Let us define $P_M(x) = m$ where $x = m + n$ is the decomposition from the previous theorem. The linearity of this operator follows from the fact that both M and $M^\perp$ are linear subspaces. Also $P_M(m) = m$ for all $m \in M$ and the image of P_M is M. Thus $P_M^2 = P_M$. Also if $P_M(x) = 0$ then $x \perp M$, i.e. $\ker P_M = M^\perp$. Similarly $P_{M^\perp}(x) = n$ where $x = m + n$ and $P_M + P_{M^\perp} = I$. □

Example 9.5. Let (e_n) be an orthonormal basis in a Hilber space and let $S \subset \mathbb{N}$ be fixed. Let $M = \text{CLin}\{e_n : n \in S\}$ and $M^\perp = \text{CLin}\{e_n : n \in \mathbb{N} \setminus S\}$. Then

$$\sum_{k=1}^{\infty} a_k e_k = \sum_{k \in S} a_k e_k + \sum_{k \notin S} a_k e_k.$$

Remark 9.6. In fact there is a one-to-one correspondence between closed linear subspaces of a Hilber space H and orthogonal projections defined by identities (9.1).

Part 3. Fourier Analysis

As we saw already any separable Hilbert space posses an orthonormal basis (infinitely many of them indeed). Are they equally good? This depends from our purposes. For solution of differential equation which arose in mathematical physics (wave, heat, Laplace equations, etc.) there is a proffered choice. The fundamental formula: $\frac{d}{dx} e^{ax} = a e^{ax}$ reduces the derivative to a multiplication by a. We could benefit from this observation if the orthonormal basis will be constructed out of exponents.

10. Fourier series

Let us consider the space $L_2[-\pi, \pi]$. As we saw in Example 5.3.iii there is an orthonormal sequence $e_n(t) = (2\pi)^{-1/2} e^{int}$ in $L_2[-\pi, \pi]$. We will show that it is an orthonormal basis, i.e.

$$f(t) \in L_2[-\pi, \pi] \quad \Leftrightarrow \quad f(t) = \sum_{k=-\infty}^{\infty} \langle f, e_k \rangle \, e_k(t),$$

with convergence in L_2 norm. To do this we show that $\text{CLin}\{e_k : k \in \mathbb{Z}\} = L_2[-\pi, \pi]$.

Let $CP[-\pi, \pi]$ denote the continuous functions f on $[-\pi, \pi]$ such that $f(\pi) = f(-\pi)$. We also define f outside of the interval $[-\pi, \pi]$ by periodicity.

Lemma 10.1. *The space $CP[-\pi, \pi]$ is dense in $L_2[-\pi, \pi]$.*

Proof. Let $f \in L_2[-\pi, \pi]$. Given $\epsilon > 0$ there exists $g \in C[-\pi, \pi]$ such that $\|f - g\| < \epsilon/2$. Form continuity of g on a compact set follows that there is M such that $|g(t)| < M$ for all $t \in [-\pi, \pi]$. We can now replace g by periodic $\tilde{g}$, which coincides with g on $[-\pi, \pi - \delta]$ for an arbitrary $\delta > 0$

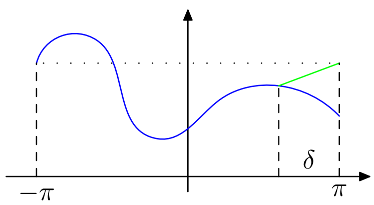

Figure 9. A modification of continuous function to periodic

and has the same bounds: $|\tilde{g}(t)| < M$, see Figure 9. Then

$$\|g - \tilde{g}\|_2^2 = \int_{\pi-\delta}^{\pi} |g(t) - \tilde{g}(t)|^2 \, dt \leq (2M)^2 \delta.$$

So if $\delta < \epsilon^2/(4M)^2$ then $\|g - \tilde{g}\| < \epsilon/2$ and $\|f - \tilde{g}\| < \epsilon$. □

Now if we could show that $\text{CLin}\{e_k : k \in \mathbb{Z}\}$ includes $CP[-\pi, \pi]$ then it also includes $L_2[-\pi, \pi]$.

Notation 10.2. Let $f \in CP[-\pi, \pi]$,write

$$f_n = \sum_{k=-n}^{n} \langle f, e_k \rangle e_k, \qquad \text{for } n = 0, 1, 2, \ldots \tag{10.1}$$

the *partial sum of the Fourier series* for f.

We want to show that $\|f - f_n\|_2 \to 0$. To this end we define nth *Fejér sum* by the formula

$$F_n = \frac{f_0 + f_1 + \cdots + f_n}{n+1}, \tag{10.2}$$

and show that

$$\|F_n - f\|_\infty \to 0.$$

Then we conclude

$$\|F_n - f\|_2 = \left(\int_{-\pi}^{\pi} |F_n(t) - f|^2 \right)^{1/2} \leq (2\pi)^{1/2} \|F_n - f\|_\infty \to 0.$$

Since $F_n \in \mathrm{Lin}((e_n))$ then $f \in \mathrm{CLin}((e_n))$ and hence $f = \sum_{-\infty}^{\infty} \langle f, e_k \rangle e_k$.

Remark 10.3. It is **not** always true that $\|f_n - f\|_\infty \to 0$ even for $f \in CP[-\pi, \pi]$.

Exercise 10.4. Find an example illustrating the above Remark.

11. Fejer's theorem

Proposition 11.1 (Fejér, age 19)**.** *Let $f \in CP[-\pi, \pi]$. Then*

$$F_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) K_n(x - t)\, dt, \qquad \textit{where} \tag{11.1}$$

$$K_n(t) = \frac{1}{n+1} \sum_{k=0}^{n} \sum_{m=-k}^{k} e^{imt}, \tag{11.2}$$

is the Fejér kernel*.*

Proof. From notation (10.1):

$$\begin{aligned} f_k(x) &= \sum_{m=-k}^{k} \langle f, e_m \rangle e_m(x) \\ &= \sum_{m=-k}^{k} \int_{-\pi}^{\pi} f(t) \frac{e^{-imt}}{\sqrt{2\pi}}\, dt\, \frac{e^{imx}}{\sqrt{2\pi}} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{m=-k}^{k} e^{im(x-t)}\, dt. \end{aligned}$$

Then from (10.2):

$$\begin{aligned} F_n(x) &= \frac{1}{n+1}\sum_{k=0}^{n} f_k(x) \\ &= \frac{1}{n+1}\frac{1}{2\pi}\sum_{k=0}^{n}\int_{-\pi}^{\pi} f(t)\sum_{m=-k}^{k} e^{im(x-t)}\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\frac{1}{n+1}\sum_{k=0}^{n}\sum_{m=-k}^{k} e^{im(x-t)}\,dt, \end{aligned}$$

which finishes the proof. □

Lemma 11.2. *The Fejér kernel is 2π-periodic, $K_n(0) = n+1$ and:*

$$K_n(t) = \frac{1}{n+1}\frac{\sin^2\frac{(n+1)t}{2}}{\sin^2\frac{t}{2}}, \qquad \text{for } t \notin 2\pi\mathbb{Z}. \tag{11.3}$$

Proof. Let $z = e^{it}$, then:

$$\begin{aligned} K_n(t) &= \frac{1}{n+1}\sum_{k=0}^{n}(z^{-k} + \cdots + 1 + z + \cdots + z^k) \\ &= \frac{1}{n+1}\sum_{j=-n}^{n}(n+1-|j|)z^j, \end{aligned}$$

by switch from counting in rows to counting in columns in Table 1. Let $w = e^{it/2}$, i.e. $z = w^2$, then

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & z^{-1} & 1 & z & & \\ & z^{-2} & z^{-1} & 1 & z & z^2 & \\ \iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

TABLE 1. Counting powers in rows and columns

$$\begin{aligned} K_n(t) &= \frac{1}{n+1}(w^{-2n} + 2w^{-2n+2} + \cdots + (n+1) + nw^2 + \cdots + w^{2n}) \\ &= \frac{1}{n+1}(w^{-n} + w^{-n+2} + \cdots + w^{n-2} + w^n)^2 \\ &= \frac{1}{n+1}\left(\frac{w^{-n-1} - w^{n+1}}{w^{-1} - w}\right)^2 \qquad \text{Could you sum a geometric progression?} \\ &= \frac{1}{n+1}\left(\frac{2i\sin\frac{(n+1)t}{2}}{2i\sin\frac{t}{2}}\right)^2, \end{aligned} \tag{11.4}$$

if $w \neq \pm 1$. For the value of $K_n(0)$ we substitute $w = 1$ into (11.4). □

The first eleven Fejér kernels are shown on Figure 10, we could observe that:

Lemma 11.3. *Fejér's kernel has the following properties:*

(i) $K_n(t) \geq 0$ *for all* $t \in \mathbb{R}$ *and* $n \in \mathbb{N}$.

(ii) $\int_{-\pi}^{\pi} K_n(t)\,dt = 2\pi$.

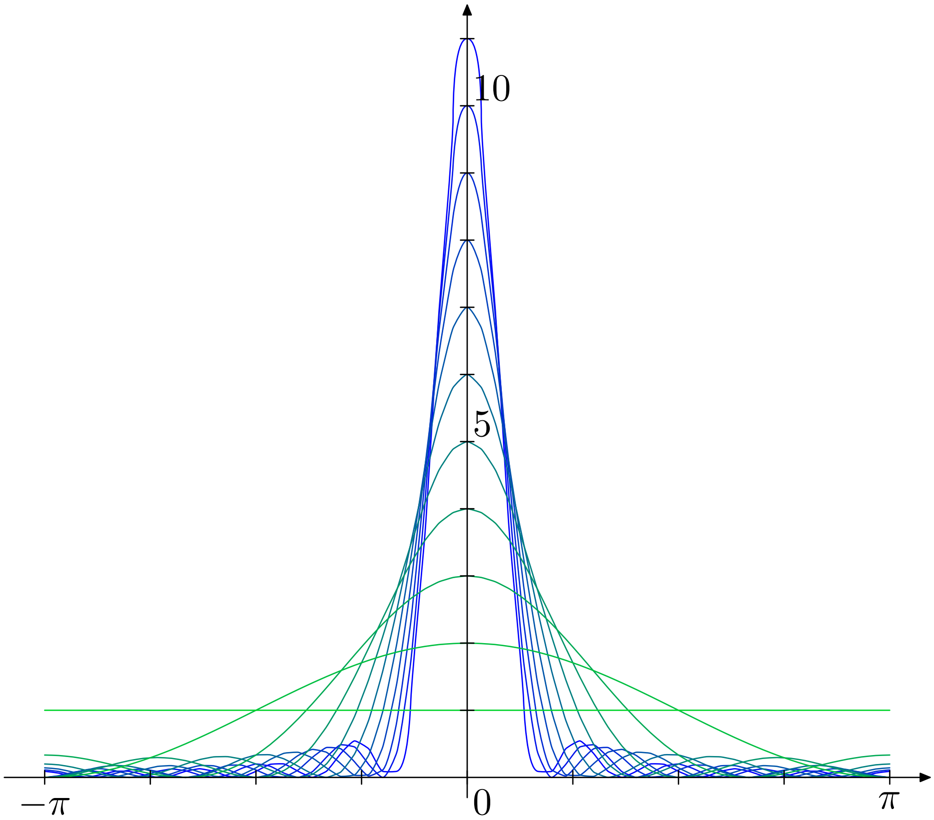

FIGURE 10. This is a plot of Fejér kernels with the parameter m running from 0 to 10.

(iii) *For any* $\delta \in (0, \pi)$

$$\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} K_n(t)\,dt \to 0 \qquad \textit{as} \qquad n \to \infty.$$

Proof. The first property immediately follows from the explicit formula (11.3). In contrast the second property is easier to deduce from expression with double sum (11.2):

$$\begin{aligned}\int_{-\pi}^{\pi} K_n(t)\,dt &= \int_{-\pi}^{\pi} \frac{1}{n+1} \sum_{k=0}^{n} \sum_{m=-k}^{k} e^{imt}\,dt \\ &= \frac{1}{n+1} \sum_{k=0}^{n} \sum_{m=-k}^{k} \int_{-\pi}^{\pi} e^{imt}\,dt \\ &= \frac{1}{n+1} \sum_{k=0}^{n} 2\pi \\ &= 2\pi,\end{aligned}$$

since the formula (5.1).

Finally if $|t| > \delta$ then $\sin^2(t/2) \geq \sin^2(\delta/2) > 0$ by monotonicity of sinus on $[0, \pi/2]$, so:

$$0 \leq K_n(t) \leq \frac{1}{(n+1)\sin^2(\delta/2)}$$

implying:

$$0 \leq \int_{\delta \leq |t| \leq \pi} K_n(t)\,dt \leq \frac{1(\pi - \delta)}{(n+1)\sin^2(\delta/2)} \to 0 \quad \text{as} \quad n \to 0.$$

Therefore the third property follows from the squeeze rule. □

Theorem 11.4 (Fejér Theorem). *Let $f \in CP[-\pi,\pi]$. Then its Fejér sums F_n (10.2) converges in supremum norm to f on $[-\pi,\pi]$ and hence in L_2 norm as well.*

Proof. Idea of the proof: if in the formula (11.1)

$$F_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)K_n(x-t)\,dt,$$

t is long way from x, K_n is small (see Lemma 11.3 and Figure 10), for t near x, K_n is big with total "weight" 2π, so the weighted average of $f(t)$ is near $f(x)$.

Here are details. Using property 11.3.ii and periodicity of f and K_n we could express trivially

$$f(x) = f(x)\frac{1}{2\pi}\int_{x-\pi}^{x+\pi} K_n(x-t)\,dt = \frac{1}{2\pi}\int_{x-\pi}^{x+\pi} f(x)K_n(x-t)\,dt.$$

Similarly we rewrite (11.1) as

$$F_n(x) = \frac{1}{2\pi}\int_{x-\pi}^{x+\pi} f(t)K_n(x-t)\,dt,$$

then

$$\begin{aligned} |f(x)-F_n(x)| &= \frac{1}{2\pi}\left|\int_{x-\pi}^{x+\pi}(f(x)-f(t))K_n(x-t)\,dt\right| \\ &\le \frac{1}{2\pi}\int_{x-\pi}^{x+\pi}|f(x)-f(t)|\,K_n(x-t)\,dt. \end{aligned}$$

Given $\epsilon>0$ split into three intervals: $I_1=[x-\pi,x-\delta]$, $I_2=[x-\delta,x+\delta]$, $I_3=[x+\delta,x+\pi]$, where δ is chosen such that $|f(t)-f(x)|<\epsilon/2$ for $t\in I_2$, which is possible by continuity of f. So

$$\frac{1}{2\pi}\int_{I_2}|f(x)-f(t)|\,K_n(x-t)\,dt \le \frac{\epsilon}{2}\frac{1}{2\pi}\int_{I_2}K_n(x-t)\,dt < \frac{\epsilon}{2}.$$

And

$$\begin{aligned} \frac{1}{2\pi}\int_{I_1\cup I_3}|f(x)-f(t)|\,K_n(x-t)\,dt &\le 2\|f\|_\infty\frac{1}{2\pi}\int_{I_1\cup I_3}K_n(x-t)\,dt \\ &= \frac{\|f\|_\infty}{\pi}\int_{\delta<|u|<\pi}K_n(u)\,du \\ &< \frac{\epsilon}{2}, \end{aligned}$$

if n is sufficiently large due to property 11.3.iii of K_n. Hence $|f(x)-F_n(x)|<\epsilon$ for a large n independent of x. □

We almost finished the demonstration that $e_n(t)=(2\pi)^{-1/2}e^{int}$ is an orthonormal basis of $L_2[-\pi,\pi]$:

Corollary 11.5 (Fourier series). *Let $f\in L_2[-\pi,\pi]$, with Fourier series*

$$\sum_{n=-\infty}^{\infty}\langle f,e_n\rangle e_n = \sum_{n=-\infty}^{\infty}c_n e^{int} \qquad \text{where} \quad c_n = \frac{\langle f,e_n\rangle}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt.$$

Then the series $\sum_{-\infty}^{\infty} \langle f, e_n \rangle e_n = \sum_{-\infty}^{\infty} c_n e^{int}$ *converges in* $L_2[-\pi, \pi]$ *to* f*, i.e*

$$\lim_{n\to\infty} \left\| f - \sum_{n=-\infty}^{\infty} c_n e^{int} \right\|_2 = 0.$$

Proof. This follows from the previous Theorem, Lemma 10.1 about density of CP in L_2, and Theorem 7.3 on orthonormal basis. □

12. Parseval's formula

The following result first appeared in the framework of $L_2[-\pi, \pi]$ and only later was understood to be a general property of inner product spaces.

Theorem 12.1 (Parseval's formula)**.** *If* f*,* $g \in L_2[-\pi, \pi]$ *have Fourier series* $f = \sum_{n=-\infty}^{\infty} c_n e^{int}$*,* $g = \sum_{n=-\infty}^{\infty} d_n e^{int}$ *then*

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)\overline{g(t)}\, dt = 2\pi \sum_{-\infty}^{\infty} c_n \overline{d_n}. \tag{12.1}$$

More generally if f *and* g *are two vectors of a Hilbert space* H *with an orthonormal basis* $(e_n)_{-\infty}^{\infty}$ *then*

$$\langle f, g \rangle = \sum_{k=-\infty}^{\infty} c_n \overline{d_n}, \qquad \textit{where} \quad c_n = \langle f, e_n \rangle\,,\ d_n = \langle g, e_n \rangle\,,$$

are the Fourier coefficients of f *and* g*.*

Proof. In fact we could just prove the second, more general, statement—the first one is its particular realisation. Let $f_n = \sum_{k=-n}^{n} c_k e_k$ and $g_n = \sum_{k=-n}^{n} d_k e_k$ will be partial sums of the corresponding Fourier series. Then from orthonormality of (e_n) and linearity of the inner product:

$$\langle f_n, g_n \rangle = \left\langle \sum_{k=-n}^{n} c_k e_k, \sum_{k=-n}^{n} d_k e_k \right\rangle = \sum_{k=-n}^{n} c_k \overline{d_k}.$$

This formula together with the facts that $f_k \to f$ and $g_k \to g$ (following from Corollary 11.5) and Lemma about continuity of the inner product implies the assertion. □

Corollary 12.2. *A integrable function* f *belongs to* $L_2[-\pi, \pi]$ *if and only if its Fourier series is convergent and then* $\|f\|^2 = 2\pi \sum_{-\infty}^{\infty} |c_k|^2$*.*

Proof. The necessity, i.e. implication $f \in L_2 \Rightarrow \langle f, f \rangle = \|f\|^2 = 2\pi \sum |c_k|^2$, follows from the previous Theorem. The sufficiency follows by Riesz–Fisher Theorem. □

Remark 12.3. The actual rôle of the Parseval's formula is shadowed by the orthonormality and is rarely recognised until we meet the *wavelets* or *coherent states*. Indeed the equality (12.1) should be read as follows:

Theorem 12.4 (Modified Parseval)**.** *The map* $\mathcal{W} : H \to \ell_2$ *given by the formula* $[\mathcal{W} f](n) = \langle f, e_n \rangle$ *is an isometry for any orthonormal basis* (e_n)*.*

We could find many other systems of vectors (e_x), $x \in X$ (very different from orthonormal bases) such that the map $\mathcal{W} : H \to L_2(X)$ given by the simple universal formula

$$[\mathcal{W} f](x) = \langle f, e_x \rangle \tag{12.2}$$

will be an isometry of Hilbert spaces. The map (12.2) is oftenly called *wavelet transform* and most famous is the *Cauchy integral formula* in complex analysis. The majority of wavelets transforms are linked with *group representations*, see our postgraduate course *Wavelets in Applied and Pure Maths*.

13. Some Application of Fourier Series

We are going to provide now few examples which demonstrate the importance of the Fourier series in many questions. The first two (Example 13.1 and Theorem 13.2) belong to pure mathematics and last two are of more applicable nature.

Example 13.1. Let $f(t) = t$ on $[-\pi, \pi]$. Then

$$\langle f, e_n \rangle = \int_{-\pi}^{\pi} t e^{-int}\, dt = \begin{cases} (-1)^n \frac{2\pi i}{n}, & n \neq 0 \\ 0, & n = 0 \end{cases} \qquad \text{(check!)},$$

so $f(t) \sim \sum_{-\infty}^{\infty} (-1)^n (i/n) e^{int}$. By a direct integration:

$$\|f\|_2^2 = \int_{-\pi}^{\pi} t^2\, dt = \frac{2\pi^3}{3}.$$

On the other hand by the previous Corollary:

$$\|f\|_2^2 = 2\pi \sum_{n\neq 0} \left| \frac{(-1)^n i}{n} \right|^2 = 4\pi \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Thus we get a beautiful formula

$$\sum_{1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Here is another important result.

Theorem 13.2 (Weierstrass Approximation Theorem)**.** *For any function $f \in C[a, b]$ and any $\epsilon > 0$ there exists a polynomial p such that $\|f - p\|_\infty < \epsilon$.*

Proof. Change variable: $t = 2\pi(x - \frac{a+b}{2})/(b - a)$ this maps $x \in [a, b]$ onto $t \in [-\pi, \pi]$. Let P denote the subspace of polynomials in $C[-\pi, \pi]$. Then $e^{int} \in \bar{P}$ for any $n \in \mathbb{Z}$ since Taylor series converges uniformly in $[-\pi, \pi]$. Consequently P contains the closed linear span in (supremum norm) of e^{int}, any $n \in \mathbb{Z}$, which is $CP[-\pi, \pi]$ by the Fejér theorem. Thus $\bar{P} \supseteq CP[-\pi, \pi]$ and we extend that to non-periodic function as follows (why we could not make use of Lemma 10.1 here, by the way?).

For any $f \in C[-\pi, \pi]$ let $\lambda = (f(\pi) - f(-\pi))/(2\pi)$ then $f_1(t) = f(t) - \lambda t \in CP[-\pi, \pi]$ and could be approximated by a polynomial $p_1(t)$ from the above discussion. Then $f(t)$ is approximated by the polynomial $p(t) = p_1(t) + \lambda t$. □

Example 13.3. The modern history of the Fourier analysis starts from the works of Fourier on the heat equation. As was mentioned in the introduction to this part, the exceptional role of Fourier coefficients for differential equations is explained by the simple formula $\partial_x e^{inx} = in e^{inx}$. We shortly review a solution of the *heat equation* to illustrate this.

Let we have a rod of the length 2π. The temperature at its point $x \in [-\pi, \pi]$ and a moment $t \in [0, \infty)$ is described by a function $u(t, x)$ on $[0, \infty) \times [-\pi, \pi]$. The mathematical equation describing a dynamics of the temperature distribution is:

$$\frac{\partial u(t,x)}{\partial t} = \frac{\partial^2 u(t,x)}{\partial x^2} \qquad \text{or, equivalently,} \qquad \left(\partial_t - \partial_x^2\right) u(t,x) = 0. \tag{13.1}$$

For any fixed moment t_0 the function $u(t_0, x)$ depends only from $x \in [-\pi, \pi]$ and according to Corollary 11.5 could be represented by its Fourier series:

$$u(t_0, x) = \sum_{n=-\infty}^{\infty} \langle u, e_n \rangle e_n = \sum_{n=-\infty}^{\infty} c_n(t_0) e^{inx} \qquad \text{where} \quad c_n(t_0) = \frac{\langle u, e_n \rangle}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} u(t_0, x) e^{-inx}\, dx,$$

with Fourier coefficients $c_n(t_0)$ depending from t_0. We substitute that decomposition into the heat equation (13.1) to receive:

$$\begin{aligned}\left(\partial_t - \partial_x^2\right) u(t,x) &= \left(\partial_t - \partial_x^2\right) \sum_{n=-\infty}^{\infty} c_n(t) e^{inx} \\ &= \sum_{n=-\infty}^{\infty} \left(\partial_t - \partial_x^2\right) c_n(t) e^{inx} \\ &= \sum_{n=-\infty}^{\infty} (c_n'(t) + n^2 c_n(t)) e^{inx} = 0. \qquad (13.2)\end{aligned}$$

Since function e^{inx} form a basis the last equation (13.2) holds if and only if

$$c_n'(t) + n^2 c_n(t) = 0 \qquad \text{for all } n \text{ and } t. \qquad (13.3)$$

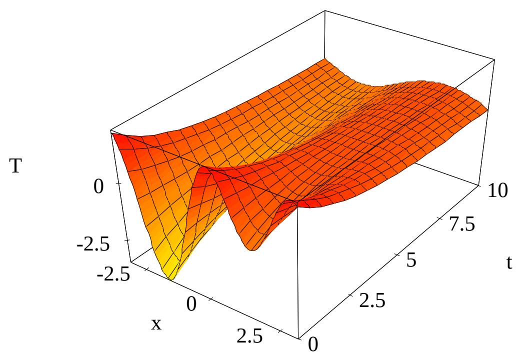

FIGURE 11. The dynamics of a heat equation:
x—coordinate on the rod,
t—time,
T—temperature.

Equations from the system (13.3) have general solutions of the form:

$$c_n(t) = c_n(0) e^{-n^2 t} \qquad \text{for all } t \in [0, \infty), \qquad (13.4)$$

producing a general solution of the heat equation (13.1) in the form:

$$u(t,x) = \sum_{n=-\infty}^{\infty} c_n(0) e^{-n^2 t} e^{inx} = \sum_{n=-\infty}^{\infty} c_n(0) e^{-n^2 t + inx}, \qquad (13.5)$$

where constant $c_n(0)$ could be defined from boundary condition. For example, if it is known that the initial distribution of temperature was $u(0,x) = g(x)$ for a function $g(x) \in L_2[-\pi, \pi]$ then $c_n(0)$ is the n-th Fourier coefficient of $g(x)$.

The general solution (13.5) helps produce both the analytical study of the heat equation (13.1) and numerical simulation. For example, from (13.5) obviously follows that

- the temperature is rapidly relaxing toward the thermal equilibrium with the temperature given by $c_0(0)$, however never reach it within a finite time;
- the "higher frequencies" (bigger thermal gradients) have a bigger speed of relaxation; etc.

The example of numerical simulation for the initial value problem with $g(x) = 2\cos(2*u) + 1.5\sin(u)$. It is clearly illustrate our above conclusions.

Example 13.4. Among the oldest periodic functions in human culture are acoustic waves of musical tones. The mathematical theory of musics (including rudiments of the Fourier analysis!) is as old as mathematics itself and was highly respected already in *Pythagoras' school* more 2500 years ago.

FIGURE 12. Two oscillation with unharmonious frequencies and the appearing dissonance. Click to listen the blue and green pure harmonics and red dissonance.

The earliest observations are that

(i) The musical sounds are made of pure harmonics (see the blue and green graphs on the Figure 12), in our language cos and sin functions form a basis;
(ii) Not every two pure harmonics are compatible, to be their frequencies should make a simple ratio. Otherwise the dissonance (red graph on Figure 12) appears.

FIGURE 13. Graphics of G5 performed on different musical instruments (click on picture to hear the sound). Samples are taken from Sound Library.

The musical tone, say G5, performed on different instruments clearly has something in common and different, see Figure 13 for comparisons. The decomposition into the pure harmonics, i.e. finding Fourier coefficient for the signal, could provide the complete characterisation, see Figure 14.

The Fourier analysis tells that:

(i) All sound have the same base (i.e. the lowest) frequencies which corresponds to the G5 tone, i.e. 788 Gz.
(ii) The higher frequencies, which are necessarily are multiples of 788 Gz to avoid dissonance, appears with different weights for different instruments.

The Fourier analysis is very useful in the signal processing and is indeed the fundamental tool. However it is not universal and has very serious limitations. Consider the simple case of the signals plotted on the Figure 15(a) and (b). They are both made out of same two pure harmonics:

Figure 14. Fourier series for G5 performed on different musical instruments (same order and colour as on the previous Figure)

(i) On the first signal the two harmonics (drawn in blue and green) follow one after another in time on Figure 15(a);
(ii) They just blended in equal proportions over the whole interval on Figure 15(b).

(a) (b) (c)

Figure 15. Limits of the Fourier analysis: different frequencies separated in time

This appear to be two very different signals. However the Fourier performed over the whole interval does not seems to be very different, see Figure 15(c). Both transforms (drawn in blue-green and pink) have two major pikes corresponding to the pure frequencies. It is not very easy to extract differences between signals from their Fourier transform (yet this should be possible according to our study).

Even a better picture could be obtained if we use *windowed Fourier transform*, namely use a sliding "window" of the constant width instead of the entire interval for the Fourier transform. Yet even better analysis could be obtained by means of *wavelets* already mentioned in Remark 12.3 in connection with Plancherel's formula. Roughly, wavelets correspond to a sliding window of a variable size—narrow for high frequencies and wide for low.

Part 4. Duality of Linear Spaces

Orthonormal basis allows to reduce any question on Hilbert space to a question on sequence of numbers. This is powerful but sometimes heavy technique. Sometime we need a smaller and faster tool to study questions which are represented by a single number, for example to demonstrate that two vectors are different it is enough to show that there is a unequal values of a single coordinate. In such cases *linear functionals* are just what we needed.

14. Dual space of a normed space

Definition 14.1. A *linear functional* on a vector space V is a linear mapping $\alpha : V \to \mathbb{C}$ (or $\alpha : V \to \mathbb{R}$ in the real case), i.e.

$$\alpha(ax + by) = a\alpha(x) + b\alpha(y), \qquad \text{for all } x, y \in V \text{ and } a, b \in \mathbb{C}.$$

Exercise 14.2. Show that $\alpha(0)$ is necessarily 0.

We will not consider any functionals but linear, thus bellow *functional* always means *linear functional.*

Example 14.3. (i) Let $V = \mathbb{C}^n$ and c_k, $k = 1, \ldots, n$ be complex numbers. Then $\alpha((x_1, \ldots, x_n)) = c_1 x_1 + \cdots + c_2 x_2$ is a linear functional.
(ii) On $C[0,1]$ a functional is given by $\alpha(f) = \int\limits_0^1 f(t)\,dt$.
(iii) On a Hilbert space H for any $x \in H$ a functional α_x is given by $\alpha_x(y) = \langle y, x \rangle$.

Theorem 14.4. *Let V be a normed space and α is a linear functional. The following are equivalent:*
(i) *α is continuous (at any point of V).*
(ii) *α is continuous at point 0.*
(iii) *$\sup\{|\alpha(x)| : \|x\| \le 1\} < \infty$, i.e. α is a* bounded linear functional.

Proof. Implication 14.4.i $\Rightarrow$ 14.4.ii is trivial.

Show 14.4.ii $\Rightarrow$ 14.4.iii. By the definition of continuity: for any $\epsilon > 0$ there exists $\delta > 0$ such that $\|v\| < \delta$ implies $|\alpha(v) - \alpha(0)| < \epsilon$. Take $\epsilon = 1$ then $|\alpha(\delta x)| < 1$ for all x with norm less than 1 because $\|\delta x\| < \delta$. But from linearity of α the inequality $|\alpha(\delta x)| < 1$ implies $|\alpha(x)| < 1/\delta < \infty$ for all $\|x\| \le 1$.

14.4.iii $\Rightarrow$ 14.4.i. Let mentioned supremum be M. For any x, $y \in V$ such that $x \neq y$ vector $(x-y)/\|x-y\|$ has norm 1. Thus $|\alpha((x-y)/\|x-y\|)| < M$. By the linearity of α this implies that $|\alpha(x) - \alpha(y)| < M\|x-y\|$. Thus α is continuous. □

Definition 14.5. The *dual space* X^* of a normed space X is the set of continuous linear functionals on X. Define a norm on it by

$$\|\alpha\| = \sup_{\|x\| \le 1} |\alpha(x)|. \tag{14.1}$$

Exercise 14.6. (i) Show that X^* is a linear space with natural operations.
(ii) Show that (14.1) defines a norm on X^*.
(iii) Show that $|\alpha(x)| \le \|\alpha\| \cdot \|x\|$ for all $x \in X$, $\alpha \in X^*$.

Theorem 14.7. *X^* is a Banach space with the defined norm (even if X was incomplete).*

Proof. Due to Exercise 14.6 we only need to show that X^* is complete. Let (α_n) be a Cauchy sequence in X^*, then for any $x \in X$ scalars $\alpha_n(x)$ form a Cauchy sequence, since $|\alpha_m(x) - \alpha_n(x)| \le \|\alpha_m - \alpha_n\| \cdot \|x\|$. Thus the sequence has a limit and we define α by $\alpha(x) = \lim_{n\to\infty} \alpha_n(x)$. Clearly α is a linear functional on X. We should show that it is bounded and $\alpha_n \to \alpha$. Given $\epsilon > 0$ there exists N such that $\|\alpha_n - \alpha_m\| < \epsilon$ for all n, $m \ge N$. If $\|x\| \le 1$ then $|\alpha_n(x) - \alpha_m(x)| \le \epsilon$, let $m \to \infty$ then $|\alpha_n(x) - \alpha(x)| \le \epsilon$, so

$$|\alpha(x)| \le |\alpha_n(x)| + \epsilon \le \|\alpha_n\| + \epsilon,$$

i.e. $\|\alpha\|$ is finite and $\|\alpha_n - \alpha\| \le \epsilon$, thus $\alpha_n \to \alpha$. □

Definition 14.8. The *kernel of linear functional* α, write $\ker\alpha$, is the set all vectors $x \in X$ such that $\alpha(x) = 0$.

Exercise 14.9. Show that
(i) $\ker\alpha$ is a subspace of X.
(ii) If $\alpha \not\equiv 0$ then $\ker\alpha$ is a *proper* subspace of X.
(iii) If α is continuous then $\ker\alpha$ is closed.

15. Self-duality of Hilbert space

Lemma 15.1 (Riesz–Fréchet). *Let H be a Hilbert space and α a continuous linear functional on H, then there exists the unique $y \in H$ such that $\alpha(x) = \langle x, y \rangle$ for all $x \in H$. Also $\|\alpha\|_{H^*} = \|y\|_H$.*

Proof. Uniqueness: if $\langle x, y\rangle = \langle x, y'\rangle \Leftrightarrow \langle x, y - y'\rangle = 0$ for all $x \in H$ then $y - y'$ is self-orthogonal and thus is zero (Exercise 5.2.i).

Existence: we may assume that $\alpha \not\equiv 0$ (otherwise take $y = 0$), then $M = \ker \alpha$ is a closed *proper* subspace of H. Since $H = M \oplus M^{\perp}$, there exists a non-zero $z \in M^{\perp}$, by scaling we could get $\alpha(z) = 1$. Then for any $x \in H$:

$$x = (x - \alpha(x)z) + \alpha(x)z, \qquad \text{with } x - \alpha(x)z \in M,\ \alpha(x)z \in M^{\perp}.$$

Because $\langle x, z\rangle = \alpha(x)\langle z, z\rangle = \alpha(x)\|z\|^2$ for any $x \in H$ we set $y = z/\|z\|^2$.

Equality of the norms $\|\alpha\|_{H^*} = \|y\|_H$ follows from the Cauchy–Bunyakovskii–Schwarz inequality in the form $\alpha(x) \leq \|x\| \cdot \|y\|$ and the identity $\alpha(y/\|y\|) = \|y\|$. □

Example 15.2. On $L_2[0,1]$ let $\alpha(f) = \langle f, t^2\rangle = \int_0^1 f(t)t^2\,dt$. Then

$$\|\alpha\| = \|t^2\| = \left(\int_0^1 (t^2)^2\,dt\right)^{1/2} = \frac{1}{\sqrt{5}}.$$

Part 5. Operators

All our previous considerations were only a preparation of the stage and now the main actors come forward to perform a play. The vectors spaces are not so interesting while we consider them in statics, what really make them exciting is the their transformations. The natural first steps is to consider transformations which respect both linear structure and the norm.

16. Linear operators

Definition 16.1. A *linear operator* T between two normed spaces X and Y is a mapping $T : X \to Y$ such that $T(\lambda v + \mu u) = \lambda T(v) + \mu T(u)$. The *kernel of linear operator* $\ker T$ and *image* are defined by

$$\ker T = \{x \in X : Tx = 0\} \qquad \operatorname{Im} T = \{y \in Y : y = Tx, \text{ for some } x \in X\}.$$

Exercise 16.2. Show that kernel of T is a linear subspace of X and image of T is a linear subspace of Y.

As usual we are interested also in connections with the second (topological) structure:

Definition 16.3. A *norm of linear operator* is defined:

$$\|T\| = \sup\{\|Tx\|_Y : \|x\|_X \leq 1\}. \tag{16.1}$$

Exercise 16.4. Show that $\|Tx\| \leq \|T\| \cdot \|x\|$ for all $x \in X$.

Example 16.5.
(i) On a normed space X define the *identity operator* by $\mathrm{I}_X : x \to x$ for all $x \in X$. Its norm is 1.
(ii) On a normed space X any linear functional define a linear operator from X to $\mathbb{C}$, its norm as operator is the same as functional.
(iii) The set of operators from $\mathbb{C}^n$ to $\mathbb{C}^m$ is given by $n \times m$ matrices which acts on vector by the matrix multiplication. All linear operators on finite-dimensional spaces are bounded.
(iv) On ℓ_2, let $S(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$ be the *right shift operator*. Clearly $\|Sx\| = \|x\|$ for all x, so $\|S\| = 1$.

(v) On $L_2[a,b]$, let $w(t) \in C[a,b]$ and define *multiplication operator* $M_w f$ by $(M_w f)(t) = w(t)f(t)$. Now:

$$\begin{aligned} \|M_w f\|^2 &= \int_a^b |w(t)|^2 |f(t)|^2 \, dt \\ &\leq K^2 \int_a^b |f(t)|^2 \, dt, \quad \text{where} \quad K = \|w\|_\infty = \sup_{[a,b]} |w(t)| , \end{aligned}$$

so $\|M_w\| \leq K$.

Exercise 16.6. Show that for multiplication operator in fact there is the equality of norms $\|M_w\|_2 = \|w(t)\|_\infty$.

Theorem 16.7. *Let $T : X \to Y$ be a linear operator. The following conditions are equivalent:*

(i) *T is continuous on X;*
(ii) *T is continuous at the point 0.*
(iii) *T is a* bounded linear operator, *i.e* $\|T\| = \sup\{\|Tx\| : \|x\|\} < \infty$.

Proof. Proof essentially follows the proof of similar Theorem 14.4. □

17. $B(H)$ as a Banach space (and even algebra)

Theorem 17.1. *Let $B(X,Y)$ be the* space of bounded linear operators *from X and Y with the norm defined above. If Y is complete, then $B(X,Y)$ is a Banach space.*

Proof. The proof repeat proof of the Theorem 14.7, which is a particular case of the present theorem for $Y = \mathbb{C}$, see Example 16.5.ii. □

Theorem 17.2. *Let $T \in B(X,Y)$ and $S \in B(Y,Z)$, where X, Y, and Z are normed spaces. Then $ST \in B(X,Z)$ and $\|ST\| \leq \|S\| \, \|T\|$.*

Proof. Clearly $(ST)x = S(Tx) \in Z$, and

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

which implies norm estimation if $\|x\| \leq 1$. □

Corollary 17.3. *Let $T \in B(X,X)$ =$B(X)$, where X is a normed space. Then for any $n \geq 1$, $T^n \in B(X)$ and $\|T^n\| \leq \|T\|^n$.*

Proof. It is induction by n with the trivial base $n = 1$ and the step following from the previous theorem. □

Remark 17.4. Some texts use notations $L(X,Y)$ and $L(X)$ instead of ours $B(X,Y)$ and $B(X)$.

Definition 17.5. Let $T \in B(X,Y)$. We say T is an *invertible operator* if there exists $S \in B(Y,X)$ such that

$$ST = \mathrm{I}_X \qquad \text{and} \qquad TS = \mathrm{I}_Y .$$

Such an S is called the *inverse operator* of T.

Exercise 17.6. Show that the inverse operator is unique (if exists at all). (Assume existence of S and S', then consider operator STS'.)

Example 17.7. We consider inverses to operators from Exercise 16.5.

(i) The identity operator I_X is the inverse of itself.
(ii) A linear functional is not invertible unless it is non-zero and X is one dimensional.
(iii) An operator $\mathbb{C}^n \to \mathbb{C}^m$ is invertible if and only if $m = n$ and corresponding square matrix is non-singular, i.e. has non-zero determinant.

(iv) The right shift S is not invertible on ℓ_2 (it is one-to-one but is not onto). But the *left shift operator* $T(x_1, x_2, \ldots) = (x_2, x_3, \ldots)$ is its *left inverse*, i.e. $TS = I$ but $TS \neq I$ since $ST(1, 0, 0, \ldots) = (0, 0, \ldots)$. T is not invertible either (it is onto but not one-to-one), however S is its *right inverse*.

(v) Operator of multiplication M_w is invertible if and only if $w^{-1} \in C[a, b]$ and inverse is $M_{w^{-1}}$. For example M_{1+t} is invertible $L_2[0, 1]$ and M_t is not.

18. Adjoints

Theorem 18.1. *Let H and K be Hilbert Spaces and $T \in B(H, K)$. Then there exists operator $T^* \in B(K, H)$ such that*

$$\langle Th, k\rangle_K = \langle h, T^*k\rangle_H \qquad \text{for all} \quad h \in H, \ k \in K.$$

Such T^* is called the adjoint operator *of T. Also $T^{**} = T$ and $\|T^*\| = \|T\|$.*

Proof. For any fixed $k \in K$ the expression $h :\rightarrow \langle Th, k\rangle_K$ defines a bounded linear functional on H. By the Riesz–Fréchet lemma there is a *unique* $y \in H$ such that $\langle Th, k\rangle_K = \langle h, y\rangle_H$ for all $h \in H$. Define $T^*k = y$ then T^* is linear:

$$\begin{aligned}
\langle h, T^*(\lambda_1 k_1 + \lambda_2 k_2)\rangle_H &= \langle Th, \lambda_1 k_1 + \lambda_2 k_2\rangle_K \\
&= \bar\lambda_1 \langle Th, k_1\rangle_K + \bar\lambda_2 \langle Th, k_2\rangle_K \\
&= \bar\lambda_1 \langle h, T^*k_1\rangle_H + \bar\lambda_2 \langle h, T^*k_2\rangle_K \\
&= \langle h, \lambda_1 T^*k_1 + \lambda_2 T^*k_2\rangle_H
\end{aligned}$$

So $T^*(\lambda_1 k_1 + \lambda_2 k_2) = \lambda_1 T^*k_1 + \lambda_2 T^*k_2$. T^{**} is defined by $\langle k, T^{**}h\rangle = \langle T^*k, h\rangle$ and the identity $\langle T^{**}h, k\rangle = \langle h, T^*k\rangle = \langle Th, k\rangle$ for all h and k shows $T^{**} = T$. Also:

$$\begin{aligned}
\|T^*k\|^2 &= \langle T^*k, T^*k\rangle = \langle k, TT^*k\rangle \\
&\leq \|k\| \cdot \|TT^*k\| \leq \|k\| \cdot \|T\| \cdot \|T^*k\|,
\end{aligned}$$

which implies $\|T^*k\| \leq \|T\| \cdot \|k\|$, consequently $\|T^*\| \leq \|T\|$. The opposite inequality follows from the identity $\|T\| = \|T^{**}\|$. □

Exercise 18.2. (i) For operators T_1 and T_2 show that

$$(T_1T_2)^* = T_2^*T_1^*, \qquad (T_1 + T_2)^* = T_1^* + T_2^* \qquad (\lambda T)^* = \bar\lambda T^*.$$

(ii) If A is an operator on a Hilbert space H then $(\ker A)^\perp = \operatorname{Im} A^*$.

19. Hermitian, unitary and normal operators

Definition 19.1. An operator $T : H \rightarrow H$ is a *Hermitian operator* or *self-adjoint operator* if $T = T^*$, i.e. $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$.

Example 19.2. (i) On ℓ_2 the adjoint S^* to the right shift operator S is given by the left shift $S^* = T$, indeed:

$$\begin{aligned}
\langle Sx, y\rangle &= \langle (0, x_1, x_2, \ldots), (y_1, y_2, \ldots)\rangle \\
&= x_1\bar y_2 + x_2 \bar y_3 + \cdots = \langle (x_1, x_2, \ldots), (y_2, y_3, \ldots)\rangle \\
&= \langle x, Ty\rangle.
\end{aligned}$$

Thus S is *not* Hermitian.

(ii) Let D be *diagonal operator* given by

$$D(x_1, x_2, \ldots) = (\lambda_1 x_1, \lambda_2 x_2, \ldots).$$

where (λ_k) is any bounded complex sequence. It is easy to check that $\|D\| = \|(\lambda_n)\|_\infty = \sup_k |\lambda_k|$ and

$$D^*(x_1, x_2, \ldots) = (\bar\lambda_1 x_1, \bar\lambda_2 x_2, \ldots),$$

thus D is Hermitian if and only if $\lambda_k \in \mathbb{R}$ for all k.

(iii) If $T : \mathbb{C}^n \to \mathbb{C}^n$ is represented by multiplication of a column vector by a matrix A, then T^* is multiplication by the matrix A^*—transpose and conjugate to A.

Exercise 19.3. Show that for any bounded operator T operators $T_1 = \frac{1}{2}(T+T^*)$ and $T_1 = \frac{1}{2i}(T-T^*)$ are Hermitians.

Definition 19.4. We say that $U : H \to H$ is a *unitary operator* on a Hilbert space H if $U^* = U^{-1}$, i.e. $U^*U = UU^* = I$.

Example 19.5. (i) If $D : \ell_2 \to \ell_2$ is a diagonal operator such that $De_k = \lambda_k e_k$, then $D^*e_k = \bar{\lambda}_k e_k$ and D is unitary if and only if $|\lambda_k| = 1$ for all k.
(ii) The shift operator S satisfies $S^*S = I$ but $SS^* \neq I$ thus S is **not** unitary.

Theorem 19.6. *For an operator $U \in B(H)$ the following are equivalent:*
(i) *U is unitary;*
(ii) *U is surjection and an* isometry*, i.e. $\|Ux\| = \|x\|$ for all $x \in H$;*
(iii) *U is a surjection and preserves the inner product, i.e. $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all x, $y \in H$.*

Proof. 19.6.i⇒19.6.ii. Clearly unitarity of operator implies its invertibility and hence surjectivity. Also

$$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = \langle x, x\rangle = \|x\|^2 .$$

19.6.ii⇒19.6.iii. Using the polarisation identity (cf. polarisation in equation (2.5)):

$$\begin{aligned} 4\langle Tx, y\rangle &= \langle T(x+y), x+y\rangle + i\langle T(x+iy), x+iy\rangle \\ &\quad - \langle T(x-y), x-y\rangle - i\langle T(x-iy), x-iy\rangle . \\ &= \sum_{k=0}^{3} i^k \left\langle T(x+i^k y), x+i^k y\right\rangle \end{aligned}$$

Take $T = U^*U$ and $T = I$, then

$$\begin{aligned} 4\langle U^*Ux, y\rangle &= \sum_{k=0}^{3} i^k \left\langle U^*U(x+i^k y), x+i^k y\right\rangle \\ &= \sum_{k=0}^{3} i^k \left\langle U(x+i^k y), U(x+i^k y)\right\rangle \\ &= \sum_{k=0}^{3} i^k \left\langle (x+i^k y), (x+i^k y)\right\rangle \\ &= 4\langle x, y\rangle . \end{aligned}$$

19.6.iii⇒19.6.i. Indeed $\langle U^*Ux, y\rangle = \langle x, y\rangle$ implies $\langle (U^*U - I)x, y\rangle = 0$ for all x,$y \in H$, then $U^*U = I$. Since U should be invertible by surjectivity we see that $U^* = U^{-1}$. □

Definition 19.7. A *normal operator* T is one for which $T^*T = TT^*$.

Example 19.8. (i) Any self-adjoint operator T is normal, since $T^* = T$.
(ii) Any unitary operator U is normal, since $U^*U = I = UU^*$.
(iii) Any diagonal operator D is normal , since $De_k = \lambda_k e_k$, $D^*e_k = \bar{\lambda}_k e_k$, and $DD^*e_k = D^*De_k = |\lambda_k|^2 e_k$.
(iv) The shift operator S is **not** normal.
(v) A finite matrix is normal (as an operator on ℓ_2^n) if and only if it has an orthonormal basis in which it is diagonal.

Part 6. Spectral Theory

As we saw operators could be added and multiplied each other, in some sense they behave like numbers, but are much more complicated. In this lecture we will associate to each operator a set of complex numbers which reflects certain (unfortunately not all) properties of this operator.

The analogy between operators and numbers become even more deeper since we could construct *functions of operators* (called *functional calculus*) in a way we build numeric functions. The most important functions of this sort is called *resolvent* (see Definition 20.5). The methods of analytical functions are very powerful in operator theory and students may wish to refresh their knowledge of complex analysis before this part.

20. The spectrum of an operator on a Hilbert space

An *eigenvalue of operator* $T \in B(H)$ is a complex number λ such that there exists a nonzero $x \in H$, called *eigenvector* with property $Tx = \lambda x$, in other words $x \in \ker(T - \lambda I)$.

In finite dimensions $T - \lambda I$ is invertible if and only if λ is **not** an eigenvalue. In infinite dimensions it is not the same: the right shift operator S is not invertible but 0 is not its eigenvalue because $Sx = 0$ implies $x = 0$ (check!).

Definition 20.1. The *resolvent set* $\rho(T)$ of an operator T is the set

$$\rho(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is invertible}\}.$$

The *spectrum of operator* $T \in B(H)$, denoted $\sigma(T)$, is the complement of the resolvent set $\rho(T)$:

$$\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is } \textbf{not} \text{ invertible}\}.$$

Example 20.2. If H is finite dimensional the from previous discussion follows that $\sigma(T)$ is the set of eigenvalues of T for any T.

Even this example demonstrates that spectrum does not provide a complete description for operator even in finite-dimensional case. For example, both operators in $\mathbb{C}^2$ given by matrices $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ have a single point spectrum $\{0\}$, however are rather different. The situation became even worst in the infinite dimensional spaces.

Theorem 20.3. *The spectrum $\sigma(T)$ of a bounded operator T is a nonempty compact (i.e. closed and bounded) subset of $\mathbb{C}$.*

For the proof we will need several Lemmas.

Lemma 20.4. *Let $A \in B(H)$. If $\|A\| < 1$ then $I - A$ is invertible in $B(H)$ and inverse is given by the* von Neumann series*:*

$$(I - A)^{-1} = I + A + A^2 + A^3 + \ldots = \sum_{k=0}^{\infty} A^k. \qquad (20.1)$$

Proof. Define the sequence of operators $B_n = I + A + \cdots + A^N$—the partial sums of the infinite series (20.1). It is a Cauchy sequence, indeed:

$$\begin{aligned} \|B_n - B_m\| &= \left\|A^{m+1} + A^{m+2} + \cdots + A^n\right\| \qquad \text{(if } n < m) \\ &\leq \left\|A^{m+1}\right\| + \left\|A^{m+2}\right\| + \cdots + \|A^n\| \\ &\leq \|A\|^{m+1} + \|A\|^{m+2} + \cdots + \|A\|^n \\ &\leq \frac{\|A\|^{m+1}}{1 - \|A\|} < \epsilon \end{aligned}$$

for a large m. By the completeness of $B(H)$ there is a limit, say B, of the sequence B_n. It is a simple algebra to check that $(I-A)B_n = B_n(I-A) = I - A^{n+1}$, passing to the limit in the norm topology, where $A^{n+1} \to 0$ and $B_n \to B$ we get:

$$(I-A)B = B(I-A) = I \quad \Leftrightarrow \quad B = (I-A)^{-1}.$$

□

Definition 20.5. The *resolvent* of an operator T is the operator valued function defined on the resolvent set by the formula:

$$R(\lambda, T) = (T - \lambda I)^{-1}. \tag{20.2}$$

Corollary 20.6. (i) *If $|\lambda| > \|T\|$ then $\lambda \in \rho(T)$, hence the spectrum is bounded.*

(ii) *The resolvent set $\rho(T)$ is open, i.e for any $\lambda \in \rho(T)$ then there exist $\epsilon > 0$ such that all μ with $|\lambda - \mu| < \epsilon$ are also in $\rho(T)$, i.e. the resolvent set is open and the spectrum is closed.*

Both statements together imply that the spectrum is compact.

Proof. (i) If $|\lambda| > \|T\|$ then $\|\lambda^{-1}T\| < 1$ and the operator $T - \lambda I = -\lambda(I - \lambda^{-1}T)$ has the inverse

$$R(\lambda, T) = (T - \lambda I)^{-1} = -\sum_{k=0}^{\infty} \lambda^{-k-1} A^k. \tag{20.3}$$

by the previous Lemma.

(ii) Indeed:

$$\begin{aligned} T - \mu I &= T - \lambda I + (\lambda - \mu) I \\ &= (T - \lambda I)(I + (\lambda - \mu)(T - \lambda I)^{-1}). \end{aligned}$$

The last line is an invertible operator because $T - \lambda I$ is invertible by the assumption and $I + (\lambda - \mu)(T - \lambda I)^{-1}$ is invertible by the previous Lemma, since $\|(\lambda - \mu)(T - \lambda I)^{-1}\| < 1$ if $\epsilon < \|(T - \lambda I)^{-1}\|$.

□

Exercise 20.7. (i) Prove the *first resolvent identity*:

$$R(\lambda, T) - R(\mu, T) = (\lambda - \mu) R(\lambda, T) R(\mu, T) \tag{20.4}$$

(ii) Use the identity (20.4) to show that $(T - \mu I)^{-1} \to (T - \lambda I)^{-1}$ as $\mu \to \lambda$.

(iii) Use the identity (20.4) to show that for $z \in \rho(t)$ the complex derivative $\frac{d}{dz} R(z, T)$ of the resolvent $R(z, T)$ is well defined, i.e. the resolvent is an analytic function operator valued function of z.

Lemma 20.8. *The spectrum is non-empty.*

Proof. Let us assume the opposite, $\sigma(T) = \emptyset$ then the resolvent function $R(\lambda, T)$ is well defined for all $\lambda \in \mathbb{C}$. As could be seen from the von Neumann series (20.3) $\|R(\lambda, T)\| \to 0$ as $\lambda \to \infty$. Thus for any vectors x, $y \in H$ the function $f(\lambda) = \langle R(\lambda, T)x, y)\rangle$ is analytic (see Exercise 20.7.iii) function tensing to zero at infinity. Then by the Liouville theorem from complex analysis $R(\lambda, T) = 0$, which is impossible. Thus the spectrum is not empty. □

Proof of Theorem 20.3. Spectrum is nonempty by Lemma 20.8 and compact by Corollary 20.6. □

Remark 20.9. Theorem 20.3 gives the maximal possible description of the spectrum, indeed any non-empty compact set could be a spectrum for some bounded operator, see Problem E.3.

21. The spectral radius formula

The following definition is of interest.

Definition 21.1. The *spectral radius* of T is

$$r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

From the Lemma 20.6.i immediately follows that $r(T) \leq \|T\|$. The more accurate estimation is given by the following theorem.

Theorem 21.2. *For a bounded operator T we have*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}. \tag{21.1}$$

We start from the following general lemma:

Lemma 21.3. *Let a sequence (a_n) of positive real numbers satisfies inequalities: $0 \leq a_{m+n} \leq a_m + a_n$ for all m and n. Then there is a limit $\lim\limits_{n\to\infty}(a_n/n)$ and its equal to $\inf\limits_{n}(a_n/n)$.*

Proof. The statements follows from the observation that for any n and $m = nk + l$ with $0 \leq l \leq n$ we have $a_m \leq k a_n + l a_1$ thus, for big m we got $a_m/m \leq a_n/n + la_1/m \leq a_n/n + \epsilon$. □

Proof of Theorem 21.2. The existence of the limit $\lim_{n\to\infty}\|T^n\|^{1/n}$ in (21.1) follows from the previous Lemma since by the Lemma 17.2 $\log\|T^{n+m}\| \leq \log\|T^n\| + \log\|T^m\|$. Now we are using some results from the complex analysis. The Laurent series for the resolvent $R(\lambda, T)$ in the neighbourhood of infinity is given by the von Neumann series (20.3). The radius of its convergence (which is equal, obviously, to $r(T)$) by the Hadamard theorem is exactly $\lim_{n\to\infty}\|T^n\|^{1/n}$. □

Corollary 21.4. *There exists $\lambda \in \sigma(T)$ such that $|\lambda| = r(T)$.*

Proof. Indeed, as its known from the complex analysis the boundary of the convergence circle of a Laurent (or Taylor) series contain a singular point, the singular point of the resolvent is obviously belongs to the spectrum. □

Example 21.5. Let us consider the left shift operator S^*, for any $\lambda \in \mathbb{C}$ such that $|\lambda| < 1$ the vector $(1, \lambda, \lambda^2, \lambda^3, \ldots)$ is in ℓ_2 and is an eigenvector of S^* with eigenvalue λ, so the open unit disk $|\lambda| < 1$ belongs to $\sigma(S^*)$. On the other hand spectrum of S^* belongs to the closed unit disk $|\lambda| \leq 1$ since $r(S^*) \leq \|S^*\| = 1$. Because spectrum is closed it should coincide with the closed unit disk, since the open unit disk is dense in it. Particularly $1 \in \sigma(S^*)$, but it is easy to see that 1 is not an eigenvalue of S^*.

Proposition 21.6. *For any $T \in B(H)$ the spectrum of the adjoint operator is $\sigma(T^*) = \{\bar\lambda : \lambda \in \sigma(T)\}$.*

Proof. If $(T - \lambda I)V = V(T - \lambda I) = I$ the by taking adjoints $V^*(T^* - \bar\lambda I) = (T^* - \bar\lambda I)V^* = I$. So $\lambda \in \rho(T)$ implies $\bar\lambda \in \rho(T^*)$, using the property $T^{**} = T$ we could invert the implication and get the statement of proposition. □

Example 21.7. In continuation of Example 21.5 using the previous Proposition we conclude that $\sigma(S)$ is also the closed unit disk, but S does not have eigenvalues at all!

22. Spectrum of Special Operators

Theorem 22.1. (i) *If U is a unitary operator then $\sigma(U) \subseteq \{|z| = 1\}$.*
(ii) *If T is Hermitian then $\sigma(T) \subseteq \mathbb{R}$.*

Proof. (i) If $|\lambda| > 1$ then $\|\lambda^{-1}U\| < 1$ and then $\lambda I - U = \lambda(I - \lambda^{-1}U)$ is invertible, thus $\lambda \notin \sigma(U)$. If $|\lambda| < 1$ then $\|\lambda U^*\| < 1$ and then $\lambda I - U = U(\lambda U^* - I)$ is invertible, thus $\lambda \notin \sigma(U)$. The remaining set is exactly $\{z : |z| = 1\}$.

(ii) Without lost of generality we could assume that $\|T\| < 1$, otherwise we could multiply T by a small real scalar. Let us consider the *Cayley transform* which maps real axis to the unit circle:

$$U = (T - iI)(T + iI)^{-1}.$$

Straightforward calculations show that U is unitary if T is Hermitian. Let us take $\lambda \notin \mathbb{R}$ and $\lambda \neq -i$ (this case could be checked directly by Lemma 20.4). Then the Cayley transform $\mu = (\lambda - i)(\lambda + i)^{-1}$ of λ is not on the unit circle and thus the operator

$$U - \mu I = (T - iI)(T + iI)^{-1} - (\lambda - i)(\lambda + i)^{-1} I = 2i(\lambda + i)^{-1}(T - \lambda I)(T + iI)^{-1},$$

is invertible, which implies invertibility of $T - \lambda I$. So $\lambda \notin \mathbb{R}$.

□

Part 7. Compactness

It is not easy to study linear operators "in general" and there are many questions about operators in Hilbert spaces raised many decades ago which are still unanswered. Therefore it is reasonable to single out classes of operators which have (relatively) simple properties. Such a class of operators more closed to finite dimensional ones will be studied here.

23. Compact operators

Let us recall some topological definition and results.

Definition 23.1. A *compact set* in a metric space is defined by the property that any its covering by a family of open sets contains a subcovering by a finite subfamily.

In the finite dimensional vector spaces $\mathbb{R}^n$ or $\mathbb{C}^n$ there is the following equivalent definition of compactness (equivalence of 23.2.i and 23.2.ii is known as *Heine–Borel theorem*):

Theorem 23.2. *If a set E in $\mathbb{R}^n$ or $\mathbb{C}^n$ has any of the following properties then it has other two as well:*

(i) *E is bounded and closed;*
(ii) *E is compact;*
(iii) *Any infinite subset of E has a limiting point belonging to E.*

Exercise* 23.3. Which equivalences from above are not true any more in the infinite dimensional spaces?

Definition 23.4. Let X and Y be normed spaces, $T \in B(X, Y)$ is a *finite rank operator* if $\operatorname{Im} T$ is a finite dimensional subspace of Y. T is a *compact operator* if whenever $(x_i)_1^\infty$ is a bounded sequence in X then its image $(Tx_i)_1^\infty$ has a convergent subsequence in Y.

The set of finite rank operators is denote by $F(X, Y)$ and the set of compact operators—by $K(X, Y)$

Exercise 23.5. Show that both $F(X, Y)$ and $K(X, Y)$ are linear subspaces of $B(X, Y)$.

We intend to show that $F(X, Y) \subset K(X, Y)$.

Lemma 23.6. *Let Z be a finite-dimensional normed space. Then there is a number N and a mapping $S : \ell_2^N \to Z$ which is invertible and such that S and S^{-1} are bounded.*

Proof. The proof is given by an explicit construction. Let $N = \dim Z$ and $z_1, z_2, \ldots, z_N$ be a basis in Z. Let us define

$$S : \ell_2^N \to Z \quad \text{by} \quad S(a_1, a_2, \ldots, a_N) = \sum_{k=1}^{N} a_k z_k,$$

then we have an estimation of norm:

$$\begin{aligned}\|Sa\| &= \left\|\sum_{k=1}^N a_k z_k\right\| \leq \sum_{k=1}^N |a_k| \, \|z_k\| \\ &\leq \left(\sum_{k=1}^N |a_k|^2\right)^{1/2} \left(\sum_{k=1}^N \|z_k\|^2\right)^{1/2}.\end{aligned}$$

So $\|S\| \leq \left(\sum_1^N \|z_k\|^2\right)^{1/2}$ and S is continuous.

Clearly S has the trivial kernel, particularly $\|Sa\| > 0$ if $\|a\| = 1$. By the Heine–Borel theorem the unit sphere in ℓ_2^N is compact, consequently the continuous function $a \mapsto \left\|\sum_1^N a_k z_k\right\|$ attains its lower bound, which has to be positive. This means there exists $\delta > 0$ such that $\|a\| = 1$ implies $\|Sa\| > \delta$, or, equivalently if $\|z\| < \delta$ then $\|S^{-1}z\| < 1$. The later means that $\|S^{-1}\| \leq \delta^{-1}$ and boundedness of S^{-1}. □

Corollary 23.7. *For any two metric spaces X and Y we have $F(X,Y) \subset K(X,Y)$.*

Proof. Let $T \in F(X,Y)$, if $(x_n)_1^\infty$ is a bounded sequence in X then $((Tx_n)_1^\infty \subset Z = \operatorname{Im} T$ is also bounded. Let $S : \ell_2^N \to Z$ be a map constructed in the above Lemma. The sequence $(S^{-1}Tx_n)_1^\infty$ is bounded in ℓ_2^N and thus has a limiting point, say a_0. Then Sa_0 is a limiting point of $(Tx_n)_1^\infty$. □

There is a simple condition which allows to determine which diagonal operators are compact (particularly the identity operator $I_X$is *not* compact if $\dim X = \infty$):

Proposition 23.8. *Let T is a diagonal operator and given by identities $Te_n = \lambda_n e_n$ for all n in a basis e_n. T is compact if and only if $\lambda_n \to 0$.*

Proof. If $\lambda_n \not\to 0$ then there exists a subsequence λ_{n_k} and $\delta > 0$ such that $|\lambda_{n_k}| > \delta$ for all k. Now the sequence (e_{n_k}) is bounded but its image $Te_{n_k} = \lambda_{n_k} e_{n_k}$ has no convergent subsequence because for any $k \neq l$:

$$\|\lambda_{n_k} e_{n_k} - \lambda_{n_l} e_{n_l}\| = (|\lambda_{n_k}|^2 + |\lambda_{n_l}|^2)^{1/2} \geq \sqrt{2}\delta,$$

i.e. Te_{n_k} is not a Cauchy sequence, see Figure 16.

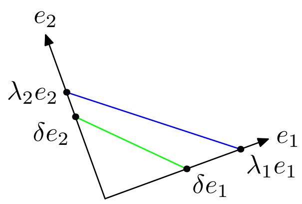

FIGURE 16. Distance between scales of orthonormal vectors

For the converse, note that if $\lambda_n \to 0$ then we can define a finite rank operator T_m, $m \geq 1$—m-"truncation" of T by:

$$T_m e_n = \begin{cases} Te_n = \lambda_n e_n, & 1 \leq n \leq m; \\ 0, & n > m. \end{cases} \tag{23.1}$$

Then obviously

$$(T - T_m)e_n = \begin{cases} 0, & 1 \leq n \leq m; \\ \lambda_n e_n, & n > m, \end{cases}$$

and $\|T - T_m\| = \sup_{n>m} |\lambda_n| \to 0$ if $m \to \infty$. All T_m are finite rank operators (so are compact) and T is also compact as their limit—by the next Theorem. □

Theorem 23.9. *Let T_m be a sequence of compact operators convergent to an operator T in the norm topology (i.e. $\|T - T_m\| \to 0$) then T is compact itself. Equivalently $K(X,Y)$ is a closed subspace of $B(X,Y)$.*

Proof. Take a bounded sequence $(x_n)_1^\infty$. From compactness

$$\begin{array}{llllll}
\text{of } T_1 & \Rightarrow \exists & \text{subsequence } (x_n^{(1)})_1^\infty \text{ of } (x_n)_1^\infty & \text{s.t.} & (T_1 x_n^{(1)})_1^\infty \text{ is convergent.} \\
\text{of } T_2 & \Rightarrow \exists & \text{subsequence } (x_n^{(2)})_1^\infty \text{ of } (x_n^{(1)})_1^\infty & \text{s.t.} & (T_2 x_n^{(2)})_1^\infty \text{ is convergent.} \\
\text{of } T_3 & \Rightarrow \exists & \text{subsequence } (x_n^{(3)})_1^\infty \text{ of } (x_n^{(2)})_1^\infty & \text{s.t.} & (T_3 x_n^{(3)})_1^\infty \text{ is convergent.} \\
\ldots & \ldots & \ldots & \ldots & \ldots
\end{array}$$

Could we find a subsequence which converges for all T_m simultaneously? The first guess "take the intersection of all above sequences $(x_n^{(k)})_1^\infty$" does not work because the intersection could be empty. The way out is provided by the *diagonal argument* (see Table 2): a subsequence $(T_m x_k^{(k)})_1^\infty$ is convergent for all m, because at latest after the term $x_m^{(m)}$ it is a subsequence of $(x_k^{(m)})_1^\infty$.

$$\begin{array}{cccccccc}
\mathbf{T_1 x_1^{(1)}} & T_1 x_2^{(1)} & T_1 x_3^{(1)} & \ldots & T_1 x_n^{(1)} & \ldots & \to & a_1 \\
T_2 x_1^{(2)} & \mathbf{T_2 x_2^{(2)}} & T_2 x_3^{(2)} & \ldots & T_2 x_n^{(2)} & \ldots & \to & a_2 \\
T_3 x_1^{(3)} & T_3 x_2^{(3)} & \mathbf{T_3 x_3^{(3)}} & \ldots & T_3 x_n^{(3)} & \ldots & \to & a_3 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & & \\
T_n x_1^{(n)} & T_n x_2^{(n)} & T_n x_3^{(n)} & \ldots & \mathbf{T_n x_n^{(n)}} & \ldots & \to & a_n \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & & \downarrow \\
 & & & & & & \searrow & \\
 & & & & & & & \mathbf{a}
\end{array}$$

TABLE 2. The "diagonal argument".

We are claiming that a subsequence $(T x_k^{(k)})_1^\infty$ of $(T x_n)_1^\infty$ is convergent as well. We use here $\epsilon/3$ *argument* (see Figure 17): for a given $\epsilon > 0$ choose $p \in \mathbb{N}$ such that $\|T - T_p\| < \epsilon/3$. Because

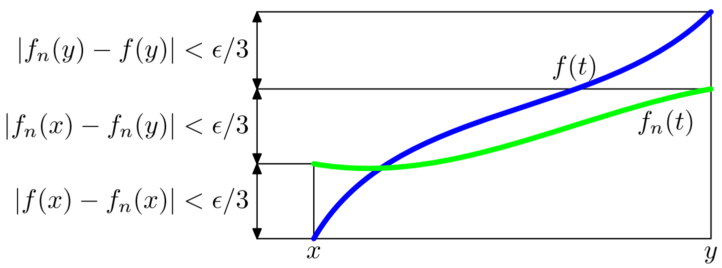

FIGURE 17. The $\epsilon/3$ argument to estimate $|f(x) - f(y)|$.

$(T_p x_k^{(k)}) \to 0$ it is a Cauchy sequence, thus there exists $n_0 > p$ such that $\left\| T_p x_k^{(k)} - T_p x_l^{(l)} \right\| < \epsilon/3$ for all k, $l > n_0$. Then:

$$\begin{aligned}
\left\| T x_k^{(k)} - T x_l^{(l)} \right\| &= \left\| (T x_k^{(k)} - T_p x_k^{(k)}) + (T_p x_k^{(k)} - T_p x_l^{(l)}) + (T_p x_l^{(l)} - T x_l^{(l)}) \right\| \\
&\leq \left\| T x_k^{(k)} - T_p x_k^{(k)} \right\| + \left\| T_p x_k^{(k)} - T_p x_l^{(l)} \right\| + \left\| T_p x_l^{(l)} - T x_l^{(l)} \right\| \\
&\leq \epsilon
\end{aligned}$$

Thus T is compact. □

24. HILBERT–SCHMIDT OPERATORS

Definition 24.1. Let $T : H \to K$ be a bounded linear map between two Hilbert spaces. Then T is said to be *Hilbert–Schmidt operator* if there exists an orthonormal basis in H such that the series $\sum_{k=1}^\infty \|T e_k\|^2$ is convergent.

Example 24.2. (i) Let $T : \ell_2 \to \ell_2$ be a diagonal operator defined by $Te_n = e_n/n$, for all $n \geq 1$. Then $\sum \|Te_n\|^2 = \sum n^{-2} = \pi^2/6$ (see Example 13.1) is finite.
(ii) The identity operator I_H is **not** a Hilbert–Schmidt operator, unless H is finite dimensional.

A relation to compact operator is as follows.

Theorem 24.3. *All Hilbert–Schmidt operators are compact. (The opposite inclusion is false, give a counterexample!)*

Proof. Let $T \in B(H, K)$ have a convergent series $\sum \|Te_n\|^2$ in an orthonormal basis $(e_n)_1^\infty$ of H. We again (see (23.1)) define the m-truncation of T by the formula

$$T_m e_n = \begin{cases} Te_n, & 1 \leq n \leq m; \\ 0, & n > m. \end{cases} \tag{24.1}$$

Then $T_m(\sum_1^\infty a_k e_k) = \sum_1^m a_k e_k$ and each T_m is a finite rank operator because its image is spanned by the finite set of vectors $Te_1, \ldots, Te_n$. We claim that $\|T - T_m\| \to 0$. Indeed by linearity and definition of T_m:

$$(T - T_m)\left(\sum_{n=1}^\infty a_n e_n\right) = \sum_{n=m+1}^\infty a_n (Te_n).$$

Thus:

$$\begin{aligned} \left\|(T - T_m)\left(\sum_{n=1}^\infty a_n e_n\right)\right\| &= \left\|\sum_{n=m+1}^\infty a_n (Te_n)\right\| \qquad (24.2)\\ &\leq \sum_{n=m+1}^\infty |a_n| \, \|(Te_n)\| \\ &\leq \left(\sum_{n=m+1}^\infty |a_n|^2\right)^{1/2} \left(\sum_{n=m+1}^\infty \|(Te_n)\|^2\right)^{1/2} \\ &\leq \left\|\sum_{n=1}^\infty a_n e_n\right\| \left(\sum_{n=m+1}^\infty \|(Te_n)\|^2\right)^{1/2} \qquad (24.3) \end{aligned}$$

so $\|T - T_m\| \to 0$ and by the previous Theorem T is compact as a limit of compact operators. □

Corollary 24.4 (from the above proof)**.** *For a Hilbert–Schmidt operator* $\|T\| \leq \left(\sum_{n=m+1}^\infty \|(Te_n)\|^2\right)^{1/2}$.

Proof. Just consider difference of T and $T_0 = 0$ in (24.2)–(24.3). □

Example 24.5. An *integral operator* T on $L_2[0, 1]$ is defined by the formula:

$$(Tf)(x) = \int_0^1 K(x, y) f(y) \, dy, \qquad f(y) \in L_2[0, 1], \tag{24.4}$$

where the continuous on $[0, 1] \times [0, 1]$ function K is called the *kernel of integral operator*.

Theorem 24.6. *Integral operator* (24.4) *is Hilbert–Schmidt.*

Proof. Let $(e_n)_{-\infty}^\infty$ be an orthonormal basis of $L_2[0, 1]$, e.g. $(e^{2\pi i n t})_{n \in \mathbb{Z}}$. Let us consider the kernel $K_x(y) = K(x, y)$ as a function of the argument y depending from the parameter x. Then:

$$(Te_n)(x) = \int_0^1 K(x, y) e_n(y) \, dy = \int_0^1 K_x(y) e_n(y) \, dy = \langle K_x, \bar{e}_n \rangle.$$

So $\|Te_n\|^2 = \int_0^1 |\langle K_x, \bar{e}_n\rangle|^2 \, dx$. Consequently:

$$
\begin{aligned}
\sum_{-\infty}^{\infty} \|Te_n\|^2 &= \sum_{-\infty}^{\infty} \int_0^1 |\langle K_x, \bar{e}_n\rangle|^2 \, dx \\
&= \int_0^1 \sum_{1}^{\infty} |\langle K_x, \bar{e}_n\rangle|^2 \, dx \qquad (24.5)\\
&= \int_0^1 \|K_x\|^2 \, dx \\
&= \int_0^1 \int_0^1 |K(x,y)|^2 \, dx\, dy < \infty
\end{aligned}
$$

Exercise 24.7. Justify the exchange of summation and integration in (24.5).

□

Remark 24.8. The definition 24.5 and Theorem 24.6 work also for any $T : L_2[a,b] \to L_2[c,d]$ with a continuous kernel $K(x,y)$ on $[c,d] \times [a,b]$.

Definition 24.9. Define *Hilbert–Schmidt norm* of a Hilbert–Schmidt operator A by $\|A\|_{HS}^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2$ (it is independent of the choice of orthonormal basis $(e_n)_1^{\infty}$, see Question F.2).

Exercise* **24.10.** Show that set of Hilbert–Schmidt operators with the above norm is a Hilbert space and find the an expression for the inner product.

Example 24.11. Let $K(x,y) = x - y$, then

$$(Tf)(x) = \int_0^1 (x-y) f(y)\, dy = x \int_0^1 f(y)\, dy - \int_0^1 y f(y)\, dy$$

is a rank 2 operator. Furthermore:

$$
\begin{aligned}
\|T\|_{HS}^2 &= \int_0^1 \int_0^1 (x-y)^2 \, dx\, dy = \int_0^1 \left[\frac{(x-y)^3}{3}\right]_{x=0}^1 dy \\
&= \int_0^1 \frac{(1-y)^3}{3} + \frac{y^3}{3}\, dy = \left[-\frac{(1-y)^4}{12} + \frac{y^4}{12}\right]_0^1 = \frac{1}{6}.
\end{aligned}
$$

On the other hand there is an orthonormal basis such that

$$Tf = \frac{1}{\sqrt{12}} \langle f, e_1\rangle e_1 - \frac{1}{\sqrt{12}} \langle f, e_2\rangle e_2,$$

and $\|T\| = \frac{1}{\sqrt{12}}$ and $\sum_1^2 \|Te_k\|^2 = \frac{1}{6}$ and we get $\|T\| \le \|T\|_{HS}$ in agreement with Corollary 24.4.

Part 8. The spectral theorem for compact normal operators

Recall from Section 19 that an operator T is normal if $TT^* = T^*T$; Hermitian ($T^* = T$) and unitary ($T^* = T^{-1}$) operators are normal.

25. Spectrum of normal operators

Theorem 25.1. *Let $T \in B(H)$ be a normal operator then*

(i) $\ker T = \ker T^*$, *so* $\ker(T - \lambda I) = \ker(T^* - \bar{\lambda} I)$ *for all* $\lambda \in \mathbb{C}$

(ii) *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*

(iii) $\|T\| = r(T)$.

Proof. (i) Obviously:

$$\begin{aligned} x \in \ker T &\Leftrightarrow \langle Tx, Tx\rangle = 0 \Leftrightarrow \langle T^*Tx, x\rangle = 0 \\ &\Leftrightarrow \langle TT^*x, x\rangle = 0 \Leftrightarrow \langle T^*x, T^*x\rangle = 0 \\ &\Leftrightarrow x \in \ker T^*. \end{aligned}$$

The second part holds because normalities of T and $T - \lambda I$ are equivalent.

(ii) If $Tx = \lambda x$, $Ty = \mu y$ then from the previous statement $T^*y = \bar{\mu} y$. If $\lambda \neq \mu$ then the identity

$$\lambda \langle x, y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle = \mu \langle x, y\rangle$$

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

(iii) Let $S = T^*T$ then normality of T implies that S is Hermitian (check!). Consequently inequality

$$\|Sx\|^2 = \langle Sx, Sx\rangle = \langle S^2x, x\rangle \leq \|S^2\| \|x\|^2$$

implies $\|S\|^2 \leq \|S^2\|$. But the opposite inequality follows from the Theorem 17.2, thus we have the equality $\|S^2\| = \|S\|^2$ and more generally by induction: $\left\|S^{2^m}\right\| = \|S\|^{2^m}$ for all m.

Now we claim $\|S\| = \|T\|^2$. From Theorem 17.2 and 18.1 we get $\|S\| = \|T^*T\| \leq \|T\|^2$. On the other hand if $\|x\| = 1$ then

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

implies the opposite inequality $\|S\| \geq \|T\|^2$. And because $(T^{2^m})^*T^{2^m} = (T^*T)^{2^m}$ we get the equality

$$\left\|T^{2^m}\right\|^2 = \left\|(T^*T)^{2^m}\right\| = \|T^*T\|^{2^m} = \|T\|^{2^{m+1}}.$$

Thus:

$$r(T) = \lim_{m\to\infty} \left\|T^{2^m}\right\|^{1/2^m} = \lim_{m\to\infty} \|T\|^{2^{m+1}/2^{m+1}} = \|T\|.$$

by the spectral radius formula (21.1).

□

Example 25.2. It is easy to see that normality is important in 25.1.iii, indeed the non-normal operator T given by the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ in $\mathbb{C}$ has one-point spectrum $\{0\}$, consequently $r(T) = 0$ but $\|T\| = 1$.

Lemma 25.3. *Let T be a compact normal operator then*

(i) *The set of of eigenvalues of T is either finite or a countable sequence tending to zero.*

(ii) *All the eigenspaces, i.e.* $\ker(T - \lambda I)$, *are finite-dimensional for all* $\lambda \neq 0$.

Remark 25.4. This Lemma is true for any compact operator, but we will not use that in our course.

Proof. (i) Let H_0 be the closed linear span of eigenvectors of T. Then T restricted to H_0 is a diagonal compact operator with the same set of eigenvalues λ_n as in H. Then $\lambda_n \to 0$ from Proposition 23.8 .

Exercise 25.5. Use the proof of Proposition 23.8 to give a direct demonstration.

Solution. Or straightforwardly assume opposite: there exist an $\delta > 0$ and infinitely many eigenvalues λ_n such that $|\lambda_n| > \delta$. By the previous Theorem there is an orthonormal sequence v_n of corresponding eigenvectors $Tv_n = \lambda_n v_n$. Now the sequence (v_n) is bounded but its image $Tv_n = \lambda_n e_n$ has no convergent subsequence because for any $k \neq l$:

$$\|\lambda_k v_k - \lambda_l e_l\| = (|\lambda_k|^2 + |\lambda_l|^2)^{1/2} \geq \sqrt{2}\delta,$$

i.e. Te_{n_k} is not a Cauchy sequence, see Figure 16. □

(ii) Similarly if $H_0 = \ker(T - \lambda I)$ is infinite dimensional, then restriction of T on H_0 is λI—which is non-compact by Proposition 23.8. Alternatively consider the infinite orthonormal sequence (v_n), $Tv_n = \lambda v_n$ as in Exercise 25.5.

□

Lemma 25.6. *Let T be a compact normal operator. Then all non-zero points $\lambda \in \sigma(T)$ are eigenvalues and there exists an eigenvalue of modulus $\|T\|$.*

Proof. Assume without lost of generality that $T \neq 0$. Let $\lambda \in \sigma(T)$, without lost of generality (multiplying by a scalar) $\lambda = 1$.

We claim that if 1 is not an eigenvalue then there exist $\delta > 0$ such that

$$\|(I - T)x\| \geq \delta \|x\| . \qquad (25.1)$$

Otherwise there exists a sequence of vectors (x_n) with unit norm such that $(I - T)x_n \to 0$. Then from the compactness of T for a subsequence (x_{n_k}) there is $y \in H$ such that $Tx_{n_k} \to y$, then $x_n \to y$ implying $Ty = y$ and $y \neq 0$—i.e. y is eigenvector with eigenvalue 1.

Now we claim $\operatorname{Im}(I - T)$ is closed, i.e. $y \in \overline{\operatorname{Im}(I - T)}$ implies $y \in \operatorname{Im}(I - T)$. Indeed, if $(I - T)x_n \to y$, then there is a subsequence (x_{n_k}) such that $Tx_{n_k} \to z$ implying $x_{n_k} \to y + z$, then $(I - T)(z + y) = y$.

Finally $I - T$ is injective, i.e $\ker(I - T) = \{0\}$, by (25.1). By the property 25.1.i, $\ker(I - T^*) = \{0\}$ as well. But because always $\ker(I - T^*) = \operatorname{Im}(I - T)^\perp$ (check!) we got surjectivity, i.e. $\operatorname{Im}(I - T)^\perp = \{0\}$, of $I - T$. Thus $(I - T)^{-1}$ exists and is bounded because (25.1) implies $\|y\| > \delta \|(I - T)^{-1}y\|$. Thus $1 \notin \sigma(T)$.

The existence of eigenvalue λ such that $|\lambda| = \|T\|$ follows from combination of Lemma 21.4 and Theorem 25.1.iii. □

26. Compact normal operators

Theorem 26.1 (The spectral theorem for compact normal operators)**.** *Let T be a compact normal operator on a Hilbert space H. Then there exists an orthonormal sequence (e_n) of eigenvectors of T and corresponding eigenvalues (λ_n) such that:*

$$Tx = \sum_n \lambda_n \langle x, e_n \rangle e_n, \qquad \text{for all } x \in H. \qquad (26.1)$$

If (λ_n) is an infinite sequence it tends to zero.

Conversely, if T is given by a formula (26.1) *then it is compact and normal.*

Proof. Suppose $T \neq 0$. Then by the previous Theorem there exists an eigenvalue λ_1 such that $|\lambda_1| = \|T\|$ with corresponding eigenvector e_1 of the unit norm. Let $H_1 = \operatorname{Lin}(e_1)^\perp$. If $x \in H_1$ then

$$\langle Tx, e_1 \rangle = \langle x, T^* e_1 \rangle = \left\langle x, \bar{\lambda}_1 e_1 \right\rangle = \lambda_1 \langle x, e_1 \rangle = 0, \qquad (26.2)$$

thus $Tx \in H_1$ and similarly $T^*x \in H_1$. Write $T_1 = T|_{H_1}$ which is again a normal compact operator with a norm does not exceeding $\|T\|$. We could inductively repeat this procedure for T_1 obtaining sequence of eigenvalues $\lambda_2, \lambda_3, \ldots$ with eigenvectors $e_2, e_3, \ldots$. If $T_n = 0$ for a finite n then theorem is already proved. Otherwise we have an infinite sequence $\lambda_n \to 0$. Let

$$x = \sum_1^n \langle x, e_k \rangle e_k + y_n \quad \Rightarrow \quad \|x\|^2 = \sum_1^n |\langle x, e_k \rangle|^2 + \|y_n\|^2, \qquad y_n \in H_n,$$

from Pythagoras's theorem. Then $\|y_n\| \le \|x\|$ and $\|Ty_n\| \le \|T_n\| \|y_n\| \le |\lambda_n| \|x\| \to 0$ by Lemma 25.3. Thus

$$Tx = \lim_{n\to\infty} \left(\sum_1^n \langle x, e_n \rangle Te_n + Ty_n \right) = \sum_1^\infty \lambda_n \langle x, e_n \rangle e_n$$

Conversely, if $Tx = \sum_1^\infty \lambda_n \langle x, e_n \rangle e_n$ then

$$\langle Tx, y \rangle = \sum_1^\infty \lambda_n \langle x, e_n \rangle \langle e_n, y \rangle = \sum_1^\infty \langle x, e_n \rangle \lambda_n \overline{\langle y, e_n \rangle},$$

thus $T^*y = \sum_1^\infty \bar{\lambda}_n \langle y, e_n \rangle e_n$. Then we got the normality of T: $T^*Tx = TT^*x = \sum_1^\infty |\lambda_n|^2 \langle y, e_n \rangle e_n$. Also T is compact because it is a uniform limit of the finite rank operators $T_n x = \sum_1^n \lambda_n \langle x, e_n \rangle e_n$. □

Corollary 26.2. *Let T be a compact normal operator on a separable Hilbert space H, then there exists a orthonormal basis g_k such that*

$$Tx = \sum_1^\infty \lambda_n \langle x, g_n \rangle g_n,$$

and λ_n are eigenvalues of T including zeros.

Proof. Let (e_n) be the orthonormal sequence constructed in the proof of the previous Theorem. Then x is perpendicular to all e_n if and only if its in the kernel of T. Let (f_n) be any orthonormal basis of $\ker T$. Then the union of (e_n) and (f_n) is the orthonormal basis (g_n) we have looked for. □

Exercise 26.3. Finish all details in the above proof.

Corollary 26.4 (Singular value decomposition)**.** *If T is any compact operator on a separable Hilbert space then there exists orthonormal sequences (e_k) and (f_k) such that $Tx = \sum_k \mu_k \langle x, e_k \rangle f_k$ where (μ_k) is a sequence of positive numbers such that $\mu_k \to 0$ if it is an infinite sequence.*

Proof. Operator T^*T is compact and Hermitian (hence normal). From the previous Corollary there is an orthonormal basis (e_k) such that $T^*Tx = \sum_n \lambda_n \langle x, e_k \rangle e_k$ for some positive $\lambda_n = \|Te_n\|^2$. Let $\mu_n = \|Te_n\|$ and $f_n = Te_n/\mu_n$. Then f_n is an orthonormal sequence (check!) and

$$Tx = \sum_n \langle x, e_n \rangle Te_n = \sum_n \langle x, e_n \rangle \mu_n f_n.$$

□

Corollary 26.5. *A bounded operator in a Hilber space is compact if and only if it is a uniform limit of the finite rank operators.*

Proof. Sufficiency follows from 23.9. *Necessity*: by the previous Corollary $Tx = \sum_n \langle x, e_n \rangle \mu_n f_n$ thus T is a uniform limit of operators $T_m x = \sum_{n=1}^m \langle x, e_n \rangle \mu_n f_n$ which are of finite rank. □

Part 9. Applications to integral equations

In this lecture we will study the *Fredholm equation* defined as follows. Let the *integral operator* with a *kernel* $K(x, y)$ defined on $[a, b] \times [a, b]$ be defined as before:

$$(T\phi)(x) = \int_a^b K(x, y)\phi(y)\, dy. \tag{26.3}$$

The Fredholm equation of the *first* and *second* kinds correspondingly are:

$$T\phi = f \qquad \text{and} \qquad \phi - \lambda T\phi = f, \tag{26.4}$$

for a function f on $[a,b]$. A special case is given by *Volterra equation* by an operator integral operator (26.3) T with a kernel $K(x,y)=0$ for all $y>x$ which could be written as:

$$(T\phi)(x)=\int_a^x K(x,y)\phi(y)\,dy. \tag{26.5}$$

We will consider integral operators with kernels K such that $\int_a^b\int_a^b K(x,y)\,dx\,dy<\infty$, then by Theorem 24.6 T is a Hilbert–Schmidt operator and in particular bounded.

As a reason to study Fredholm operators we will mention that solutions of differential equations in mathematical physics (notably heat and wave equations) requires a decomposition of a function f as a linear combination of functions $K(x,y)$ with "coefficients" ϕ. This is an continuous analog of a discrete decomposition into Fourier series.

Using ideas from the proof of Lemma 20.4 we define *Neumann series* for the resolvent:

$$(I-\lambda T)^{-1}=I+\lambda T+\lambda^2T^2+\cdots, \tag{26.6}$$

which is valid for all $\lambda<\|T\|^{-1}$.

Example 26.1. Solve the Volterra equation

$$\phi(x)-\lambda\int_0^x y\phi(y)\,dy=x^2, \qquad \text{on } L_2[0,1].$$

In this case $I-\lambda T\phi=f$, with $f(x)=x^2$ and:

$$K(x,y)=\begin{cases} y, & 0\le y\le x;\\ 0, & x<y\le 1.\end{cases}$$

Straightforward calculations shows:

$$\begin{aligned}(Tf)(x) &= \int_0^x y\cdot y^2\,dy=\frac{x^4}{4},\\ (T^2f)(x) &= \int_0^x y\frac{y^4}{4}\,dy=\frac{x^6}{24},\ldots\end{aligned}$$

and generally by induction:

$$(T^nf)(x)=\int_0^x y\frac{y^{2n}}{2^{n-1}n!}\,dy=\frac{x^{2n+2}}{2^n(n+1)!}.$$

Hence:

$$\begin{aligned}\phi(x) &= \sum_0^\infty \lambda^nT^nf=\sum_0^\infty\frac{\lambda^nx^{2n+2}}{2^n(n+1)!}\\ &= \frac{2}{\lambda}\sum_0^\infty\frac{\lambda^{n+1}x^{2n+2}}{2^{n+1}(n+1)!}\\ &= \frac{2}{\lambda}(e^{\lambda x^2/2}-1) \qquad \text{for all } \lambda\in\mathbb{C}\setminus\{0\},\end{aligned}$$

because in this case $r(T)=0$. For the Fredholm equations this is not always the case, see Tutorial problem F.4.

Among other integral operators there is an important subclass with *separable kernel*, namely a kernel which has a form:

$$K(x,y)=\sum_{j=1}^{n} g_j(x)h_j(y). \tag{26.7}$$

In such a case:

$$\begin{aligned}(T\phi)(x) &= \int_a^b \sum_{j=1}^{n} g_j(x)h_j(y)\phi(y)\,dy\\ &= \sum_{j=1}^{n} g_j(x)\int_a^b h_j(y)\phi(y)\,dy,\end{aligned}$$

i.e. the image of T is spanned by $g_1(x)$, ..., $g_n(x)$ and is finite dimensional, consequently the solution of such equation reduces to linear algebra.

Example 26.2. Solve the Fredholm equation (actually find eigenvectors of T):

$$\begin{aligned}\phi(x) &= \lambda\int_0^{2\pi} \cos(x+y)\phi(y)\,dy\\ &= \lambda\int_0^{2\pi} (\cos x\cos y-\sin x\sin y)\phi(y)\,dy.\end{aligned}$$

Clearly $\phi(x)$ should be a linear combination $\phi(x)=A\cos x+B\sin x$ with coefficients A and B satisfying to:

$$\begin{aligned}A &= \lambda\int_0^{2\pi} \cos y(A\cos y+B\sin y)\,dy,\\ B &= -\lambda\int_0^{2\pi} \sin y(A\cos y+B\sin y)\,dy.\end{aligned}$$

Basic calculus implies $A=\lambda\pi A$ and $B=-\lambda\pi B$ and the only nonzero solutions are:

$$\begin{aligned}\lambda=\pi^{-1}\quad A\neq 0\quad B=0\\ \lambda=-\pi^{-1}\quad A=0\quad B\neq 0\end{aligned}$$

We develop some Hilbert–Schmidt theory for integral operators.

Theorem 26.3. *Suppose that $K(x,y)$ is a continuous function on $[a,b]\times[a,b]$ and $K(x,y)=\overline{K(y,x)}$ and operator T is defined by (26.3). Then*

(i) *T is a self-adjoint Hilbert–Schmidt operator.*
(ii) *All eigenvalues of T are real and satisfy $\sum_n \lambda_n^2<\infty$.*
(iii) *The eigenvectors v_n of T can be chosen as an orthonormal basis of $L_2[a,b]$, are continuous for nonzero λ_n and*

$$T\phi=\sum_{n=1}^{\infty}\lambda_n\langle\phi,v_n\rangle v_n \qquad \text{where} \qquad \phi=\sum_{n=1}^{\infty}\langle\phi,v_n\rangle v_n$$

Proof. (i) The condition $K(x,y)=\overline{K(y,x)}$ implies the Hermitian property of T:

$$\begin{aligned}\langle T\phi,\psi\rangle &= \int_a^b \left(\int_a^b K(x,y)\phi(y)\,dy\right)\bar{\psi}(x)\,dx\\ &= \int_a^b\int_a^b K(x,y)\phi(y)\bar{\psi}(x)\,dx\,dy\\ &= \int_a^b \phi(y)\left(\int_a^b \overline{K(y,x)\psi(x)}\,dx\right)dy\\ &= \langle \phi, T\psi\rangle .\end{aligned}$$

The Hilbert–Schmidt property (and hence compactness) was proved in Theorem 24.6.

(ii) Spectrum of T is real as for any Hermitian operator, see Theorem 22.1.ii and finiteness of $\sum_n \lambda_n^2$ follows from Hilbert–Schmidt property

(iii) The existence of orthonormal basis consisting from eigenvectors (v_n) of T was proved in Corollary 26.2. If $\lambda_n\neq 0$ then:

$$\begin{aligned}v_n(x_1)-v_n(x_2) &= \lambda_n^{-1}((Tv_n)(x_1)-(Tv_n)(x_2))\\ &= \frac{1}{\lambda_n}\int_a^b (K(x_1,y)-K(x_2,y))v_n(y)\,dy\end{aligned}$$

and by Cauchy–Schwarz-Bunyakovskii inequality:

$$|v_n(x_1)-v_n(x_2)|\le \frac{1}{|\lambda_n|}\|v_n\|_2\int_a^b |K(x_1,y)-K(x_2,y)|\,dy$$

which tense to 0 due to (uniform) continuity of $K(x,y)$. □

Theorem 26.4. *Let T be as in the previous Theorem. Then if $\lambda\neq 0$ and $\lambda^{-1}\notin\sigma(T)$, the unique solution ϕ of the* Fredholm equation of the second kind $\phi-\lambda T\phi=f$ *is*

$$\phi=\sum_1^\infty \frac{\langle f,v_n\rangle}{1-\lambda\lambda_n}v_n. \tag{26.8}$$

Proof. Let $\phi=\sum_1^\infty a_n v_n$ where $a_n=\langle\phi,v_n\rangle$, then

$$\phi-\lambda T\phi=\sum_1^\infty a_n(1-\lambda\lambda_n)v_n=f=\sum_1^\infty\langle f,v_n\rangle v_n$$

if and only if $a_n=\langle f,v_n\rangle/(1-\lambda\lambda_n)$ for all n. Note $1-\lambda\lambda_n\neq 0$ since $\lambda^{-1}\notin\sigma(T)$.

Because $\lambda_n\to 0$ we got $\sum_1^\infty |a_n|^2$ by its comparison with $\sum_1^\infty|\langle f,v_n\rangle|^2=\|f\|^2$, thus the solution exists and is unique by the Riesz–Fisher Theorem. □

See Exercise F.5 for an example.

Theorem 26.5 (Fredholm alternative)**.** *Let $T\in K(H)$ be compact normal and $\lambda\in\mathbb{C}\setminus\{0\}$. Consider the equations:*

$$\phi-\lambda T\phi = 0 \tag{26.9}$$

$$\phi-\lambda T\phi = f \tag{26.10}$$

then ***either***

(A) *the only solution to* (26.9) *is* $\phi = 0$ *and* (26.10) *has a unique solution for any* $f \in H$; ***or***
(B) *there exists a nonzero solution to* (26.9) *and* (26.10) *can be solved if and only if* f *is orthogonal all solutions to* (26.9).

Proof. (A) If $\phi = 0$ is the only solution of (26.9), then λ^{-1} is not an eigenvalue of T and then by Lemma 25.6 is neither in spectrum of T. Thus $I - \lambda T$ is invertible and the unique solution of (26.10) is given by $\phi = (I - \lambda T)^{-1} f$.

(B) A nonzero solution to (26.9) means that $\lambda^{-1} \in \sigma(T)$. Let (v_n) be an orthonormal basis of eigenvectors of T for eigenvalues (λ_n). By Lemma 25.3.ii only a finite number of λ_n is equal to λ^{-1}, say they are $\lambda_1, \ldots, \lambda_N$, then

$$(I - \lambda T)\phi = \sum_{n=1}^{\infty}(1 - \lambda\lambda_n) \langle \phi, v_n\rangle v_n = \sum_{n=N+1}^{\infty}(1 - \lambda\lambda_n) \langle \phi, v_n\rangle v_n.$$

If $f = \sum_1^\infty \langle f, v_n\rangle v_n$ then the identity $(I - \lambda T)\phi = f$ is only possible if $\langle f, v_n\rangle = 0$ for $1 \le n \le N$. Conversely from that condition we could give a solution

$$\phi = \sum_{n=N+1}^{\infty} \frac{\langle f, v_n\rangle}{1 - \lambda\lambda_n} v_n + \phi_0, \qquad \text{for any } \phi_0 \in \mathrm{Lin}(v_1, \ldots, v_N),$$

which is again in H because $f \in H$ and $\lambda_n \to 0$.

□

Example 26.6. Let us consider

$$(T\phi)(x) = \int_0^1 (2xy - x - y + 1)\phi(y)\, dy.$$

Because the kernel of T is real and symmetric $T = T^*$, the kernel is also separable:

$$(T\phi)(x) = x\int_0^1 (2y - 1)\phi(y)\, dy + \int_0^1 (-y + 1)\phi(y)\, dy,$$

and T of the rank 2 with image of T spanned by 1 and x. By direct calculations:

$$\begin{array}{rcl} T: \ 1 & \mapsto & \frac{1}{2} \\ T: \ x & \mapsto & \frac{1}{6}x + \frac{1}{6}, \end{array} \qquad \text{or } T \text{ is given by the matrix} \qquad \begin{pmatrix} \frac{1}{2} & \frac{1}{6} \\ 0 & \frac{1}{6} \end{pmatrix}$$

According to linear algebra decomposition over eigenvectors is:

$$\lambda_1 = \frac{1}{2} \quad \text{with vector} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\lambda_2 = \frac{1}{6} \quad \text{with vector} \quad \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix}$$

with normalisation $v_1(y) = 1$, $v_2(y) = \sqrt{12}(y - 1/2)$ and we complete it to an orthonormal basis (v_n) of $L_2[0, 1]$. Then

- If $\lambda \ne 2$ or 6 then $(I - \lambda T)\phi = f$ has a unique solution (cf. equation (26.8)):

$$\begin{aligned} \phi &= \sum_{n=1}^{2} \frac{\langle f, v_n\rangle}{1 - \lambda\lambda_n} v_n + \sum_{n=3}^{\infty} \langle f, v_n\rangle v_n \\ &= \sum_{n=1}^{2} \frac{\langle f, v_n\rangle}{1 - \lambda\lambda_n} v_n + \left(f - \sum_{n=1}^{2} \langle f, v_n\rangle v_n\right) \\ &= f + \sum_{n=1}^{2} \frac{\lambda\lambda_n}{1 - \lambda\lambda_n} \langle f, v_n\rangle\, v_n. \end{aligned}$$

- If $\lambda = 2$ then the solutions exist provided $\langle f, v_1 \rangle = 0$ and are:
$$\phi = f + \frac{\lambda\lambda_2}{1-\lambda\lambda_2} \langle f, v_2 \rangle v_2 + Cv_1 = f + \frac{1}{2} \langle f, v_2 \rangle v_2 + Cv_1, \qquad C \in \mathbb{C}.$$
- If $\lambda = 6$ then the solutions exist provided $\langle f, v_2 \rangle = 0$ and are:
$$\phi = f + \frac{\lambda\lambda_1}{1-\lambda\lambda_1} \langle f, v_1 \rangle v_1 + Cv_2 = f - \frac{3}{2} \langle f, v_2 \rangle v_2 + Cv_2, \qquad C \in \mathbb{C}.$$

Part 10. Tutorial Problems

These are tutorial problems intended for self-assessment of the course understanding.

Appendix A. Tutorial problems I

All spaces are complex, unless otherwise specified.

A.1. Show that $\|f\| = |f(0)| + \sup |f'(t)|$ defines a norm on $C^1[0,1]$, which is the space of (real) functions on $[0,1]$ with continuous derivative.

A.2. Show that the formula $\langle (x_n), (y_n) \rangle = \sum_{n=1}^{\infty} x_n \overline{y}_n / n^2$ defines an inner product on ℓ_∞, the space of bounded (complex) sequences. What norm does it produce?

A.3. Use the Cauchy–Schwarz inequality for a suitable inner product to prove that for all $f \in C[0,1]$ the inequality

$$\left| \int_0^1 f(x) x \, dx \right| \leq C \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2}$$

holds for some constant $C > 0$ (independent of f) and find the smallest possible C that holds for all functions f (hint: consider the cases of equality).

A.4. We define the following norm on ℓ_∞, the space of bounded complex sequences:

$$\|(x_n)\|_\infty = \sup_{n \geq 1} |x_n|.$$

Show that this norm makes ℓ_∞ into a Banach space (i.e., a complete normed space).

A.5. Fix a vector $(w_1, \ldots, w_n)$ whose components are strictly positive real numbers, and define an inner product on $\mathbb{C}^n$ by

$$\langle x, y \rangle = \sum_{k=1}^{n} w_k x_k \overline{y}_k.$$

Show that this makes $\mathbb{C}^n$ into a Hilbert space (i.e., a complete inner-product space).

Appendix B. Tutorial problems II

B.1. Show that the supremum norm on $C[0,1]$ isn't given by an inner product, by finding a counterexample to the parallelogram law.

B.2. In ℓ_2 let $e_1 = (1,0,0,\ldots)$, $e_2 = (0,1,0,0,\ldots)$, $e_3 = (0,0,1,0,0,\ldots)$, and so on. Show that $\mathrm{Lin}\,(e_1, e_2, \ldots) = c_{00}$, and that $\mathrm{CLin}\,(e_1, e_2, \ldots) = \ell_2$. What is $\mathrm{CLin}\,(e_2, e_3, \ldots)$?

B.3. Let $C[-1,1]$ have the standard L_2 inner product, defined by

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

Show that the functions 1, t and $t^2 - 1/3$ form an orthogonal (not orthonormal!) basis for the subspace P_2 of polynomials of degree at most 2 and hence calculate the best L_2-approximation of the function t^4 by polynomials in P_2.

B.4. Define an inner product on $C[0,1]$ by

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

Use the Gram–Schmidt process to find the first 2 terms of an orthonormal sequence formed by orthonormalising the sequence 1, t, t^2, ...

B.5. Consider the plane P in $\mathbb{C}^4$ (usual inner product) spanned by the vectors $(1,1,0,0)$ and $(1,0,0,-1)$. Find orthonormal bases for P and $P^\perp$, and verify directly that $(P^\perp)^\perp = P$.

Appendix C. Tutorial Problems III

C.1. Let a and b be arbitrary real numbers with $a < b$. By using the fact that the functions $\frac{1}{\sqrt{2\pi}}e^{inx}$, $n \in \mathbb{Z}$, are orthonormal in $L_2[0, 2\pi]$, together with the change of variable $x = 2\pi(t-a)/(b-a)$, find an orthonormal basis in $L_2[a,b]$ of the form $e_n(t) = \alpha e^{in\lambda t}$, $n \in \mathbb{Z}$, for suitable real constants α and λ.

C.2. For which real values of α is

$$\sum_{n=1}^{\infty} n^\alpha e^{int}$$

the Fourier series of a function in $L_2[-\pi,\pi]$?

C.3. Calculate the Fourier series of $f(t) = e^t$ on $[-\pi,\pi]$ and use Parseval's identity to deduce that

$$\sum_{n=-\infty}^{\infty} \frac{1}{n^2+1} = \frac{\pi}{\tanh \pi}.$$

C.4. Using the fact that (e_n) is a complete orthonormal system in $L_2[-\pi,\pi]$, where $e_n(t) = \exp(int)/\sqrt{2\pi}$, show that $e_0, s_1, c_1, s_2, c_2, \ldots$ is a complete orthonormal system, where $s_n(t) = \sin nt/\sqrt{\pi}$ and $c_n(t) = \cos nt/\sqrt{\pi}$. Show that every $L_2[-\pi,\pi]$ function f has a Fourier series

$$a_0 + \sum_{n=1}^{\infty} a_n \cos nt + b_n \sin nt,$$

converging in the L_2 sense, and give a formula for the coefficients.

C.5. Let $C(\mathbb{T})$ be the space of continuous (complex) functions on the circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ with the supremum norm. Show that, for any polynomial $f(z)$ in $C(\mathbb{T})$

$$\int_{|z|=1} f(z)\, dz = 0.$$

Deduce that the function $f(z) = \bar{z}$ is *not* the uniform limit of polynomials on the circle (i.e., Weierstrass's approximation theorem doesn't hold in this form).

Appendix D. Tutorial Problems IV

D.1. Define a linear functional on $C[0,1]$ (continuous functions on $[0,1]$) by $\alpha(f) = f(1/2)$. Show that α is bounded if we give $C[0,1]$ the supremum norm. Show that α is not bounded if we use the L_2 norm, because we can find a sequence (f_n) of continuous functions on $[0,1]$ such that $\|f_n\|_2 \le 1$, but $f_n(1/2) \to \infty$.

D.2. The *Hardy space* H_2 is the Hilbert space of all power series $f(z) = \sum_{n=0}^\infty a_n z^n$, such that $\sum_{n=0}^\infty |a_n|^2 < \infty$, where the inner product is given by

$$\left\langle \sum_{n=0}^{\infty} a_n z^n, \sum_{n=0}^{\infty} b_n z^n \right\rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}.$$

Show that the sequence $1, z, z^2, z^3, \ldots$ is an orthonormal basis for H_2.

Fix w with $|w| < 1$ and define a linear functional on H_2 by $\alpha(f) = f(w)$. Write down a formula for the function $g(z) \in H_2$ such that $\alpha(f) = \langle f, g\rangle$. What is $\|\alpha\|$?

D.3. The *Volterra operator* $V : L_2[0,1] \to L_2[0,1]$ is defined by

$$(Vf)(x) = \int_0^x f(t)\,dt.$$

Use the Cauchy–Schwarz inequality to show that $|(Vf)(x)| \le \sqrt{x}\|f\|_2$ (hint: write $(Vf)(x) = \langle f, J_x\rangle$ where J_x is a function that you can write down explicitly).

Deduce that $\|Vf\|_2^2 \le \frac{1}{2}\|f\|_2^2$, and hence $\|V\| \le 1/\sqrt{2}$.

D.4. Find the adjoints of the following operators:

(i) $A : \ell_2 \to \ell_2$, defined by $A(x_1, x_2, \ldots) = (0, \frac{x_1}{1}, \frac{x_2}{2}, \frac{x_3}{3}, \ldots)$;
 and, on a general Hilbert space H:

(ii) The rank-one operator R, defined by $Rx = \langle x, y\rangle z$, where y and z are fixed elements of H;

(iii) The projection operator P_M, defined by $P_M(m+n) = m$, where $m \in M$ and $n \in M^\perp$, and $H = M \oplus M^\perp$ as usual.

D.5. Let $U \in B(H)$ be a unitary operator. Show that (Ue_n) is an orthonormal basis of H whenever (e_n) is.

Let $\ell_2(\mathbb{Z})$ denote the Hilbert space of two-sided sequences $(a_n)_{n=-\infty}^{\infty}$ with

$$\|(a_n)\|^2 = \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty.$$

Show that the *bilateral right shift*, $V : \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z})$ defined by $V((a_n)) = (b_n)$, where $b_n = a_{n-1}$ for all $n \in \mathbb{Z}$, is unitary, whereas the usual right shift S on $\ell_2 = \ell_2(\mathbb{N})$ is not unitary.

Appendix E. Tutorial Problems V

E.1. Let $f \in C[-\pi,\pi]$ and let M_f be the multiplication operator on $L_2(-\pi,\pi)$, given by $(M_f g)(t) = f(t)g(t)$, for $g \in L_2(-\pi,\pi)$. Find a function $\tilde{f} \in C[-\pi,\pi]$ such that $M_f^* = M_{\tilde{f}}$.

Show that M_f is always a normal operator. When is it Hermitian? When is it unitary?

E.2. Let T be any operator such that $T^n = 0$ for some integer n (such operators are called *nilpotent*). Show that $I - T$ is invertible (hint: consider $I + T + T^2 + \ldots + T^{n-1}$). Deduce that $I - T/\lambda$ is invertible for any $\lambda \ne 0$.

What is $\sigma(T)$? What is $r(T)$?

E.3. Let (λ_n) be a fixed bounded sequence of complex numbers, and define an operator on ℓ_2 by $T((x_n)) = ((y_n))$, where $y_n = \lambda_n x_n$ for each n. Recall that T is a bounded operator and $\|T\| = \|(\lambda_n)\|_\infty$. Let $\Lambda = \{\lambda_1, \lambda_2, \ldots\}$. Prove the following:

(i) Each λ_k is an eigenvalue of T, and hence is in $\sigma(T)$.

(ii) If $\lambda \notin \overline{\Lambda}$, then the inverse of $T - \lambda I$ exists (and is bounded).

Deduce that $\sigma(T) = \overline{\Lambda}$. Note, that then *any non-empty compact set could be a spectrum of some bounden operator*.

E.4. Let S be an *isomorphism* between Hilbert spaces H and K, that is, $S : H \to K$ is a linear bijection such that S and S^{-1} are bounded operators. Suppose that $T \in B(H)$. Show that T and STS^{-1} have the same spectrum and the same eigenvalues (if any).

E.5. Define an operator $U : \ell_2(\mathbb{Z}) \to L_2(-\pi,\pi)$ by $U((a_n)) = \sum_{n=-\infty}^{\infty} a_n e^{int}/\sqrt{2\pi}$. Show that U is a bijection and an isometry, i.e., that $\|Ux\| = \|x\|$ for all $x \in \ell_2(\mathbb{Z})$.

Let V be the bilateral right shift on $\ell_2(\mathbb{Z})$, the unitary operator defined on Question D.5. Let $f \in L_2(-\pi,\pi)$. Show that $(UVU^{-1}f)(t) = e^{it}f(t)$, and hence, using Question E.4, show that $\sigma(V) = \mathbb{T}$, the unit circle, but that V has no eigenvalues.

Appendix F. Tutorial Problems VI

F.1. Show that $K(X)$ is a closed linear subspace of $B(X)$, and that AT and TA are compact whenever $T \in K(X)$ and $A \in B(X)$. (This means that $K(X)$ is a closed ideal of $B(X)$.)

F.2. Let A be a Hilbert–Schmidt operator, and let $(e_n)_{n\geq 1}$ and $(f_m)_{m\geq 1}$ be orthonormal bases of A. By writing each Ae_n as $Ae_n = \sum_{m=1}^{\infty} \langle Ae_n, f_m\rangle f_m$, show that

$$\sum_{n=1}^{\infty} \|Ae_n\|^2 = \sum_{m=1}^{\infty} \|A^* f_m\|^2.$$

Deduce that the quantity $\|A\|_{HS}^2 = \sum_{n=1}^{\infty} \|Ae_n\|^2$ is independent of the choice of orthonormal basis, and that $\|A\|_{HS} = \|A^*\|_{HS}$. ($\|A\|_{HS}$ is called the *Hilbert–Schmidt norm* of A.)

F.3.
(i) Let $T \in K(H)$ be a compact operator. Using Question F.1, show that T^*T and TT^* are compact Hermitian operators.
(ii) Let $(e_n)_{n\geq 1}$ and $(f_n)_{n\geq 1}$ be orthonormal bases of a Hilbert space H, let $(\alpha_n)_{n\geq 1}$ be any bounded complex sequence, and let $T \in B(H)$ be an operator defined by

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

Prove that T is Hilbert–Schmidt precisely when $(\alpha_n) \in \ell_2$. Show that T is a compact operator if and only if $\alpha_n \to 0$, and in this case write down spectral decompositions for the compact Hermitian operators T^*T and TT^*.

F.4. Solve the Fredholm integral equation $\phi - \lambda T\phi = f$, where $f(x) = x$ and

$$(T\phi)(x) = \int_0^1 xy^2 \phi(y)\, dy \qquad (\phi \in L_2(0,1)),$$

for small values of λ by means of the Neumann series.

For what values of λ does the series converge? Write down a solution which is valid for all λ apart from one exception. What is the exception?

F.5. Suppose that h is a 2π-periodic $L_2(-\pi,\pi)$ function with Fourier series $\sum_{n=-\infty}^{\infty} a_n e^{int}$. Show that each of the functions $\phi_k(y) = e^{iky}$, $k \in \mathbb{Z}$, is an eigenvector of the integral operator T on $L_2(-\pi,\pi)$ defined by

$$(T\phi)(x) = \int_{-\pi}^{\pi} h(x-y)\phi(y)\, dy,$$

and calculate the corresponding eigenvalues.

Now let $h(t) = -\log(2(1-\cos t))$. Assuming, without proof, that $h(t)$ has the Fourier series $\sum_{n\in\mathbb{Z}, n\neq 0} e^{int}/|n|$, use the Hilbert–Schmidt method to solve the Fredholm equation $\phi - \lambda T\phi = f$, where $f(t)$ has Fourier series $\sum_{n=-\infty}^{\infty} c_n e^{int}$ and $1/\lambda \notin \sigma(T)$.

Appendix G. Tutorial Problems VII

G.1. Use the Gram–Schmidt algorithm to find an orthonormal basis for the subspace X of $L_2(-1,1)$ spanned by the functions t, t^2 and t^4.

Hence find the best $L_2(-1,1)$ approximation of the constant function $f(t) = 1$ by functions from X.

G.2. For $n = 1, 2, \ldots$ let ϕ_n denote the linear functional on ℓ_2 defined by

$$\phi_n(x) = x_1 + x_2 + \ldots + x_n,$$

where $x = (x_1, x_2, \ldots) \in \ell_2$. Use the Riesz–Fréchet theorem to calculate $\|\phi_n\|$.

G.3. Let T be a bounded linear operator on a Hilbert space, and suppose that $T = A + iB$, where A and B are self-adjoint operators. Express T^* in terms of A and B, and hence solve for A and B in terms of T and T^*.

Deduce that every operator T can be written $T = A + iB$, where A and B are self-adjoint, in a unique way.

Show that T is normal if and only if $AB = BA$.

G.4. Let P_n be the subspace of $L_2(-\pi, \pi)$ consisting of all polynomials of degree at most n, and let T_n be the subspace consisting of all trigonometric polynomials of the form $f(t) = \sum_{k=-n}^{n} a_k e^{ikt}$. Calculate the spectrum of the differentiation operator D, defined by $(Df)(t) = f'(t)$, when

(i) D is regarded as an operator on P_n, and
(ii) D is regarded as an operator on T_n.

Note that both P_n and T_n are finite-dimensional Hilbert spaces.

Show that T_n has an orthonormal basis of eigenvectors of D, whereas P_n does not.

G.5. Use the Neumann series to solve the Volterra integral equation $\phi - \lambda T\phi = f$ in $L_2[0, 1]$, where $\lambda \in \mathbb{C}$, $f(t) = 1$ for all t, and $(T\phi)(x) = \int_0^x t^2\phi(t)\, dt$. (You should be able to sum the infinite series.)

Part 11. Solutions of Tutorial Problems

Solutions of the tutorial problems will be distributed due in time on the paper.

Part 12. Course in the Nutshell

Please let me know of any errors, omissions or obscurities!

Appendix H. Some useful results and formulae (1)

H.1. A *norm* on a vector space, $\|x\|$, satisfies $\|x\| \geq 0$, $\|x\| = 0$ if and only if $x = 0$, $\|\lambda x\| = |\lambda|\,\|x\|$, and $\|x + y\| \leq \|x\| + \|y\|$ (*triangle inequality*). A norm defines a metric and a complete normed space is called a *Banach space.*

H.2. An *inner-product space* is a vector space (usually complex) with a scalar product on it, $\langle x, y\rangle \in \mathbb{C}$ such that $\langle x, y\rangle = \overline{\langle y, x\rangle}$, $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$, $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$, $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ if and only if $x = 0$. This defines a norm by $\|x\|^2 = \langle x, x\rangle$. A complete inner-product space is called a *Hilbert space.* A Hilbert space is automatically a Banach space.

H.3. The *Cauchy–Schwarz inequality.* $|\langle x, y\rangle| \leq \|x\|\,\|y\|$ with equality if and only if x and y are linearly dependent.

H.4. Some *examples of Hilbert spaces.* (i) Euclidean $\mathbb{C}^n$. (ii) ℓ_2, sequences (a_k) with $\|(a_k)\|_2^2 = \sum |a_k|^2 < \infty$. In both cases $\langle (a_k), (b_k)\rangle = \sum a_k\overline{b_k}$. (iii) $L_2[a, b]$, functions on $[a, b]$ with $\|f\|_2^2 = \int_a^b |f(t)|^2\, dt < \infty$. Here $\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\, dt$. (iv) Any closed subspace of a Hilbert space.

H.5. Other *examples of Banach spaces.* (i) $C_b(X)$, continuous bounded functions on a topological space X. (ii) $\ell_\infty(X)$, all bounded functions on a set X. The supremum norms on $C_b(X)$ and $\ell_\infty(X)$ make them into Banach spaces. (iii) Any closed subspace of a Banach space.

H.6. On *incomplete spaces.* The inner-product (L_2) norm on $C[0, 1]$ is incomplete. c_{00} (sequences eventually zero), with the ℓ_2 norm, is another incomplete i.p.s.

H.7. The *parallelogram identity.* $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ in an inner-product space. Not in general normed spaces.

H.8. On *subspaces.* Complete $\Longrightarrow$ closed. The closure of a linear subspace is still a linear subspace. $\operatorname{Lin}(A)$ is the smallest subspace containing A and $\operatorname{CLin}(A)$ is its closure, the smallest closed subspace containing A.

H.9. *From now on we work in inner-product spaces.*

H.10. The *orthogonality.* $x \perp y$ if $\langle x, y\rangle = 0$. An *orthogonal sequence* has $\langle e_n, e_m\rangle = 0$ for $n \neq m$. If all the vectors have norm 1 it is an *orthonormal sequence (o.n.s.)*, e.g. $e_n = (0, \ldots, 0, 1, 0, 0, \ldots) \in \ell_2$ and $e_n(t) = (1/\sqrt{2\pi})e^{int}$ in $L_2(-\pi, \pi)$.

H.11. *Pythagoras's theorem*: if $x \perp y$ then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

H.12. The *best approximation* to x by a linear combination $\sum_{k=1}^{n} \lambda_k e_k$ is $\sum_{k=1}^{n}\langle x, e_k\rangle e_k$ if the e_k are orthonormal. Note that $\langle x, e_k\rangle$ is the Fourier coefficient of x w.r.t. e_k.

H.13. *Bessel's inequality.* $\|x\|^2 \geq \sum_{k=1}^{n} |\langle x, e_k\rangle|^2$ if $e_1, \ldots, e_n$ is an o.n.s.

H.14. *Riesz–Fischer theorem.* For an o.n.s. (e_n) in a Hilbert space, $\sum \lambda_n e_n$ converges if and only if $\sum |\lambda_n|^2 < \infty$; then $\| \sum \lambda_n e_n\|^2 = \sum |\lambda_n|^2$.

H.15. A *complete o.n.s.* or *orthonormal basis (o.n.b.)* is an o.n.s. (e_n) such that if $\langle y, e_n\rangle = 0$ for all n then $y = 0$. In that case every vector is of the form $\sum \lambda_n e_n$ as in the R-F theorem. Equivalently: the closed linear span of the (e_n) is the whole space.

H.16. *Gram–Schmidt orthonormalization process.* Start with $x_1, x_2, \ldots$ linearly independent. Construct $e_1, e_2, \ldots$ an o.n.s. by inductively setting $y_{n+1} = x_{n+1} - \sum_{k=1}^{n}\langle x_{n+1}, e_k\rangle e_k$ and then normalizing $e_{n+1} = y_{n+1}/\|y_{n+1}\|$.

H.17. On *orthogonal complements.* $M^\perp$ is the set of all vectors orthogonal to everything in M. If M is a closed linear subspace of a Hilbert space H then $H = M \oplus M^\perp$. There is also a linear map, P_M the projection from H onto M with kernel $M^\perp$.

H.18. *Fourier series.* Work in $L_2(-\pi, \pi)$ with o.n.s. $e_n(t) = (1/\sqrt{2\pi})e^{int}$. Let $CP(-\pi, \pi)$ be the continuous periodic functions, which are dense in L_2. For $f \in CP(-\pi, \pi)$ write $f_m = \sum_{n=-m}^{m}\langle f, e_n\rangle e_n$, $m \geq 0$. We wish to show that $\|f_m - f\|_2 \to 0$, i.e., that (e_n) is an o.n.b.

H.19. The *Fejér kernel.* For $f \in CP(-\pi, \pi)$ write $F_m = (f_0 + \ldots + f_m)/(m + 1)$. Then $F_m(x) = (1/2\pi) \int_{-\pi}^{\pi} f(t)K_m(x - t)\,dt$ where $K_m(t) = (1/(m + 1)) \sum_{k=0}^{m} \sum_{n=-k}^{k} e^{int}$ is the Fejér kernel. Also $K_m(t) = (1/(m + 1))[\sin^2(m + 1)t/2]/[\sin^2 t/2]$.

H.20. *Fejér's theorem.* If $f \in CP(-\pi, \pi)$ then its Fejér sums tend uniformly to f on $[-\pi, \pi]$ and hence in L_2 norm also. Hence $\operatorname{CLin}((e_n)) \supseteq CP(-\pi, \pi)$ so must be all of $L_2(-\pi, \pi)$. Thus (e_n) is an o.n.b.

H.21. *Corollary.* If $f \in L_2(-\pi, \pi)$ then $f(t) = \sum c_n e^{int}$ with convergence in L_2, where $c_n = (1/2\pi) \int_{-\pi}^{\pi} f(t)e^{-int}\,dt$.

H.22. *Parseval's formula.* If f, $g \in L_2(-\pi, \pi)$ have Fourier series $\sum c_n e^{int}$ and $\sum d_n e^{int}$ then $(1/2\pi)\langle f, g\rangle = \sum c_n \bar{d}_n$.

H.23. *Weierstrass approximation theorem.* The polynomials are dense in $C[a, b]$ for any $a < b$ (in the supremum norm).

Appendix I. Some useful results and formulae (2)

I.1. On *dual spaces.* A *linear functional* on a vector space X is a linear mapping $\alpha : X \to \mathbb{C}$ (or to $\mathbb{R}$ in the real case), i.e., $\alpha(ax + by) = a\alpha(x) + b\alpha(y)$. When X is a normed space, α is continuous if and only if it is *bounded*, i.e., $\sup\{|\alpha(x)| : \|x\| \leq 1\} < \infty$. Then we define $\|\alpha\|$ to be this sup, and it is a norm on the space X^* of bounded linear functionals, making X^* into a Banach space.

I.2. *Riesz–Fréchet theorem.* If $\alpha : H \to \mathbb{C}$ is a bounded linear functional on a Hilbert space H, then there is a unique $y \in H$ such that $\alpha(x) = \langle x, y\rangle$ for all $x \in H$; also $\|\alpha\| = \|y\|$.

I.3. On *linear operator.* These are linear mappings $T : X \to Y$, between normed spaces. Defining $\|T\| = \sup\{\|T(x)\| : \|x\| \le 1\}$, finite, makes the bounded (i.e., continuous) operators into a normed space, $B(X,Y)$. When Y is complete, so is $B(X,Y)$. We get $\|Tx\| \le \|T\|\,\|x\|$, and, when we can compose operators, $\|ST\| \le \|S\|\,\|T\|$. Write $B(X)$ for $B(X,X)$, and for $T \in B(X)$, $\|T^n\| \le \|T\|^n$. *Inverse* $S = T^{-1}$ when $ST = TS = I$.

I.4. On *adjoints.* $T \in B(H,K)$ determines $T^* \in B(K,H)$ such that $\langle Th, k\rangle_K = \langle h, T^*k\rangle_H$ for all $h \in H$, $k \in K$. Also $\|T^*\| = \|T\|$ and $T^{**} = T$.

I.5. On *unitary operator.* Those $U \in B(H)$ for which $UU^* = U^*U = I$. Equivalently, U is surjective and an isometry (and hence preserves the inner product).

Hermitian operator or *self-adjoint operator.* Those $T \in B(H)$ such that $T = T^*$.

On *normal operator.* Those $T \in B(H)$ such that $TT^* = T^*T$ (so including Hermitian and unitary operators).

I.6. On *spectrum.* $\sigma(T) = \{\lambda \in \mathbb{C} : (T - \lambda I) \text{ is not invertible in } B(X)\}$. Includes all *eigenvalues* λ where $Tx = \lambda x$ for some $x \ne 0$, and often other things as well. On *spectral radius:* $r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$. Properties: $\sigma(T)$ is closed, bounded and nonempty. Proof: based on the fact that $(I - A)$ is invertible for $\|A\| < 1$. This implies that $r(T) \le \|T\|$.

I.7. The *spectral radius formula.* $r(T) = \inf_{n\ge1} \|T^n\|^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n}$.

Note that $\sigma(T^n) = \{\lambda^n : \lambda \in \sigma(T)\}$ and $\sigma(T^*) = \{\overline{\lambda} : \lambda \in \sigma(T)\}$. The spectrum of a unitary operator is contained in $\{|z| = 1\}$, and the spectrum of a self-adjoint operator is real (proof by *Cayley transform*: $U = (T - iI)(T + iI)^{-1}$ is unitary).

I.8. On *finite rank operator.* $T \in F(X,Y)$ if $\operatorname{Im} T$ is finite-dimensional.

On *compact operator.* $T \in K(X,Y)$ if: whenever (x_n) is bounded, then (Tx_n) has a convergent subsequence. Now $F(X,Y) \subseteq K(X,Y)$ since bounded sequences in a finite-dimensional space have convergent subsequences (because when Z is f.d., Z is isomorphic to ℓ_2^n, i.e., $\exists S : \ell_2^n \to Z$ with S, S^{-1} bounded). Also limits of compact operators are compact, which shows that a diagonal operator $Tx = \sum \lambda_n \langle x, e_n\rangle e_n$ is compact iff $\lambda_n \to 0$.

I.9. *Hilbert–Schmidt operators.* T is H–S when $\sum \|Te_n\|^2 < \infty$ for some o.n.b. (e_n). All such operators are compact—write them as a limit of finite rank operators T_k with $T_k \sum_{n=1}^\infty a_n e_n = \sum_{n=1}^k a_n (Te_n)$. This class includes integral operators $T : L_2(a,b) \to L_2(a,b)$ of the form

$$(Tf)(x) = \int_a^b K(x,y) f(y)\, dy,$$

where K is continuous on $[a,b] \times [a,b]$.

I.10. On *spectral properties of normal operators.* If T is normal, then (i) $\ker T = \ker T^*$, so $Tx = \lambda x \implies T^*x = \overline{\lambda} x$; (ii) eigenvectors corresponding to distinct eigenvalues are orthogonal; (iii) $\|T\| = r(T)$.

If $T \in B(H)$ is compact normal, then its set of eigenvalues is either finite or a sequence tending to zero. The eigenspaces are finite-dimensional, except possibly for $\lambda = 0$. All nonzero points of the spectrum are eigenvalues.

I.11. On *spectral theorem for compact normal operators.* There is an orthonormal sequence (e_k) of eigenvectors of T, and eigenvalues (λ_k), such that $Tx = \sum_k \lambda_k \langle x, e_k\rangle e_k$. If (λ_k) is an infinite sequence, then it tends to 0. All operators of the above form are compact and normal.

Corollary. In the spectral theorem we can have the same formula with an orthonormal *basis*, adding in vectors from $\ker T$.

I.12. On *general compact operators.* We can write $Tx = \sum \mu_k \langle x, e_k \rangle f_k$, where (e_k) and (f_k) are orthonormal sequences and (μ_k) is either a finite sequence or an infinite sequence tending to 0. *Hence $T \in B(H)$ is compact if and only if it is the norm limit of a sequence of finite-rank operators.*

I.13. On *integral equations.* Fredholm equations on $L_2(a,b)$ are $T\phi = f$ or $\phi - \lambda T\phi = f$, where $(T\phi)(x) = \int_a^b K(x,y)\phi(y)\,dy$. Volterra equations similar, except that T is now defined by $(T\phi)(x) = \int_a^x K(x,y)\phi(y)\,dy$.

I.14. *Neumann series.* $(I - \lambda T)^{-1} = 1 + \lambda T + \lambda^2 T^2 + \ldots$, for $\|\lambda T\| < 1$.

On *separable kernels.* $K(x,y) = \sum_{j=1}^n g_j(x) h_j(y)$. The image of T (and hence its eigenvectors for $\lambda \neq 0$) lies in the space spanned by $g_1, \ldots, g_n$.

I.15. *Hilbert–Schmidt theory.* Suppose that $K \in C([a,b] \times [a,b])$ and $K(y,x) = \overline{K(x,y)}$. Then (in the Fredholm case) T is a self-adjoint Hilbert–Schmidt operator and eigenvectors corresponding to nonzero eigenvalues are continuous functions. If $\lambda \neq 0$ and $1/\lambda \notin \sigma(T)$, the the solution of $\phi - \lambda T\phi = f$ is

$$\phi = \sum_{k=1}^{\infty} \frac{\langle f, v_k \rangle}{1 - \lambda \lambda_k} v_k.$$

I.16. *Fredholm alternative.* Let T be compact and normal and $\lambda \neq 0$. Consider the equations (i) $\phi - \lambda T\phi = 0$ and (ii) $\phi - \lambda T\phi = f$. Then EITHER (A) The only solution of (i) is $\phi = 0$ and (ii) has a unique solution for all f OR (B) (i) has nonzero solutions ϕ and (ii) can be solved if and only if f is orthogonal to every solution of (i).

Appendix J. Examinable material

All definitions and results are examinable, and you may be asked to do exercises and examples similar to those on the problem sheets.

The following (long) **proofs** are **NOT** examinable (the rest are):

9.3	Orthogonal complements
11.2–11.4	Fejér kernel
13.3	The heat equation
13.4	Spectral decomposition of audio samples
15.1	Riesz–Fréchet
21.2	Spectral radius
23.6	Finite-dimensional normed spaces
23.9	Limits of compact operators
25.1.iii	$\|T\| = r(T)$ if T is normal
25.6	Eigenvalues of compact normal operators
26.1	Implication $\Longrightarrow$ of the spectral theorem
26.4	Singular Value Decomposition
	Any supplementary material beyond this point

If in doubt, you should check with me.

Part 13. Supplementary Sections

Appendix K. Reminder from Complex Analysis

The analytic function theory is the most powerful tool in the operator theory. Here we briefly recall few facts of complex analysis used in this course. Use any decent textbook on complex variables for a concise exposition. The only difference with our version that we consider function $f(z)$ of a complex variable z t*aking value in an arbitrary normed space* V over the field $\mathbb{C}$. By the direct inspection we could check that all standard proofs of the listed results work as well in this more general case.

Definition K.1. A function $f(z)$ of a complex variable z taking value in a normed vector space V is called *differentiable* at a point z_0 if the following limit (called *derivative* of $f(z)$ at z_0) exists:

$$f'(z_0) = \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}. \tag{K.1}$$

Definition K.2. A function $f(z)$ is called *holomorphic* (or *analytic*) in an open set $\Omega \subset \mathbb{C}$ it is differentiable at any point of Ω.

Theorem K.3 (Laurent Series)**.** *Let a function $f(z)$ be analytical in the annulus $r < z < R$ for some real $r < R$, then it could be uniquely represented by the Laurent series:*

$$f(z) = \sum_{k=-\infty}^{\infty} c_k z^k, \qquad \textit{for some } c_k \in V. \tag{K.2}$$

Theorem K.4 (Cauchy–Hadamard)**.** *The radii r' and R', $(r' < R')$ of convergence of the Laurent series* (K.2) *are given by*

$$r' = \liminf_{n\to\infty} \|c_n\|^{1/n} \qquad \textit{and} \qquad \frac{1}{R'} = \limsup_{n\to\infty} \|c_n\|^{1/n}. \tag{K.3}$$

References

[1] Nicholas Young. *An introduction to Hilbert space*. Cambridge University Press, Cambridge, 1988. MR **90e:**46001. 1

[2] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987. MR **88k:**00002. 1

[3] Béla Bollobás. *Linear analysis*. Cambridge University Press, Cambridge, second edition, 1999. An introductory course. MR **2000g:**46001. 1

[4] Erwin Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons Inc., New York, 1989. MR **90m:**46003. 1

[5] Alexander A. Kirillov and Alexei D. Gvishiani. *Theorems and Problems in Functional Analysis*. Problem Books in Mathematics. Springer-Verlag, New York, 1982. 1

[6] Michael Reed and Barry Simon. *Functional Analysis*, volume 1 of *Methods of Modern Mathematical Physics*. Academic Press, Orlando, second edition, 1980. 1

Index