Introduction to Methods of Hilbert Spaces

J. Banasiak

University of KwaZulu-Natal, Durban
Technical University of Łódź

Contents

Chapter 1

Basic properties

1.1 Definitions

Hilbert spaces are a sub-class of Banach spaces with norm defined by a bilinear (or rather, sesquilinear) functional which induces a number of useful geometric properties of the space. Let us start with a more general concept of a *unitary space.*

Definition 1.1.1. A non-empty set H is called a unitary space if it is a complex linear space equipped with a complex-valued function $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ satisfying the following properties:

1. $\forall_{x \in H} \langle x, x \rangle \geq 0$ with $\langle x, x \rangle = 0$ if and only if $x = 0$;

2. $\forall_{x,y,z \in H} \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;

3. $\forall_{x,y \in H, \lambda \in \mathbb{C}} \langle \lambda x, y \rangle = \lambda \langle x, y \rangle$;

4. $\forall_{x,y \in H} \langle x, y \rangle = \overline{\langle y, x \rangle}$;

The functional $\langle \cdot, \cdot \rangle$ is called the *scalar product* or the *inner product.*

We collect a few basic properties of the inner product. Since $0x = 0$ for any $x \in H$, properties 3. and 4. imply

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

for any $x, y \in H$. Properties 2. and 3. can be summarized by saying that for any $y \in H$ the mapping $x \mapsto \langle x, y \rangle$ is a linear functional on H.

We note that it is possible to consider real unitary spaces. In this case H is a real vector space, $\langle\cdot,\cdot\rangle : H \times H \mapsto \mathbb{R}$ so that 4. becomes commutativity. A real inner product is a linear functional with respect to both variables.

Since, by 1., $\langle x, x\rangle \geq 0$, we can introduce a functional $\|\cdot\| : H \to \mathbb{R}_+$ by

$$\|x\| = \sqrt{\langle x, x\rangle}. \tag{1.1.1}$$

So far there is no real justification for using the norm symbol $\|\cdot\|$ here. We see, however, that by 1. $\|x\| = 0$ if and only if $x = 0$ and that $\|\lambda x\| = |\lambda|\|x\|$ for any $x \in H$ and $\lambda \in \mathbb{C}$. Hence, to prove that it is a norm on H, we only have to prove the triangle inequality. This follows from the Schwarz lemma.

Lemma 1.1.2. *For any $x, y \in H$ we have*

$$|\langle x, y\rangle| \leq \|x\|\|y\|, \tag{1.1.2}$$

and the equality occurs if and only if x and y are colinear.

Proof. If $y = 0$, then (1.1.2) is obvious. Suppose then that $x, y \neq 0$. For any $\lambda \in \mathbb{C}$ we have by 1.

$$\begin{aligned} 0 &\leq \langle x + \lambda y, x + \lambda y\rangle \\ &= \|x\|^2 + |\lambda|^2\|y\|^2 + \lambda\langle x, y\rangle + \bar{\lambda}\langle y, x\rangle. \end{aligned}$$

Now, take $\lambda = -\langle x, y\rangle/\|y\|^2$, then

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

so that $|\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2$.

If in the inequality we have equality sign, then reversing the steps we find $\|x + \lambda y\| = 0$ which means that $x = -\langle y$, where we used comments below (1.1.1), provided $y \neq 0$. If $y = 0$, then the elements are obviously dependent.

If $y = \mu x$ for some $y \neq 0, \mu \neq 0$, then

$$|\langle x, y\rangle| = |\langle x, \mu x\rangle| = |\mu|\|x\|\|x\| = \|x\|\|\mu x\| = \|x\|\|y\|.$$

If either element equals 0, then the equality is obvious. ∎

Proposition 1.1.3. *An inner product space H equipped with the norm (1.1.1) is a normed space.*

Proof. In view of the comments above, it suffice to prove that (1.1.1) satisfies the triangle inequality. In this respect, we have for arbitrary $x, y \in H$

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

where we used the Schwarz inequality to pass to the last line. Hence

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

and the proposition is proved. ∎

Thus the use of $\|\cdot\|$ is fully justified and we can use all terminology and results of the normed space theory to unitary spaces. In particular, any unitary space becomes a metric/topological space with metric defined by

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

and topology generated by the basis of open balls $B(y, r) = \{x \in H;\ \|x - y\| < r\}$, $y \in H, r > 0$. We have

Proposition 1.1.4. *The scalar product is a continuous functional over* $H \times H$.

Proof. Since $\langle x, y\rangle - \langle x_0, y_0\rangle = \langle x - x_0, y\rangle + \langle x_0, y - y_0\rangle + \langle x - x_0, y - y_0\rangle$,

$$|\langle x, y\rangle - \langle x_0, y_0\rangle| \leq \|x - x_0\|\|y\| + \|x_0\|\|y - y_0\| + \|x - x_0\|\|y - y_0\|,$$

and the statement follows. ∎

Definition 1.1.5. We say that a unitary space is a Hilbert space if it is complete with respect to the norm (1.1.1).

We know that every metric/normed space admits a completion; that is, it is isometric with a dense subspace of a complete metric space/Banach space. The same is true for unitary spaces.

Theorem 1.1.6. *The completion of a unitary space is a Hilbert space. That is, any unitary space H is isometric with a dense subspace of a Hilbert space.*

Proof. Let us recall that the completion of a unitary space understood as a normed space is the space of all Cauchy sequences $(x_n)_{n\in\mathbb{N}}$, denoted H_1, modulo the equivalence relation $\sim$ defined as $(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$ if and only if $\lim_{n\to\infty} \|x_n - y_n\| = 0$. If $[x] \in \mathcal{H} := H_1/\sim$ we have

$$\|[x]\| = \lim_{n\to\infty} \|x_n\|,$$

where $(x_n)_{n\in\mathbb{N}}$ is an arbitrary element of $[x]$. We shall see that this construction extends the inner product on $\mathcal{H}$ which generates the above norm. Natural definition is

$$\langle [x], [y] \rangle = \lim_{n\to\infty} \langle x_n, y_n \rangle. \tag{1.1.3}$$

This limit exists. Indeed,

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x_m, y_m\rangle| &\leq |\langle x_n, y_n - y_m\rangle| + |\langle x_n - x_m, y_m\rangle| \\ &\leq \|x_n\|\|y_n - y_m\| + \|x_n - x_m\|\|y_m\| \end{aligned}$$

and the last line tends to 0 since $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are bounded being Cauchy sequences. Thus $(\langle x_n, y_n\rangle)_{n\in\mathbb{N}}$ is a Cauchy sequence in (complete) $\mathbb{C}$ and the sequence converges. The limit does not depend on the representatives of $[x]$ and $[y]$. Indeed, taking other representatives $(\xi_n)_{n\in\mathbb{N}} \in [x]$ and $(\eta_n)_{n\in\mathbb{N}} \in [y]$, we have

$$\langle x_n, y_n\rangle = \langle x_n - \xi_n + \xi_n, y_n - \eta_n + \eta_n\rangle = \langle x_n - \xi_n, y_n\rangle + \langle \xi_n, y_n - \eta_n\rangle + \langle \xi_n, \eta_n\rangle$$

with

$$|\langle x_n - \xi_n, y_n\rangle + \langle \xi_n, y_n - \eta_n\rangle| \leq \|y_n\|\|x_n - \xi_n\| + \|\xi_n\|\|y_n - \eta_n\| \to 0.$$

Since the scalar product on H is continuous, all algebraic properties of the definition of the scalar product carry over to the scalar product on $\mathcal{H}$, defined by (1.1.3). We shall check axiom 1. If $[x] = 0$, then any $(x_n)_{n\in\mathbb{N}} \in [x]$ converges to 0 (as $(0, 0, \ldots, 0, \ldots) \in [x]$). Thus, $\langle [x], [x]\rangle = 0$. Conversely, if

$$0 = \langle [x], [x]\rangle = \lim_{n\to\infty} \langle x_n, x_n\rangle = \lim_{n\to\infty} \|x_n\|^2$$

which means that $(x_n)_{n\in\mathbb{N}} \in [x]$ is a null sequence and so $[x] = 0$.

Finally, for $(x_n)_{n\in\mathbb{N}} \in [x]$,

$$\|[x]\| = \lim_{n\to\infty} \|x_n\| = \lim_{n\to\infty} \sqrt{\langle x_n, x_n\rangle} = \sqrt{\langle [x], [x]\rangle}.$$

∎

1.2 Geometry of unitary spaces

A norm in a unitary space has special properties which typically do not occur in other normed spaces.

Proposition 1.2.1. *The unit ball in a unitary space H is strictly convex.*

Proof. Let us take $x, y \in H$ with $\|x\| = \|y\| = 1$ and consider the segment joining x and z: $S = \{z \in H;\ z = \alpha x + \beta y, 0 \leq \alpha, \beta, \alpha + \beta = 1\}$. Then

$$\|z\|^2 = \alpha^2 + \beta^2 + 2\alpha\beta\Re\langle x, y\rangle < (\alpha + \beta)^2 = 1$$

with equality possible only when x and y are colinear (or α or β are 0). ■

While there are nonunitary normed spaces with strictly convex unit balls, the next property sets unitary spaces apart from all other normed spaces.

Theorem 1.2.2. *Let H be a complex normed space with norm $\|\cdot\|$. Then H is a unitary space if and only if for any $x, y \in H$ the parallelogram law holds:*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{1.2.1}$$

Proof. Assume $\|\cdot\|$ is given by (1.1.1). Then

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

The proof in the opposite direction is much more complicated. Let us start from real spaces. Then the inner product can be expressed by

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

so, if we suspect that a given norm $\|\cdot\|$ is generated by an unknown scalar product, then it must be given by the above formula. The problem is to check that

$$p(x, y) = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$$

satisfies the axiom of a (real) inner product, provided the norm obeys the parallelogram law. First we notice that

$$p(x, x) = \|x\|^2 \geq 0$$

so that the norm satisfies axiom 1. It is also clear that $p(x, y) = p(y, x)$. To prove homogeneity and additivity, we observe that the parallelogram law gives

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

as well as

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

Therefore

$$\begin{aligned} & p(x + w, y) + p(x - w, y) \\ & \quad = \frac{1}{4}\left(\|x + y + w\|^2 - \|x - y + w\|^2 - \|x - y - w\|^2 + \|x + y - w\|^2\right) \\ & \quad = \frac{1}{2}\left(\|x + y\|^2 - \|x - y\|^2\right) = 2p(x, y) \end{aligned}$$

If we set $x = w$, then we get

$$p(2x, y) = 2p(x, y)$$

and taking $x = \frac{1}{2}(x_1 + x_2)$ and $w = \frac{1}{2}(x_1 - x_2)$ we find

$$\begin{aligned} p(x + w, y) + p(x - w, y) &= p(x_1, y) + p(x_2, y) = 2p\left(\frac{1}{2}(x_1 + x_2), y\right) \\ &= p(x_1 + x_2, y). \end{aligned}$$

Moreover, the last identity shows, by induction, that

$$p(nx, y) = np(x, y)$$

for any $x, y \in H$ and natural n and

$$p(x, y) = p\left(m\frac{x}{m}, y\right) = mp\left(\frac{x}{m}, y\right)$$

which, combined, gives

$$p\left(\frac{n}{m}x, y\right) = \frac{n}{m}p(x, y)$$

for any natural n, m, that is,

$$rp(x, y) = p(rx, y)$$

for any positive rational number. Since rationals are dense in real numbers and p is a continuous function we get

$$\lambda p(x, y) = p(\lambda x, y)$$

for any $\lambda \geq 0$. If $\lambda < 0$ then we write

$$\begin{aligned}\lambda p(x,y) - p(\lambda x, y) &= \lambda p(x,y) - |\lambda| p(-x,y) = \lambda p(x,y) + \lambda p(-x,y) \\ &= \lambda(p(x,y) + p(-x,y)) = \lambda p(0,y) = 0\end{aligned}$$

and p satisfies all axioms of a real inner product.

Let us consider p on a complex inner space H. Then

$$p(ix,y) = \frac{1}{4}\left(\|ix+y\|^2 - \|ix-y\|^2\right) = \frac{1}{4}\left(\|x-iy\|^2 - \|x+iy\|^2\right) = -p(x,iy).$$

Now, let us reflect how one can build a complex scalar product if one has a real one. We need to incorporate both x and ix so let us try

$$\langle x,y\rangle = Ap(x,y) + Bp(ix,y)$$

and find A and B for which $\langle(\alpha+i\beta)x,y\rangle = (\alpha+i\beta)\langle x,y\rangle$. We should have

$$\begin{aligned}\langle(\alpha+i\beta)x,y\rangle &= A\alpha p(x,y) + A\beta p(ix,y) + B\alpha p(ix,y) + B\beta p(-x,y) \\ &= A\alpha p(x,y) + A\beta p(ix,y) + B\alpha p(ix,y) - B\beta p(x,y) \\ &= (\alpha+i\beta)(Ap(x,y) + Bp(ix,y))\end{aligned}$$

which yields $A\alpha - B\beta = (\alpha+i\beta)A$ and $A\beta + B\alpha = (\alpha+i\beta)B$ and hence $A = iB$. Taking $A = 1$ we arrive at

$$\langle x,y\rangle = p(x,y) + ip(ix,y) \tag{1.2.2}$$

which is homogeneous in the first variable, additive (as a linear combination of additive functionals). Furthermore, $p(ix,x) = -p(x,ix) = -p(ix,x)$ so that $p(ix,x) = 0$ hence

$$\langle x,x\rangle = p(x,x) + ip(ix,x) = p(x,x) = \|x\|^2$$

and

$$\langle x,y\rangle = p(x,y) + ip(ix,y) = p(y,x) - ip(x,iy) = p(y,x) - ip(iy,x) = \overline{\langle y,x\rangle}.$$

This shows that $\langle x,y\rangle$ defined in (1.2.2) is an inner product generating the norm. ■

It is worthwhile to write explicitly the derived expressions for the inner product in terms of the norm.

Corollary 1.2.3. *If $\|\cdot\|$ is the norm in the unitary space H, given by (1.1.1), then for any $x, y \in H$*

$$\begin{aligned} \Re\langle x,y\rangle &= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \\ \Im\langle x,y\rangle &= -\frac{1}{4}\left(\|ix+y\|^2 - \|ix-y\|^2\right). \end{aligned}$$

Hence, if H is a real space, then

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) \tag{1.2.3}$$

and

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right) - i\frac{1}{4}\left(\|ix+y\|^2 - \|ix-y\|^2\right). \tag{1.2.4}$$

The inner product allows to introduce in a unitary space the concept of orthogonality.

Definition 1.2.4. Let x, y be two points in a unitary space H. If $\langle x,y\rangle = 0$ then we say that x and y are orthogonal and write $x \perp y$. If $M \subset H$ and x is orthogonal to all elements of M, then we say that x is orthogonal to M and write $x \perp M$. If $M, N \subset H$ and for every $x \in N$ we have $x \perp M$, we say that N and M are orthogonal and write $M \perp N$ (we then also have $N \perp M$. By $M^\perp$ we denote the set of all elements that are orthogonal to M and call it the orthogonal complement of M.

We observe that $M^\perp$ is a closed linear subspace of H as well as that if $N \perp M$ then $N \subset M^\perp$. [!]

Proposition 1.2.5. *(Pythagoras theorem) For $x, y \in H$, if $x \perp y$, then*

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2. \tag{1.2.5}$$

If H is a real space, then (1.2.5) implies $x \perp y$.

Proof. We have

$$\|x-y\|^2 = \langle x-y, x-y\rangle = \|x\|^2 + \|y\|^2 - 2\Re\langle x,y\rangle$$

so clearly $x \perp y$ implies (1.2.5). If H is a real space, then $\Re\langle x,y\rangle$ is to be replaced by $\langle x,y\rangle$ and the argument can be reversed. ■

1.3 Basic examples

Let us discuss several examples of normed spaces which are, or which are not, unitary or Hilbert spaces.

1.3.1 The spaces $\mathbb{R}^n$ and $\mathbb{C}^n$.

The space $\mathbb{R}^n$ and, respectively, $\mathbb{C}^n$, equipped with the inner products

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

and, respectively,

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i},$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ are in $\mathbb{R}^n$ (respectively $\mathbb{C}^n$) are inner product spaces. The fact that they are Hilbert spaces follows from the fact that they are finite dimensional, and thus complete.

1.3.2 The space l_2.

The simplest infinite-dimensional extension of $\mathbb{C}^n$ is the complex space l_2 consisting of infinite sequences $x = (x_n)_{n\in\mathbb{N}}$ for which the

$$\|x\| = \sum_{n=0}^{\infty} |x_n|^2 < +\infty. \tag{1.3.1}$$

This space becomes a unitary space if equipped with the inner product

$$\langle x, y\rangle = \sum_{n=0}^{\infty} x_n \overline{y_n}, \tag{1.3.2}$$

where $x = (x_n)_{n\in\mathbb{N}}, y = (y_n)_{n\in\mathbb{N}}$. The fact that it is a well defined inner product on l_2 follows from the comparison criterion of series convergence, since $|x_n y_n| \leq \frac{1}{2}(x_n^2 + y_n^2)$ for $n = 1, 2, \ldots$. Once this is established, all axioms of the inner product follow.

The space l_2 is complete. Indeed, consider a Cauchy sequence $(x^{(k)})_{k\in\mathbb{N}}$ with $x^{(k)} = (x_n^{(k)})_{n\in\mathbb{N}}$. For any ϵ there is N such that

$$\|x^{(k)} - x^{(m)}\|^2 = \sum_{n=0}^{\infty} |x_n^{(k)} - x_n^{(m)}|^2 \leq \epsilon^2$$

for $k, m \geq N$. This means that for any n we have $|x_n^{(k)} - x_n^{(m)}| \leq \epsilon$ as long as $k, m \geq N$. Since ϵ is arbitrary, this means that each numerical sequence $(x_n^{(k)})_{k\in\mathbb{N}}$, $n \in \mathbb{N}$, is a Cauchy sequence with limit, say, x_n. Since for any finite r we have

$$\sum_{n=0}^{r} |x_n^{(k)} - x_n^{(m)}|^2 \leq \epsilon^2,$$

we can pass to the limit with $k \to \infty$ so that

$$\sum_{n=0}^{r} |x_n - x_n^{(m)}|^2 \leq \epsilon^2.$$

The above is valid for any finite r hence

$$\sum_{n=0}^{\infty} |x_n - x_n^{(m)}|^2 \leq \epsilon^2,$$

as the terms of the series are non-negative. This shows that $(x_n)_{n\in\mathbb{N}} - (x_n^{(k)})_{n\in\mathbb{N}} \in l_2$ and, by $(x_n)_{n\in\mathbb{N}} = (x_n)_{n\in\mathbb{N}} - (x_n^{(k)})_{n\in\mathbb{N}} + (x_n^{(k)})_{n\in\mathbb{N}}$, $(x_n)_{n\in\mathbb{N}} \in l_2$. Further, the last equality shows that $(x_n^{(k)})_{n\in\mathbb{N}}$ converges to $(x_n)_{n\in\mathbb{N}}$ in l_2 and thus l_2 is complete.

We note that l_2 is also separable as the set $\{(\delta_{ij})_{i\in\mathbb{N}}\}_{j\in\mathbb{N}}$ is linearly dense in l_2.

The space l_2 is the prototype of a Hilbert space. It was introduced and investigated by D. Hilbert in 1912 in his work on integral equations. However, the axiomatic definition was given only in 1927 by J. von Neumann in the context of foundations of quantum mechanics. It is also a universal separable Hilbert space in the sense that any other separable Hilbert space is isometric with l_2. The isometry is achieved through the Fourier series expansion which will be discussed at a later stage.

1.3.3 The space $L_2(\Omega)$

. In general, we can consider a measure space Ω with positive measure μ but for simplicity leat us focus on $\Omega = [0, 1]$ with Lebesgue measure $d\mu = dt$.

Let us begin with the linear space $\mathcal{C}$ of continuous functions on $[0, 1]$. We know that equipped with the norm

$$\|x(t)\| = \sup_{t\in[0,1]} |x(t)|, \qquad , x \in \mathcal{C}$$

the space $\mathcal{C}$ becomes a Banach space, denoted by $C([0,1])$. Let us first check whether $C([0,1])$ is a Hilbert space. Consider $x(t) = 1$ and $y(t) = t$. Then

$$\begin{aligned} \|x\|^2 &= 1, \quad \|y\|^2 = 1, \\ \|x+y\|^2 &= 2^2, \quad \|x-y\|^2 = 1 \end{aligned}$$

and

$$\|x+y\|^2 + \|x-y\|^2 = 5 \neq 4 = 2(\|x\|^2 + \|y\|^2).$$

Thus, the parallelogram law is not satisfied and the norm cannot be generated by a scalar product.

Let us consider a scalar product on $\mathcal{C}$ which mimics the scalar product on l_2: Define

$$\langle x, y\rangle = \int_0^1 x(t)y(t)dt \tag{1.3.3}$$

in the real case and

$$\langle x, y\rangle = \int_0^1 x(t)\overline{y(t)}dt \tag{1.3.4}$$

for complex valued functions. Note that for continuous functions we can use the Riemann integral as well as the Lebesgue integral.

Since the product of continuous functions is a continuous function and the interval $[0,1]$ is bounded, both inner products are well defined on $\mathcal{C} \times \mathcal{C}$. Furthermore, since integral of a continuous nonnegative function is zero if and only if the function is identically zero; that is

$$\int_0^1 |x(t)|^2 dt = 0 \quad \text{if and only if} \quad \forall_{t\in[0,1]} x(t) = 0,$$

we see that the first axiom of the inner product is satisfied. Linearity is clear from properties of integration. For the skew-symmetry, we have

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

Denote the norm defined by this inner product by

$$\|x\|_2 = \sqrt{\int_0^1 |x(t)|^2 dt}. \tag{1.3.5}$$

Let us consider functions shown in Fig. 1.1:

$$x_m(t) = \begin{cases} 0 & \text{for } 0 \le t \le \frac{1}{2}, \\ \frac{m}{t} - \frac{m}{2} & \text{for } \frac{1}{2} < t \le \frac{1}{2} + \frac{1}{m}, \\ 1 & \text{for } \frac{1}{2} + \frac{1}{m} < t \le 1 \end{cases}$$

Let $n > m$. We have

$$\|x_n - x_m\|_2^2 = \int\limits_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{n}} \left((n-m)t + \frac{m-n}{2}\right)^2 dt + \int\limits_{\frac{1}{2}+\frac{1}{n}}^{\frac{1}{2}+\frac{1}{m}} \left(1 - mt + \frac{m}{2}\right)^2 dt \le \frac{4}{m}.$$

Hence $\|x_n - x_m\|_2 \to 0$ as $m \to \infty$ (and thus $n \to \infty$) and $(x_n)_{n\in\mathbb{N}}$ is a

Figure 1.1: Cauchy sequence of continuous functions

Cauchy sequence. Assume that there is a continuous function x such that $\|x - x_n\|_2 \to 0$. Then, for any ϵ there is N such that for any $n \ge N$

$$\|x - x_m\|_2^2 = \int\limits_0^{\frac{1}{2}} |x(t)|^2 dt + \int\limits_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{m}} \left(x(t) - mt + \frac{m}{2}\right)^2 dt + \int\limits_{\frac{1}{2}+\frac{1}{m}}^{1} (1 - x(t))^2 \, dt \le \epsilon.$$

Since each term is nonnegative, each must be smaller then ϵ. Hence, $x(t) = 0$ on $\left[0, \frac{1}{2}\right]$ and on each interval $\left[\frac{1}{2} + \frac{1}{m}, 1\right]$, and hence on $\left(\frac{1}{2}, 1\right]$, $x(t) = 1$. This, however, contradicts the definition of continuous function.

Thus, $\mathcal{C}$ equipped with the inner product (1.3.4) is an incomplete unitary space. By Theorem 1.1.6 it can, however, be completed. By construction , the elements of the completion are classes of equivalence $[x]$ of sequences $(x_n)_{n\in\mathbb{N}}$ of continuous functions which are Cauchy with respect to the norm (1.3.5) and such that $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \in [x]$ if and only if $\lim_{n\to\infty} \|x_n - y_n\|_2 = 0$. Let us provide a more friendly description of this completion. Let us take any representative $(x_n)_{n\in\mathbb{N}} \in [x]$. It is a Cauchy sequence and hence there exists a subsequence $(x_{n_i})_{i\in\mathbb{N}}$ such that

$$\|x_{n_i+1} - x_{n_i}\|_2 \leq 2^{-i}.$$

The construction of this sequence is done by induction. By definition, for $\epsilon = 2^{-1}$ there is N_1 such that $\|x_n - x_m\|_2 \leq 2^{-1}$ for $n, m \geq N_1$. We take $n_1 = N_1$. Then, for $\epsilon = 2^{-2}$, there is $N_2 > N_1$ such that $\|x_n - x_m\|_2 \leq 2^{-2}$ for $n, m \geq N_2$. Then we take $n_2 = N_2$. In this way we constructed a strictly increasing sequence $n_k = N_k, n_{k-1} = N_{k-1}$ such that $\|x_{n_k} - x_{n_k-1}\|_2 \leq 2^{-(k-1)}$ and $\|x_n - x_m\|_2 \leq 2^{-k}$ for $n, m \geq N_k$. We define

$$g_k(t) = \sum_{i=1}^{k} |x_{n_i+1}(t) - x_{n_i}(t)|. \tag{1.3.6}$$

Since each term is nonnegative, we can consider the function

$$g(t) = \sum_{i=1}^{\infty} |x_{n_i+1}(t) - x_{n_i}(t)|, \tag{1.3.7}$$

where $g_k(t)$ can be infinite at some points. However, using the triangle inequality

$$\|g_k\|_2 \leq \sqrt{\sum_{i=1}^{k} \int_0^1 |x_{n_i+1}(t) - x_{n_i}(t)|^2 dt} \leq \sqrt{\sum_{i=1}^{\infty} \frac{1}{2^i}} < 1,$$

but then from the Lebesgue monotone convergence theorem we have $\|g\|_2 \leq 1$. Thus, $g(t)$ is finite almost everywhere. Now, we have

$$x_{n_k}(t) = x_{n_1}(t) + g_k(t)$$

hence the subsequence $(x_{n_k}(t))_{k\in\mathbb{N}}$ is convergent almost everywhere to a measurable function which we denote $x(t)$ (to fix attention we define $x(t) = 0$ whenever $g_k(t)$ does not converge to a finite limit). Next we show that x is the limit of $(x_n)_{n\in\mathbb{N}}$ in the norm $\|\cdot\|_2$. Since $|x(t) - x_{n_k}(t)| \to 0$ almost everywhere as $n_k \to \infty$ we have

$$\int_0^1 |x(t) - x_n(t)|^2 dt \leq \liminf_{k\to\infty} \int_0^1 |x_{n_k}(t) - x_n(t)|^2 dt < \epsilon$$

by Fatou lemma and the fact that $(x_n)_{n\in\mathbb{N}}$ is Cauchy so that for any ϵ we can find N such that $\|x_n - x_{n_k}\| < \epsilon$ whenever $n, n_k \geq N$. Furthermore, we obtain that $\int_0^1 |x(t)|^2 dt < +\infty$. This argument also shows that that if $(y_n)_{n\in\mathbb{N}}$ is another sequence satisfying $\|x_n - y_n\|_2 \to 0$, $\|y_n - x\|_2 \to 0$. Hence, we succeeded in relating with any $[x]$ in the completion of $\mathcal{C}$ a unique (almost everywhere) function x with integrable square of modulus in such a way that $\|[x]\| = \|x\|_2$ and $\langle [x], [y]\rangle = \langle x, y\rangle$. In this interpretation the completion becomes the closure in a bigger space.

Conversely, consider a measurable function x such that $\int_0^1 |x(t)|^2 dt < 0$. Space of such functions we denote by $L_2([0,1])$, equipped with the scalar product (1.3.4). We have to recall two facts. First, the Luzin theorem which states that given a measurable function f with support of finite measure, for any ϵ there is a continuous function ϕ of compact support such that the measure of the set $\{t, f(t) \neq g(t)\}$ is smaller than ϵ and $\phi(t) \leq g(t)$. Further, we recall that any non-negative measurable function f there is a nondecreasing sequence of simple functions $(s_n)_{n\in\mathbb{N}}$ which converges to f everywhere. Thus, for every $0 \leq x \in L_2([0,1])$, we have a monotone sequence $(s_n)_{n\in\mathbb{N}}$ such that $0 \leq s_n \leq x$ from where each s_n is square integrable and hence its support has finite measure. By dominated convergence theorem, $s_n \to f$ in $L_2([0,1])$. Hence, given $x \in L_2([0,1])$ and $\epsilon > 0$ we find a simple function s such that $\|x - s\|_2 < \epsilon/2$. Since the support of s has clearly finite measure, for any $\epsilon' > 0$ there is a continuous function ϕ such that $\|s - \phi\|_2 < 2\sqrt{\epsilon}\max s$. Taking $\epsilon/2 = 2\sqrt{\epsilon}\max s$ we get $\|x - \phi\|$ which gives density of continuous functions in $L_2([0,1])$.

This way we identified the completion of $\mathcal{C}$ in $\|\cdot\|$ with $L_2([0,1])$.

Remark 1.3.1. To be more precise, if we leave it like this, then $\|\cdot\|_2$ is not a norm on $L_2([0,1])$ because from $\|x(t)\|_2 = 0$ only follows that $x(t) = 0$ almost everywhere, so the first axiom of norm is not satisfied (the function x defined as $x(t) = 1$ for rational t and $x(t) = 0$ for irrational t, clearly is not a zero function but integral of $x^2(t)$ is zero). This is solved by considering $L_2([0,1])$ as consisting of classes of equivalence of functions equal almost everywhere but in most practical applications this distinction is not essential.

Chapter 2

Projections and approximations

We start with proving that any nonempty, convex and closed subset of a Hilbert space contains one and only one element with smallest norm.

Theorem 2.0.2. *Let $A \neq \emptyset$ be a closed convex set in a Hilbert space. Then there is a unique element $x \in A$ satisfying*

$$\|x\| = \inf_{y \in A} \|x\| \tag{2.0.1}$$

Proof. Define $\delta = \inf\{\|y\|;\ y \in A\}$. The set is bounded from below hence the infimum exists. First we prove uniqueness. Consider $x, y \in A$ with $\|x\| = \|y\| = \delta$ and the midpoint $(x+y)/4$ which, by convexity, belongs to A so that $\|(x+y)/2\| \geq \delta$. Using the parallelogram law (1.2.1) we get

$$\frac{1}{4}\|x-y\|^2 = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 - \left\|\frac{x+y}{2}\right\|^2 \leq 0,$$

and hence $x = y$.

To prove existence, we note that from the definition of infimum there is a sequence $(y_n)_{n\in\mathbb{N}} \subset A$ satisfying $\lim_{n\to\infty} \|y_n\| = \delta$. Using again the parallelogram law, we find as above

$$\|y_n - y_m\|^2 = 2\|y_n\|^2 + 2\|y_m\|^2 - 4\left\|\frac{y_n + y_m}{2}\right\|^2 \leq 2\|y_n\|^2 + \|y_m\|^2 - 4\delta^2$$

which shows that $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Indeed, for any $\epsilon > 0$ there is N such that for any $n, m > N$ we have $\delta^2 \leq \|y_n\|^2, \|y_m\|^{\leq}\delta^2 + \epsilon^2$. Thus,

$$\|y_n - y_m\|^2 \leq 2\|y_n\|^2 + 2\|y_m\|^2 - 4\delta^2 \leq 4\delta^2 + 4\epsilon^2 - 4\delta^2 = 4\epsilon^2, \quad n, m > N.$$

Since H is a complete space and A is a closed subset of H, we get an element $y \in A$ such that $\lim_{n\to\infty} y_n = x$. Since the norm is a continuous functional, we get

$$\|x\| = \lim_{n\to\infty} \|y_n\| = \delta.$$

■

Corollary 2.0.3. *Let M be a closed convex subset of a Hilbert space H. For any $x \in H$ there is a unique $y \in M$ such that*

$$\|x - y\| = \inf_{z\in M} \|x - z\| \tag{2.0.2}$$

Proof. We want to minimize $\|x - z\| = \|z - x\|$ where $z \in M$. If we denote $z' = z - x$, then $z' \in -x + M$ (translation of M by $-x$). This set is also convex: $u, v \in -x + M$ means that $u = -x + u', v = -x + v'$ and

$$\alpha u + \beta v = -(\alpha + \beta)x + \alpha u' + \beta v' = -x + \alpha u' + \beta v' \in -x + M$$

and thus there is a unique element $\xi \in -x + M$ with the least norm:

$$\|\xi\| = \inf_{z'\in -x+M} \|z'\|$$

But then there unique $y \in M$ such that $\xi = y - x$ so that

$$\|y - x\| = \inf_{z'\in -x+M} \|z'\| = \inf_{z\in M} \|z - x\|$$

■

Theorem 2.0.4. *Let M be a closed subspace of a Hilbert space H. Any $x \in H$ can be written in the form*

$$x = y + z, \tag{2.0.3}$$

where $y \in M$, $z \in M^\perp$ are uniquely determined by x.

Proof. If $x \in M$, then $x = y$ and $z = 0$ gives the required decomposition. It is, moreover, a unique decomposition since if $y' \in M$, $z' \in M^\perp$ is another pair giving $x = y' + z'$, then $0 = (y - y') - z'$, so that $z' = y - y'$. We obtain $\langle y - y', y - y'\rangle = \langle y - y', z'\rangle = 0$ yielding $y = y'$.

Hence, let $x \notin M$. The translated hyperplane $x + M$ is closed and convex and thus it contains an element z of the least norm:

$$\|z\| = \inf_{\xi\in x+M} \|\xi\|$$

Since M is a linear subspace, there is unique $y \in M$ such that $z = x - y$ and, writing $\xi = x - \eta$ with arbitrary $\eta \in M$, we have

$$\|y - x\| = \|z\| = \inf_{\xi \in x+M} \|\xi\| = \inf_{\eta \in M} \|\eta - x\| \tag{2.0.4}$$

so that $y = x - z$ is unique element in M minimizing the distance from M to x. We have to prove that $z \perp M$. From the construction of $z \in x + M$ we have

$$0 \leq \langle z, z \rangle = \|z\|^2 \leq \|z - \alpha y\|^2 = \langle z - \alpha y, z - \alpha y \rangle$$

for any $y \in M$, $\alpha \in \mathbb{C}$ (so that $z - \alpha y \in x + M$). Take for simplicity $\|y\| = 1$. Then, multiplying out and simplifying

$$|\alpha|^2 - \alpha \langle y, z \rangle - \bar{\alpha} \langle z, y \rangle \geq 0$$

hence, taking $\alpha = \langle z, y \rangle$, we obtain

$$-|\langle y, z \rangle|^2 \geq 0$$

and $\langle y, z \rangle = 0$.

It remains to prove uniqueness. If $x = y + z = y' + z'$, then $M \ni (y - y') = -(z - z') \in M^\perp$ and, multiplying by $(y - y')$, by orthogonality, $y = y'$. ■

Since the elements y, z, constructed in the above theorem are unique, we can define operators

$$y = Px, \qquad z = Qx. \tag{2.0.5}$$

such that

$$x = Px + Qx. \tag{2.0.6}$$

We summarize properties of these operators in the following theorem.

Corollary 2.0.5. *The operators P and Q have the following properties.*

1. *If $x \in M$, then $Px = x, Qx = 0$. If $x \in M^\perp$, then $Px = 0, Qx = x$;*
2. *$\|x - Px\| = \inf\{\|x - y\|;\ y \in M\}$;*
3. *$\|x\|^2 = \|Px\|^2 + \|Qx\|^2$;*
4. *P and Q are linear operators.*

Proof. Property 1. follows from the first part of the proof above. Property 2. follows is the formulation of (2.0.4). For the property 3., from the construction we have $Px \perp Qx$ so that, by (2.0.6),

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

To prove linearity, we take $x, y, \alpha x + \beta y$ for arbitrary $x, y \in H$, $\alpha, \beta \in \mathbb{C}$ so that

$$\begin{aligned} x &= Px + Qx, \\ y &= Py + Qy, \\ \alpha x + \beta y &= P(\alpha x + \beta y) + Q(\alpha x + \beta y) \end{aligned}$$

Multiplying the first equation by α, the second by β and subtracting from the third, we get

$$0 = P(\alpha x + \beta y) - \alpha Px - \beta Py + Q(\alpha x + \beta y) - \alpha Qx - \beta Qy,$$

that is,

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

Since left hand side is orthogonal to the right hand side, each must equal zero, which gives linearity. ■

We have seen that, for each fixed $y \in H$, the mapping $x \mapsto \langle x, y \rangle$ is a continuous linear functional on H. It is of great importance that all continuous functionals are of this form.

Theorem 2.0.6 (Riesz representation theorem)**.** *If L is a continuous linear functional on a Hilbert space H, then there is exactly one element $y \in H$ such that*

$$Lx = \langle x, y \rangle. \tag{2.0.7}$$

Proof. If $Lx = 0$ for all $x \in H$, then clearly we can put $y = 0$. Hence, assume that $L \neq 0$ and consider

$$M = \{x \in H;\ Lx = 0\}.$$

Since L is continuous and linear, M is a proper closed linear subspace of H. Thus, there is $x' \notin M$ and hence from Theorem 2.0.4 there is $y' \in M^\perp$. Let y'' be another element of $M^\perp$ which is linearly independent of y'; that is, for any $0 \neq \alpha, \beta \in \mathbb{C}$, $0 \neq z = \alpha y' + \beta y'' \in M^\perp$. Then, $0 \neq Lz = \alpha Ly' + \beta Ly''$. However, for $\alpha = -Ly''$ and $\beta = Ly'$ this combination is zero, and hence $z \in M$ which is impossible. Thus, $M^\perp$ is one-dimensional. To fix attention, we take $y' \in M^\perp$ with $\|y'\| = 1$. Then, for any $x \in H$ we have a unique decomposition

$$x = \alpha y' + z$$

with $z \in M$, and thus

$$Lx = \alpha Ly'$$

and, on the other hand, $\langle x, y' \rangle = \alpha \langle y', y' \rangle + \langle z, y' \rangle = \alpha$. Hence,

$$Lx = \langle x, (\overline{Ly'}) y' \rangle$$

and we can take $y = (\overline{Ly'})y'$.

We need to prove uniqueness. Considering two elements y, v such that $Lx = \langle x, y \rangle = \langle x, v \rangle$, we find

$$\langle x, y - v, = \rangle 0$$

for all $x \in H$. Taking $x = y - v$, we get $\|y - v\|^2 = 0$, thus $y = v$. ■

2.1 Orthonormal sets and Fourier series

A set M in a Hilbert space is called orthonormal if each element of M is a unit element: for any $x \in M$

$$\|x\| = 1$$

and if for any two elements $x \neq y \in M$ we have

$$x \perp y.$$

We note the following simple fact:

Proposition 2.1.1. *If the set $M = \{e_1, \ldots, e_k\}$ is orthonormal and $x \in \mathcal{L}inM$, then*

$$x = \sum_{i=1}^{k} c_k e_k \tag{2.1.1}$$

where $c_i = \langle x, e_i \rangle$, $i = 1, \ldots, k$ are called Fourier coefficients of x. Moreover

$$\|x\|^2 = \sum_{i=1}^{k} |c_i|^2. \tag{2.1.2}$$

Proof. Since $x \in \mathcal{L}inM$, there must be $\{c_1, \ldots, c_k\}$ such that

$$x = \sum_{i=1}^{k} c_k e_k.$$

Since $\{e_i\}_{1\le i\le k}$ is orthonormal, we have

$$\langle x, e_i\rangle = c_i\langle e_i, e_i\rangle = c_i, \qquad i = 1,\ldots,k.$$

Eq. (2.1.2) we obtain by

$$\|x\|^2 = \langle \sum_{i=1}^{k} c_k e_k, \sum_{i=1}^{k} c_k e_k\rangle = \sum_{i,j=1}^{k} \langle c_i e_i, c_j e_j\rangle = \sum_{i=1}^{k} |c_i|^2.$$

■

Motivation–an approximation problem. From Corollary 2.0.3 we know that for any closed convex non-empty set $A \subset H$ and $x \in H$ there exists a unique element $y \in A$ which is closest to x. y is called the best approximation to x in A. A problem of practical importance is how to find y. It has a relatively simple answer if A is a linear subspace spanned by vectors $\{v_1,\ldots,v_k\}$; that is $A = \mathcal{L}in\{v_1,\ldots,v_k\}$. A is closed as a finite dimensional linear space. To avoid trivial case, we assume that $x \notin A$. Hence, we know that there is $y \in A$ such that

$$\|x - y\| = \inf_{z\in A} \|x - z\|.$$

However, $y = \sum_{i=1}^{k} c_i v_k$ so that question is to find $\{c_1,\ldots,c_k\}$ among all k-tuples $\{\lambda_1,\ldots,\lambda_k\}$ satisfying

$$\|x - \sum_{i=1}^{k} c_i v_k\| \le \|x - \sum_{i=1}^{k} \lambda_i v_k\|.$$

From the proof of Theorem 2.0.4 we infer that $x - \sum_{i=1}^{k} c_i v_k \perp A$ which is equivalent to

$$\langle x - \sum_{i=1}^{k} c_i v_k, v_j\rangle = 0, \qquad j = 1,\ldots,k,$$

which yields the system of algebraic equations for $\{c_1,\ldots,c_k\}$:

$$\sum_{j=1}^{k} a_{ij} c_j = b_i \tag{2.1.3}$$

where $a_{ij} = \langle v_i, v_j\rangle$, $b_i = \langle x, v_i\rangle$. Since we know that the system has exactly one solution, the matrix $\{a_{ij}\}_{1\le i,j\le k}$ is invertible and thus (2.1.3) uniquely determines $(c_1,\ldots,c_k)$.

If δ is the smallest distance from x to A, then we obtain

$$\begin{aligned}\delta^2 &= \langle x-y, x-y\rangle = \langle x, x-y\rangle = \langle x, x-\sum_{i=1}^{k} c_j v_j\rangle = \|x\|^2 - \sum_{i=1}^{k} \bar{c_j} b_j \\ &= \|x\|^2 - \sum_{i=1}^{k} c_j \bar{b_j}\end{aligned}$$

However, this could be cumbersome for higher dimensional A and thus of interest are bases of A which are orthonormal as then $a_{ij} = \delta_{ij}$ so that $c_j = b_j$ and

$$\delta^2 = \|x\|^2 - \sum_{i=1}^{k} b_j^2. \tag{2.1.4}$$

Summarizing, we have the following result

Theorem 2.1.2. *If $\{e_1, \ldots, e_k\}$ is an orthonormal system in H and $x \in H$, then for any set of scalars $\{\lambda_1, \ldots, \lambda_k\}$ we have*

$$\|x - \sum_{i=1}^{k} \langle x, e_i\rangle e_i\| \leq \|x - \sum_{i=1}^{k} \lambda_i e_i\|. \tag{2.1.5}$$

Inequality (2.1.5) turns into equality if and only if $\lambda_i = \langle x, e_i\rangle$ for $i = 1, \ldots, k$. The vector

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

is the orthogonal projection of x onto $\mathcal{L}in\{e_1, \ldots, e_k\}$ and, denoting by δ the distance between x and this subspace, we obtain

$$\sum_{i=1}^{k} b_j^2 = \|x\|^2 - \delta^2. \tag{2.1.6}$$

Corollary 2.1.3. *If $\{e_i\}_{i\in\mathbb{N}}$ is an orthonormal sequence in H, then*

$$\sum_{i=1}^{\infty} |\langle x, e_i\rangle|^2 \leq \|x\|^2 \tag{2.1.7}$$

Proof. From (2.1.6) we have

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

for any finite k so monotonicity of the series gives the convergence and the limit estimate. ∎

Theorem 2.1.4. *Let $\{e_i\}_{i\in\mathbb{N}}$ is an orthonormal sequence in H and let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of scalars. The series*

$$\sum_{i=1}^{\infty} \alpha_i e_i \tag{2.1.8}$$

converges (unconditionally) if and only if

$$\sum_{i=1}^{\infty} |\alpha_i|^2 < +\infty. \tag{2.1.9}$$

In such a case

$$\left\| \sum_{i=1}^{\infty} \alpha_i e_i \right\|^2 = \sum_{i=1}^{\infty} |\alpha_i|^2. \tag{2.1.10}$$

Proof. If we take $m > n$ and use (2.1.2), we get

$$\left\| \sum_{i=n}^{m} \alpha_i e_i \right\|^2 = \sum_{i=n}^{m} |\alpha_i|^2$$

hence the series (2.1.8) satisfies the Cauchy criterion if and only if the scalar series (2.1.9) does. Thus, the completeness of H gives the first statement (without unconditionality). Taking $n = 1$ and letting $m \to \infty$ proves (2.1.10).

The scalar series (2.1.9) converges absolutely and thus unconditionally. Let $z = \sum_{n=1}^{\infty} \alpha_{j_n} e_{j_n}$ be any rearrangement of the series (2.1.8) (which converges by virtue of the previous remark and the first part of the proof. Consider

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

where the first two terms on the right hand side equal $\sum_{j=1}^{\infty} |\alpha_j|^2$. If we consider $s_m = \sum_{j=1}^{m} \alpha_j e_j$ and $t_m = \sum_{n=1}^{m} \alpha_{j_n} e_{j_n}$ we see that

$$\langle s_m, t_m\rangle = \sum_{n\le m, j_n \le m} \|\alpha_{j_n}\|^2.$$

Finally, we observe that for any j_n there is m such that $n \le m$ and $j_n \le m$ (namely, $m = \max\{n, j_n\}$) and thus, by continuity of the scalar product,

$$\lim_{m\to\infty} \langle s_m, t_m\rangle = \sum_{n=1}^{\infty} \|\alpha_{j_n}\|^2 = \|z\|^2 = \|x\|^2$$

which shows that $< x, z >= \|x\|^2$. But then $< z, x >= \|x\|^2$ and hence $x = z$. ∎

Theorem 2.1.5. *Let M be an orthonormal set in a Hilbert space H. Then,*

1. *for any $x \in H$, $\langle x, y\rangle = 0$ for all but a countable number of $y \in M$;*

2. *The sum*
$$Px = \sum_{y\in M_x} \langle x, y\rangle y \tag{2.1.11}$$
where $M_x = \{y \in M;\ \langle x, y\rangle \neq 0\}$ is well defined in the sense that it does not depend on the way the countable set M_x is numbered;

3. *the operator P is the projection on $\overline{\mathcal{L}inM}$.*

Proof. a) For a given $x \in H$, consider $M_n = \{y \in M;\ \langle x, y\rangle \geq 1/n\}$ for some n. From Bessel's inequality we infer that the number of such points cannot exceed $\|x\|^2/n^2$. Indeed, assume we had more points (even uncountably many). Then taking $k > \|x\|^2 n^2$ of them and numbering them from 1 to k in an arbitrary order (as there are only finitely many) as $\{y_1, y_2, \ldots, y_k\}$ we would have
$$\sum_{i=1}^{k} |\langle x, y_i\rangle|^2 \geq \frac{k}{n^2} > \|x\|^2$$
which contradicts the Bessel inequality. Since $M = \bigcup_{n=1}^{\infty} M_n$, M is at most countable.

b) follows from a) and the previous theorem as for each $x \in H$ in the sum (2.1.11) has only countably many components which can be numbered into a sequence whose sum does not depend on the way in which it was arranged.

c) Denote $\mathcal{M} = \overline{\mathcal{L}inM}$. If $x \perp \mathcal{M}$, then $Px = 0$ by linearity and continuity of the scalar product. If $x \in \mathcal{M}$, then for any $\epsilon > 0$ there are scalars $\lambda_1, \ldots, \lambda_n$ and $\{y_1, \ldots, y_n\} \subset M$ such that
$$\left\| x - \sum_{j=1}^{n} \lambda_j y_j \right\| < \epsilon.$$
Then, by (2.1.5),
$$\left\| x - \sum_{j=1}^{n} \langle x, y_j\rangle y_j \right\| < \epsilon.$$

We may assume that in the sum above we have $y_j \in M_x$ for $j = 1, \ldots, n$, otherwise we could remove elements with $\langle x, y_j \rangle = 0$. Now, arrange the set M_x into a sequence as $\{y_1, \ldots, y_n, \ldots\}$. By (2.1.6) we have

$$\left\| x - \sum_{j=1}^{n} \langle x, y_j \rangle y_j \right\|^2 = \|x\|^2 - \sum_{j=1}^{n} |\langle x, y_j \rangle|^2$$

which means that adding additional terms to the series does not increase the left hand side. Hence

$$\left\| x - \sum_{j=1}^{\infty} \langle x, y_j \rangle y_j \right\|^2 < \epsilon$$

and, since ϵ was arbitrary, we eventually get $\|x - Px\| = 0$ so that $Px = x$ for $x \in \mathcal{M}$. Hence, P is a an orthogonal projection onto $\mathcal{M} = \overline{\mathcal{L}inM}$. ■

Our interest is to be able to decompose arbitrary element of H into a sum (series) of orthogonal elements. This is possible in finite dimensional spaces, as demonstrated in (2.1.1). Following this idea, we say that M is an *orthonormal basis* for H if it is an orthonormal set and for any $x \in H$

$$x = \sum_{y \in M_x} \langle x, y \rangle y. \tag{2.1.12}$$

Theorem 2.1.6. *Let M be an orthonormal set in a Hilbert space H. Then the following conditions are equivalent:*

a) M is complete; that is, $M^{\perp} = \{0\}$;

b) $\overline{\mathcal{L}inM} = H$;

c) M is an orthonormal basis;

d) For any $x \in H$,

$$\|x\|^2 = \sum_{y \in M_x} |\langle x, y \rangle|^2. \tag{2.1.13}$$

Remark 2.1.7. The relation (2.1.13) is called *Parseval's formula.*

Proof. a)→ b). Assume that $\overline{\mathcal{L}inM} \neq H$, then there is $H \ni x \notin \overline{\mathcal{L}inM}$. But then $x = y + z$ with $z \neq 0$ and $z \in \overline{\mathcal{L}inM}^{\perp} = M^{\perp}$.

b)→ c). If b) holds, then by Theorem 2.1.5 3., $Px = x$ for any x, that is, M is an orthonormal basis.

c)→ d). Suppose c) holds. We arrange M_x in a sequence $(x_n)_{n\in\mathbb{N}}$ and, writing (2.1.6),

$$\left\|x-\sum_{j=1}^{n}\langle x,y_j\rangle y_j\right\|^2=\|x\|^2-\sum_{j=1}^{n}|\langle x,y_j\rangle|^2$$

and using the definition of the basis, we get (2.1.13).

d)→ a). If (2.1.13) holds and $x\perp\overline{\mathcal{L}inM}$ then

$$0=\sum_{y\in M_x}|\langle x,y\rangle|^2=\|x\|^2,$$

hence $x=0$. ■

Theorem 2.1.8. *In any separable Hilbert space with non-zero dimension there exists an orthonormal basis (finite or denumerable).*

Proof. Assume first that $dimH=n<\infty$. Then there exists a basis in H, say $B=\{v_1,v_2,\ldots,v_n\}$. Such basis can be transformed into an orthonormal basis of H constructed by a process called the Gram-Schmidt orthonormalization process. We proceed by induction. Define

$$e_1=\frac{v_1}{\|v_1\|},\quad e_2=\frac{v_2-\langle v_2,e_1\rangle e_1}{v_2-\langle v_2,e_1\rangle e_1}.$$

Since v_1 and v_2 are linearly independent, $v_2-\langle v_2,e_1\rangle e_1\neq 0$ so e_2 is well defined. Clearly $\|e_2\|=1$ and $\langle e_1,e_2\rangle=0$, so that $\{e_1,e_2\}$ are linearly independent. Moreover, $v_i\in\mathcal{L}in\{e_1,e_2\}$ i $e_i\in\mathcal{L}in\{v_1,v_2\}$; that is both sets span the same twodimensional space. Assume that we have defined orthonormal elements $\{e_1,\ldots,e_k\}$ spanning the same space as $\{v_1,\ldots,v_k\}$, $k<n$ and define

$$e_{k+1}=\frac{v_{k+1}-\sum\limits_{i=1}^{k}\langle v_{k+1},e_k\rangle e_k}{\left\|v_{k+1}-\sum\limits_{i=1}^{k}\langle v_{k+1},e_k\rangle e_k\right\|}.$$

As before $v_{k+1}-\sum\limits_{i=1}^{k}\langle v_{k+1},e_k\rangle e_k\neq 0$ as otherwise, v_{k+1} would be a linear combination of $\{e_1,\ldots,e_k\}$ and thus, by the induction assumption, of $\{v_1,\ldots,v_k\}$. Clearly, $\langle e_{k+1},e_i\rangle=0$ for any $i=1,\ldots,k$ and $e_{k+1}\in\mathcal{L}in\{v_1,v_2,\ldots,v_{k+1}\}$ as well as $v_{k+1}\in\mathcal{L}in\{e_1,e_2,\ldots,e_{k+1}\}$. In this way we exhaust all elements of B constructing an orthonormal basis spanning the same space.

Let us pass to the case when $dim H = \infty$ but H is separable. We take a denumerable dense set $Z = \{v_1, v_2, \ldots\}$ and define a subsequence in the following way: $v_{n_1} = v_1$, v_{n_2} is the first element of Z linearly independent of v_{n_1}, v_{n_3} the first element of Z not belonging to $\mathcal{L}in\{v_{n_1}, v_{n_2}\}$ and we proceed in this way by induction. The sequence $(v_{n_k})_{k\in\mathbb{N}}$ is infinite. Indeed, otherwise we would have a number n_0 such that $a_n \in X_0 = \mathcal{L}in\{v_{n_1}, \ldots, v_{n_{n_0}}\}$; that is, $Z \subset X_0$. However, X_0 is finite dimensional and thus closed and therefore $H = \bar{Z} \subset X_0 \subset H$ and $H = X_0$ would be finite dimensional.

By construction, the elements of $(v_{n_k})_{k\in\mathbb{N}}$ are linearly independent. Moreover,

$$v_n \in \mathcal{L}in\{v_{n_1}, \ldots, v_{n_i}\}$$

whenever $n \le n_i$; that is any element Z is in the linear span of the elements of the sequence $(v_{n_k})_{k\in\mathbb{N}}$. We can apply the Gram-Schmidt orthonormalization procedure to $(v_{n_k})_{k\in\mathbb{N}}$ in the same way as we did for finite dimensional case. It remains to prove that that the obtained set $\{e_i\}_{i\in\mathbb{N}}$ is complete. Since $v_n \in \mathcal{L}in\{v_{n_1}, \ldots, v_{n_i}\} = \mathcal{L}in\{e_1, \ldots, e_i\}$, we find that for any n, $v_n \in \mathcal{L}in\{e_1, e_2, \ldots\}$. Thus $Z \subset \overline{\mathcal{L}in\{e_1, e_2, \ldots\}}$ and, $H = \bar{Z} \subset \overline{\mathcal{L}in\{e_1, e_2, \ldots\}} \subset H$. Hence, the condition b) of Theorem 2.1.6 is satisfied and therefore $\{e_1, e_2, \ldots\}$ is an orthonormal basis. ■

Remark 2.1.9. In fact, any Hilbert space has an orthonormal basis but the proof requires using Zorn's lemma and the basis need not be countable. On the other hand, in separable spaces bases are always countable. In fact, let $\{x_\alpha\}$ be an uncountable orthonormal basis. Let B_α be the ball centred at x_α and with radius $1/2$. Since

$$\|x_\alpha - x_\beta\|^2 = \|x_\alpha\|^2 + \|x_\beta\|^2 = 2, \qquad \alpha \neq \beta$$

the balls B_α are mutually disjoint. Hence, any sequence would miss some balls and thus no sequence could be dense in H.

Theorem 2.1.10. *Any separable Hilbert space is isometrically isomorphic to l_2 (and thus any two separable Hilbert spaces are isometrically isomorphic).*

Proof. Let H be a separable Hilbert space. There exists an orthonormal basis $\{e_1, e_2, \ldots\}$ in H. Hence

$$x = \sum_{n=1}^{\infty} b_n e_n$$

where $b_n = \langle x, e_n \rangle$ with

$$\|x\|^2 = \sum_{n=1}^{\infty} |b_n|^2. \tag{2.1.14}$$

We define $T_H x = (b_n)_{n\in\mathbb{N}}$. T_H is clearly a linear operator from H into l_2 which, by (2.1.14) is an isometry. By Theorem 2.1.4 it is surjective. It is also injective as the only x mapped onto $(b_n)_{n\in\mathbb{N}} = (0, 0, \ldots)$ is x which orthogonal to all $e_k, k = 1, 2, \ldots$; that is, $x = 0$ by the (2.1.13). Thus T_H is also an isometry.

Isometry between spaces H_1 and H_2 is obtained by composition $T_{H_2}^{-1}T_{H_1}$. ■

2.2 Trigonometric series

We know that the system $B_1 = \left\{\frac{e^{int}}{\sqrt{2\pi}}\right\}_{n\in\mathbb{Z}}$ is orthonormal in $L_2([-\pi, \pi])$. To shorten notation, we denote $I = [-\pi, \pi]$. Using Euler formulae, it is easy to see that any linear combination (with complex coefficients)

$$f(t) = \sum_{n=-N}^{N} c_n e^{int} \tag{2.2.1}$$

can be written as a *trigonometric polynomial*

$$f(t) = a_0 + \sum_{n=1}^{N}(a_n \cos nt + i \sin nt). \tag{2.2.2}$$

If we denote

$$B_2 = \left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos t, \frac{1}{\sqrt{\pi}}\sin t, \frac{1}{\sqrt{\pi}}\cos 2t, \frac{1}{\sqrt{\pi}}\sin 2t, \ldots\right\}$$

we see that also B_2 is an orthonormal set in $L_2(I)$ and $\overline{\mathcal{L}inB_1} = \overline{\mathcal{L}inB_2}$. Our aim is to show that B_2 (and thus B_1) is a basis in $L_2(I)$. By Theorem 2.1.6 b), it is enough to show that the set of trigonometric polynomials is dense in $L_2(I)$. By the construction of $L_2(I)$ (as the completion of the space of continuous functions), the space $\mathcal{C}(I)$ is dense in $L_2(I)$. In fact, the set $\mathcal{C}_0(I)$ of continuous functions vanishing at the endpoints of I is dense in $L_2(I)$. Thus it suffices to show that for any ϵ and any continuous function f on I, there is a trigonometric polynomial P such that $\|f - P\|_\infty \leq \epsilon$ (where $\|\cdot\|_\infty$ is the norm in the space of continuous functions $C(I)$. Indeed, for any continuous function g we have

$$\|g\|_2^2 = \sqrt{\int_{-\pi}^{\pi} |g(t)|^2 dt} \leq \sqrt{2\pi}\|g\|_\infty$$

so that $\|f - P\|_2 \le \sqrt{2\pi}\|f - P\|_\infty \le \sqrt{2\pi}\epsilon$ and the approximation carries over to the L_2 norm.

Theorem 2.2.1. *For any $f \in C_0(I)$ and $\epsilon > 0$ there is a trigonometric polynomial P such that*

$$|f(t) - P(t)| \le \epsilon \tag{2.2.3}$$

for any $t \in I$.

Proof. We extend f to $\mathbb{R}$ is a periodic way. Since $f(-\pi) = f(\pi) = 0$ such an extension, still denoted by f, is continuous on $\mathbb{R}$ (and even uniformly continuous!). Assume that f We start from observation that if Q is a trigonometric polynomial, then

$$P(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)Q(s)ds$$

is also a trigonometric polynomial. We have

$$\int_{-\pi}^{\pi} f(t-s)Q(s)ds = \int_{t-\pi}^{t+\pi} f(r)Q(t-r)dr = \int_{-\pi}^{\pi} f(r)Q(t-r)dr$$

where we used the fact that $g(r) = f(r)Q(t-r)$ is a periodic function and the integral of a periodic function over the interval of the period length does not depend on the position of the interval. But $Q(t) = \sum_{n=-N}^{N} a_n e^{int}$ so

$$\int_{-\pi}^{\pi} f(r)Q(t-r)dr = \sum_{n=-N}^{N} a_n e^{int}\int_{-\pi}^{\pi} f(r)e^{-inr}dr = \sum_{n=-N}^{N} A_n e^{int}$$

where $A_n = a_n\int_{-\pi}^{\pi} f(r)e^{-inr}dr$.

Let us fix $\epsilon > 0$. Since f is uniformly continuous on I, there is $\delta > 0$ such that if $|t - s| < \delta$, then $|f(t-s) - f(t)| < \epsilon$. Assume now that Q has been chosen so that $Q \ge 0$ on I and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} Q(s)ds = 1. \tag{2.2.4}$$

Then

$$P(t) - f(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (f(t-s) - f(t))Q(s)ds$$

and

$$
\begin{aligned}
|P(t) - f(t)| &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t-s) - f(t)| Q(s) ds \\
&= \frac{1}{2\pi} \int_{-\delta}^{\delta} |f(t-s) - f(t)| Q(s) ds + \frac{1}{2\pi} \int_{[-\pi,-\delta)\cup(\delta,\pi]} |f(t-s) - f(t)| Q(s) ds \\
&= I_1 + I_2
\end{aligned}
$$

From uniform continuity, positivity of Q and (2.2.4) we have

$$
I_1 \leq \frac{\epsilon}{2\pi} \int_{-\delta}^{\delta} Q(s) ds \leq \frac{\epsilon}{2\pi} \int_{-\pi}^{\pi} Q(s) ds = \epsilon.
$$

For the second integral we have

$$
I_2 \leq 2\|f\|_\infty \frac{1}{2\pi} \int_{[-\pi,-\delta)\cup(\delta,\pi]} Q(s) ds
$$

and we have to show that we can find a trigonometric polynomial which can be arbitrarily small outside a given interval about 0.

Such trigonometric polynomials can be constructed in various ways. Consider functions

$$
Q_k(t) = c_k \left(\frac{1 + \cos t}{2} \right)^k,
$$

where c_k are constants picked up so as that (2.2.4) holds. Clearly $Q_k \geq 0$. Further,

$$
\begin{aligned}
\left(\frac{1 + \cos t}{2} \right)^k &= \left(\frac{\cos^2 t/2 + \sin^2 t/2 + \cos^2 t/2 - \sin^2 t/2}{2} \right)^k \\
&= \cos^{2k} t/2 = \frac{\left(e^{it/2} + e^{-it/2}\right)^{2k}}{4^k} = \frac{1}{4^k} \sum_{j=0}^{2k} \binom{2k}{j} e^{it(j-k)},
\end{aligned}
$$

which, by the discussion at the beginning of this section, is a trigonometric polynomial.

We have to check that Q_k satisfy the last requirement. First we observe that Q_k is an even function so that

$$\begin{aligned} 1 &= \frac{c_k}{2\pi}\int_{-\pi}^{\pi}\left(\frac{1+\cos t}{2}\right)^k dt = \frac{c_k}{\pi}\int_{0}^{\pi}\left(\frac{1+\cos t}{2}\right)^k dt \\ &\geq \frac{c_k}{\pi}\int_{0}^{\pi}\left(\frac{1+\cos t}{2}\right)^k \sin t dt = \frac{2c_k}{\pi}\int_{0}^{1} u^k du = \frac{2c_k}{\pi(k+1)}, \end{aligned}$$

hence

$$c_k \leq \frac{\pi(k+1)}{2}.$$

Next, Q_k decreases on $[0,\pi]$ and thus

$$Q_k(t) \leq Q_k(\delta) \leq \frac{\pi(k+1)}{2}\left(\frac{1+\cos\delta}{2}\right)^k$$

where we can take $0 < \delta \leq |t| \leq \pi$ since Q_k is even. Now, $0 \leq \cos\delta < 1$ for $0 < \delta \leq \pi$ and thus $q := (1+\cos\delta)/2 < 1$. Now, $(k+1)q^k \to \infty$ as $k \to \infty$ and we can take k large enough to have $Q_k(t) < \epsilon/2\|f\|_\infty$ and for such k

$$I_2 < \epsilon.$$

Thus

$$|f(t) - P(t)| \leq \epsilon$$

if

$$P(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)Q_k(t-s)ds$$

and the theorem is proved. ■