4 Hilbert spaces

The proof of the Hilbert basis theorem is not mathematics, it is theology.
— Camille Jordan

Wir müssen wissen, wir werden wissen.
— David Hilbert

We now continue to study a special class of Banach spaces, namely Hilbert spaces, in which the presence of a so-called "inner product" allows us to define angles between elements. In particular, we can introduce the geometric concept of orthogonality. This has far-reaching consequences.

4.1 Definition and Examples

Definition 4.1. Let H be a vector space over $\mathbb{K}$. An **inner product** (or a **scalar product**) on H is a map $(\cdot\,|\,\cdot)\colon H \times H \to \mathbb{K}$ such that the following three properties hold.

(IP1) For all $x \in H$, one has $(x\,|\,x) \geq 0$, and $(x\,|\,x) = 0$ if and only if $x = 0$.

(IP2) For all $x, y \in H$, one has $(x\,|\,y) = \overline{(y\,|\,x)}$.

(IP3) For all $x, y, z \in H$ and $\lambda \in \mathbb{K}$, one has $(\lambda x + y\,|\,z) = \lambda(x\,|\,z) + (y\,|\,z)$.

The pair $(H, (\cdot\,|\,\cdot))$ is called **inner product space** or **pre-Hilbert space**.

Example 4.2. (a) On $H = \mathbb{K}^N$,

$$(x\,|\,y) := \sum_{k=1}^{N} x_k \overline{y}_k$$

defines an inner product.

(b) On $H = \ell^2$,

$$(x \mid y) := \sum_{k=1}^{\infty} x_k \overline{y}_k$$

for $x = (x_k)$ and $y = (y_k)$ defines an inner product.

(c) On $C([a,b])$,

$$(f \mid g) := \int_a^b f(t)\overline{g(t)}\,\mathrm{d}t$$

defines an inner product. If $w\colon [a,b] \to \mathbb{R}$ is such that there exist constants $0 < \varepsilon < M$, such that $\varepsilon \leq w(t) \leq M$ for all $t \in [a,b]$, then also

$$(f \mid g)_w := \int_a^b f(t)\overline{g(t)}w(t)\,\mathrm{d}t$$

defines an inner product on $C([a,b])$.

(d) If (Ω, Σ, μ) is a measure space, then

$$(f \mid g) := \int_\Omega f\overline{g}\,\mathrm{d}\mu$$

defines an inner product on $L^2(\Omega, \Sigma, \mu)$.

Note that for examples (b), (c) and (d), it follows from Hölder's inequality that the sum, respectively the integral, is well-defined.

Remark 4.3. Condition (IP1) is called *definiteness* of $(\cdot \mid \cdot)$, (IP2) is called *symmetry*. Note that if $\mathbb{K} = \mathbb{R}$, then (IP2) reduces to $(x \mid y) = (y \mid x)$. (IP3) states that $(\cdot \mid \cdot)$ is linear in the first component. Note that it follows from (IP2) and (IP3) that

$$(x \mid \lambda y + z) = \overline{\lambda}(x \mid y) + (x \mid z).$$

If $(H, (\cdot \mid \cdot))$ is an inner product space, we set

$$\|x\| := \sqrt{(x \mid x)}.$$

Lemma 4.4 (Cauchy–Schwarz)**.** *Let $(H, (\cdot \mid \cdot))$ be an inner product space. Then*

$$|(x \mid y)| \leq \|x\| \cdot \|y\| \tag{4.1}$$

for all $x, y \in H$ and equality holds if and only if x and y are linearly dependent.

Proof. Let $\lambda \in \mathbb{K}$. Then

$$\begin{aligned} 0 \leq (x + \lambda y \mid x + \lambda y) &= (x \mid x) + \lambda(y \mid x) + \overline{\lambda}(x \mid y) + |\lambda|^2 (y \mid y) \\ &= \|x\|^2 + 2\,\mathrm{Re}(\overline{\lambda}(x \mid y)) + |\lambda|^2 \|y\|^2. \end{aligned}$$

If $y = 0$, then $(x \mid y) = 0$ and the claimed inequality holds true. If $y \neq 0$, we may put $\lambda = -(x \mid y)\|y\|^{-2}$. Then we obtain

$$0 \leq \|x\|^2 - 2\,\mathrm{Re}\,\frac{|(x \mid y)|^2}{\|y\|^2} + \frac{|-(x \mid y)|^2}{\|y\|^4}\|y\|^2 = \|x\|^2 - \frac{|(x \mid y)|^2}{\|y\|^2},$$

which establishes (4.1). By (IP1), equality holds if and only if $x = -\lambda y$. □

Lemma 4.5. *Let $(H, (\cdot \mid \cdot))$ be an inner product space. Then $\|\cdot\| := \sqrt{(\cdot \mid \cdot)}$ defines a norm on H.*

Proof. If $\|x\| = 0$, then $(x \mid x) = 0$ and therefore $x = 0$ by (IP1). This proves (N1). For (N2), let $x \in H$ and $\lambda \in \mathbb{K}$ be given. Then $\|\lambda x\|^2 = (\lambda x \mid \lambda x) = \lambda\overline{\lambda}(x \mid x) = |\lambda|^2\|x\|^2$. Property (N3) holds since

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}\,(x \mid y) + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2,\end{aligned}$$

where we used the Cauchy–Schwarz inequality. □

Definition 4.6. A **Hilbert space** is an inner product space $(H, (\cdot \mid \cdot))$ which is complete with respect to the norm $\|\cdot\| := \sqrt{(\cdot \mid \cdot)}$.

Example 4.7. The inner product spaces from Example 4.2 (a), (b) and (d) are Hilbert spaces, see for example Theorem 3.105. The inner product space from Example 4.2 (c) is not complete and therefore not a Hilbert space. It suffices to note that $\mathrm{C}([a,b])$ is dense in $L^2([a,b])$ and that the given inner product gives rise to the restriction of the usual norm in $L^2([a,b])$ to $\mathrm{C}([a,b])$.

Lemma 4.8 (Parallelogram identity)**.** *Let H be an inner product space. Then*

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

for all $x, y \in H$.

In fact, also a converse holds: If the parallelogramm identity holds in a normed space $(X, \|\cdot\|)$, then there exists a unique inner product $(\cdot \mid \cdot)$ on X such that $\|x\|^2 = (x \mid x)$.

Proof.

$$\begin{aligned}&\|x+y\|^2 + \|x-y\|^2 \\ &\qquad = \|x\|^2 + 2\,\mathrm{Re}\,(x \mid y) + \|y\|^2 + \|x\|^2 - 2\,\mathrm{Re}\,(x \mid y) + \|y\|^2 \\ &\qquad = 2(\|x\|^2 + \|y\|^2)\end{aligned}$$

□

Exercise 4.9. Show that in $(\mathbb{R}^d, \|\cdot\|_p)$, $(\ell^p, \|\cdot\|_p)$ and $(L^p([0,1]), \|\cdot\|_p)$ the parallelogram identity fails if $p \neq 2$. Hence in this case there is no inner product $(\cdot \mid \cdot)_p$ such that $\|x\|_p = \sqrt{(x \mid x)_p}$.

Lemma 4.10. *Let H be an inner product space. Then $(\cdot \mid \cdot)\colon H \times H \to \mathbb{K}$ is continuous.*

Proof. Let $x_n \to x$ and $y_n \to y$. Then $M := \sup_{n\in\mathbb{N}} \|x_n\| < \infty$. By the Cauchy–Schwarz inequality,

$$\begin{aligned}|(x_n \,|\, y_n) - (x \,|\, y)| &\le |(x_n \,|\, y_n - y)| + |(x_n - x \,|\, y)| \\ &\le M\|y_n - y\| + \|y\|\|x_n - x\| \to 0.\end{aligned}$$ □

4.2 Orthogonal Projection

In this section we show that in a Hilbert space one can project onto any closed subspace. In other words, for any point there exists a unique best approximation in any given closed subspace. For this the geometric properties arising from the inner product are crucial.

In fact, an inner product allows to define angles between two vectors. Consider $H = \mathbb{R}^2$ and describe the angle between two vectors $x, y \in \mathbb{R}^2$ with $\|x\| = \|y\| = 1$ in terms of the inner product!

Definition 4.11. Let $(H, (\cdot \,|\, \cdot))$ be an inner product space. We say that two vectors $x, y \in H$ are **orthogonal** (and write $x \perp y$) if $(x \,|\, y) = 0$. Given a subset $S \subset H$, the **annihilator** $S^\perp$ of S is defined by

$$S^\perp := \{y \in H : y \perp x \text{ for all } x \in S\}.$$

If S is a subspace, then $S^\perp$ is also called the **orthogonal complement** of S.

In an inner product space the following fundamental and classic geometric identity holds.

Lemma 4.12 (Pythagoras)**.** *Let H be an inner product space. If $x \perp y$, then*

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

Proof.

$$\|x + y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\,(x \,|\, y) + \|y\|^2 = \|x\|^2 + \|y\|^2.$$ □

Proposition 4.13. *Let H be an inner product space and $S \subset H$.*

(a) *$S^\perp$ is a closed, linear subspace of H.*

(b) *$\overline{\operatorname{span}}\, S \subset (S^\perp)^\perp$.*

(c) *$\overline{\operatorname{span}}\, S \cap S^\perp = \{0\}$.*

Proof. (a) If $x, y \in S^\perp$ and $\lambda \in \mathbb{K}$, then, for $z \in S$, we have $(\lambda x + y \,|\, z) = \lambda(x \,|\, z) + (y \,|\, z) = 0$, hence $\lambda x + y \in S^\perp$. This shows that $S^\perp$ is a linear subspace. If (x_n) is a sequence in $S^\perp$ which converges to x, then we infer from Lemma 4.10 that $(x \,|\, z) = \lim\,(x_n \,|\, z) = 0$ for all $z \in S$.

(b) By (a), $(S^\perp)^\perp$ is a closed linear subspace which contains S. Thus $\overline{\operatorname{span}}\, S \subset (S^\perp)^\perp$.

(c) If $x \in \overline{\operatorname{span}}\, S \cap S^\perp$, then, by (b), $x \in S^\perp \cap (S^\perp)^\perp$ and hence $x \perp x$. But this means $(x \,|\, x) = 0$. By (IP1) it follows that $x = 0$. □

We now come to the main result of this section.

Theorem 4.14. *Let $(H, (\cdot \mid \cdot))$ be a Hilbert space and $K \subset H$ be a closed linear subspace. Then for every $x \in H$, there exists a unique element $P_K x$ of K such that*

$$\|P_K x - x\| = \min\{\|y - x\| : y \in K\}.$$

Give an example of a subspace of ℓ^2 that is not closed!

Proof. Let $d := \inf\{\|y - x\| : y \in K\}$. By the definition of the infimum, there exists a sequence (y_n) in K with $\|y_n - x\| \to d$. Applying the parallelogram identity 4.8 to the vectors $x - y_n$ and $x - y_m$, we obtain

$$\begin{aligned} 2(\|x - y_n\|^2 + \|x - y_m\|^2) \\ &= \|(x - y_n) + (x - y_m)\|^2 + \|x - y_n - (x - y_m)\|^2 \\ &= \|2x - y_n - y_m\|^2 + \|y_n - y_m\|^2 \\ &= 4\left\|x - \frac{1}{2}(y_n + y_m)\right\|^2 + \|y_n - y_m\|^2. \end{aligned}$$

Since $z_{nm} := \frac{1}{2}(y_n + y_m) \in K$, we have $\|x - z_{nm}\|^2 \geq d^2$ and thus

$$\|y_n - y_m\|^2 \leq 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4d^2.$$

In a Hilbert space the orthogonal projection can be more generally defined for all closed and convex subsets. But then the orthogoal projection does not need to be a linear map, of course.

By the choice of the sequence (y_n), the right-hand side of this equation converges to 0 as $n, m \to \infty$, proving that (y_n) is a Cauchy sequence. Since H is complete, (y_n) converges to some vector $P_K x$. Since K is closed, $P_K x \in K$. We have thus proved existence.

As for uniqueness, if $\|z - x\| = \min\{\|y - x\| : y \in K\}$, then, by the parallelogram identity,

$$2\|x - P_K x\|^2 + 2\|x - z\|^2 = 4\left\|x - \frac{1}{2}(P_K x + z)\right\|^2 + \|z - P_K x\|^2$$

and thus

$$4d^2 = 4\left\|x - \frac{1}{2}(P_K x + z)\right\|^2 + \|P_K x - z\|^2 \geq 4d^2 + \|P_K x - z\|^2$$

proving that $\|P_K x - z\| = 0$; hence $P_K x = z$. □

Definition 4.15. The map $P_K \colon H \to H$ from Theorem 4.14 is called the **orthogonal projection onto K**.

We now collect some properties of P_K.

Proposition 4.16. *Let H be a Hilbert space, K be a closed subspace of H and P_K be the orthogonal projection onto K.*

(a) *For all $x, y \in H$, we have $P_K x = y$ if and only if $y \in K$ and $x - y \in K^\perp$.*

(b) *P_K is a bounded linear operator on H.*

(c) $P_K^2 = P_K$ *and* $(P_K x \,|\, y) = (x \,|\, P_K y)$ *for all* $x, y \in H$.

Proof. (a) If $y \in K$ and $x - y \in K^\perp$, then for every $z \in K$ we have $y - z \in K$ and thus $x - y \perp y - z$. By Pythagoras,

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

Thus $\|x - y\| = \min\{\|x - z\| : z \in K\}$, proving that $P_K x = y$.

Conversely, if $P_K x = y$, then clearly $y \in K$. Assume that $x - y \notin K^\perp$. Then there exists $z \in K \setminus \{0\}$ with $(x - y \,|\, z) \neq 0$. We may assume that $(x - y \,|\, z) = 1$ (otherwise, we divide z by $\overline{(x - y \,|\, z)}$). Then, for $\lambda \in \mathbb{R}$,

$$\begin{aligned} \|x - y - \lambda z\|^2 &= \|x - y\|^2 - 2 \operatorname{Re} \lambda (x - y \,|\, z) + \lambda^2 \|z\|^2 \\ &= \|x - y\|^2 - 2\lambda + \lambda^2 \|z\|^2. \end{aligned}$$

The latter is strictly less than $\|x - y\|^2$ for small $\lambda > 0$, for example if $\lambda^2 \|z\|^2 < 2\lambda$, i.e. $\lambda < 2\|z\|^{-2}$. Hence we find an element in K, namely $y + \|z\|^{-2} z$, for example, which is closer to x than to y. But then $y \neq P_K x$.

(b) Let $x, y \in H$ and $\lambda \in \mathbb{K}$. By (a), $x - P_K x, y - P_K y \in K^\perp$. Since $K^\perp$ is a subspace by Proposition 4.13, $\lambda x - \lambda P_K x + y - P_K y = (\lambda x + y) - (\lambda P_K x + P_K y) \in K^\perp$. Since $\lambda P_K x + P_K y \in K$, it follows from (a) that $P_K(\lambda x + y) = \lambda P_K x + P_K y$, i.e. P_K is linear. As for the boundedness, observe that $x = P_K x + (x - P_K x)$ where $P_K x \perp x - P_K x$ by (a). Thus, by Pythagoras,

$$\|x\|^2 = \|P_K x\|^2 + \|x - P_K x\|^2 \geq \|P_K x\|^2,$$

proving the boundedness of P_K.

(c) $P_K x \in K$ and $0 = P_K x - P_K x \in K^\perp$. Hence, by (a), $P_K P_K x = P_K x$.

For the second part, observe that

$$(P_K x \,|\, y) = (P_K x \,|\, P_K y) + (P_K x \,|\, y - P_K y) = (P_K x \,|\, P_K y)$$

since $y - P_K y \in K^\perp$ and $P_K x \in K$ by (a). Similarly, one sees that $(x \,|\, P_K y) = (P_K x \,|\, P_K y)$. □

We can now refine Proposition 4.13 for linear subspaces.

Corollary 4.17. *If* H *is a Hilbert space and* K *is a linear subspace of* H*, then* $\overline{K} = (K^\perp)^\perp$.

Proof. We have seen already that $\overline{K} \subset (K^\perp)^\perp$. Now let $y \in (K^\perp)^\perp$. Then $y = P_{\overline{K}} y + (I - P_{\overline{K}}) y =: y_1 + y_2$. Thus $\|y_2\|^2 = (y_2 \,|\, y_2) = (y_2 \,|\, y) - (y_2 \,|\, y_1) = 0$, since $y_2 \in \overline{K}^\perp = K^\perp$ and $y \in (K^\perp)^\perp$ and $y_1 \in \overline{K}$. It follows that $y_2 = 0$, hence $y = y_1 \in \overline{K}$. This shows $(K^\perp)^\perp \subset \overline{K}$. □

An important consequence of Theorem 4.14 is the following result, which shows that in a Hilbert space all bounded linear functionals can be expressed in a specific way in terms of the inner product.

Theorem 4.18 (Fréchet–Riesz)**.** *Let H be a Hilbert space. Then $\varphi \in H^*$ if and only if there exists a $y \in H$ such that $\varphi(x) = (x \,|\, y)$ for all $x \in H$.*

Recall that $H^* = \mathcal{L}(H;\mathbb{K})$ is the space of bounded linear functionals. The theorem of Fréchet–Riesz says that a Hilbert space is essentially its own dual.

Proof. If $\varphi(x) = (x \,|\, y)$, then φ is continuous as a consequence of Lemma 4.10.

Conversely, let $\varphi \in H^*$ be given. Then $K := \ker \varphi$ is a closed subspace of H. If $K = H$, pick $y = 0$. If $K \neq H$, there exists an $x_0 \in H$ with $\varphi(x_0) \neq 0$. Put $z = x_0 - P_K x_0$. Since $x_0 \notin K$, we have $z \neq 0$ and may thus define $w = \|z\|^{-1} z$. Then $\|w\| = 1$ and $w \in K^\perp$. In particular, $\varphi(w) \neq 0$.

Now for $x \in H$, we have $\varphi(x) = \frac{\varphi(x)}{\varphi(w)} \varphi(w)$. Define $\lambda := \frac{\varphi(x)}{\varphi(w)}$. Then, by linearity, $\varphi(x - \lambda w) = 0$ and thus $x - \lambda w \in K$. Put $y := \overline{\varphi(w)} w$. Then

$$\begin{aligned}(x \,|\, y) &= \varphi(w)(x \,|\, w) \\ &= \varphi(w)\big((x - \lambda w \,|\, w) + (\lambda w \,|\, w)\big) \\ &= \varphi(w)\lambda \|w\|^2 = \varphi(x). \end{aligned}$$ □

Exercise 4.19. Let H be a Hilbert space and K be a closed, linear subspace of H with $K \neq H$. Given $x_0 \in H \setminus K$, show that there exists $\varphi \in H^*$ such that $\varphi(x_0) = 1$ and $\varphi(x) = 0$ for all $x \in K$.

This shows that functionals in H^* can be used to separate points from subspaces (and, more generally, closed convex sets). Give a geometric interpretation in $\mathbb{R}^2$!

4.3 Orthonormal Bases

In this section we study orthonormal bases, which are a generalisation of the rectangular Cartesian coordinate systems of $\mathbb{R}^N$ to general Hilbert spaces. An orthonormal basis is very useful practically since it allows to effectively compute coefficients and associated orthogonal projections.

Definition 4.20. Let H be an inner product space. An **orthonormal system** is a subset $S \subset H$ such that

(i) $\|x\| = 1$ for all $x \in S$ (i.e. every vector in S is *normalized*), and

(ii) for every $x, y \in S$ with $x \neq y$ we have $(x \,|\, y) = 0$ (i.e. distinct vectors are *orthogonal*). An **orthonormal basis** of H is an orthonormal system S such that $\overline{\operatorname{span}}\, S = H$.

We frequently consider an orthonormal system to be indexed, i.e., $S = (x_j)_{j \in J}$. For $J = \mathbb{N}$, this leads to sequences.

Note the closure in the definition of an orthonormal basis. Whereas in an ordinary vector space basis (also called *Hamel basis*) every vector can be (uniquely) written as a *finite* linear combination of basis elements, this does not need to be true for an orthonormal basis. However, it will turn out that every element can be uniquely written as a (at most) countable series of multiples of the basis elements.

Example 4.21. Let $H = \ell^2$. Then define $S := \{e_j : j \in \mathbb{N}\}$, where $e_j = (0,\dots,0,1,0,\dots)$ and the 1 is at position j, is an orthonormal basis of H. Indeed, $\|e_j\| = 1$ and $(e_k \,|\, e_j) = 0$ for $k \neq j$. Finally, let $x = (x_1, x_2, \dots) \in \ell^2$. Then $y_n := (x_1,\dots,x_n,0,0,\dots) \in \operatorname{span} S$. Moreover, $\|x - y_n\|^2 = \sum_{k=n+1}^{\infty} |x_k|^2 \to 0$ as $n \to \infty$, proving that $\operatorname{span} S$ is dense in H.

Exercise 4.22. Let $H = L^2((0,2\pi), \mathscr{B}((0,2\pi)), \lambda)$ with the inner product

$$(f \,|\, g) = \int_0^{2\pi} f(t)\overline{g(t)}\,\mathrm{d}t.$$

Show that $S := \{e_k : k \in \mathbb{Z}\}$ where $e_k\colon [0,2\pi] \to \mathbb{C}$ is given by $e_k(t) := (2\pi)^{-1/2} e^{ikt}$ is an orthonormal basis of H.

*This orthonormal basis of $L^2(0,2\pi)$, which is also known as the **Fourier basis**, is very important for applications. Note that the functions e^{ikt} can be expressed in terms of sine and cosine functions, and vice versa. So the Fourier series from Example 4.34 can be understood as the description of functions as the superposition of trigonometric functions of different integer frequency.*

Hint: Use that by Corollary 2.71 the functions e_k are dense in the continuous 2π periodic functions. Then use a truncation argument and the density of the continuous functions in $L^2((0,2\pi))$.

Lemma 4.23. *Let S be an orthonormal system and $e_1,\dots,e_n \in S$ be distinct. If for scalars $\lambda_1,\dots,\lambda_n \in \mathbb{K}$ we have $\sum_{k=1}^n \lambda_k e_k = 0$, then $\lambda_1 = \dots = \lambda_n = 0$, i.e. S is linearly independent.*

Proof. If $\sum_{k=1}^n \lambda_k e_k = 0$, then for $1 \le j \le n$, we have

$$0 = (0 \,|\, e_j) = \Big(\sum_{k=1}^n \lambda_k e_k \,\Big|\, e_j\Big) = \sum_{k=1} \lambda_k (e_k \,|\, e_j) = \lambda_j$$

since S is orthonormal. □

Theorem 4.24 (Gram–Schmidt procedure)**.** *Let $(x_k)_{k\in\mathbb{N}}$ be a linearly independent sequence in an inner product space H. Then there exists an orthonormal system $(e_k)_{k\in\mathbb{N}}$ such that* $\operatorname{span}\{x_1,\dots,x_n\} = \operatorname{span}\{e_1,\dots,e_n\}$ *for all $n \in \mathbb{N}$. An analogous result holds for finite linearly independent sets.*

The iterative argument used in the Gram–Schidt procedure can be understood in the following way. Let $x, e_n \in H$ and $\|e_n\| = 1$. Then $(x \,|\, e_n)$ measures 'how much e_n there is in x'. Then $y := x - (x \,|\, e_n)e_n$ is what remains of x if 'all e_n is removed'. Note that $(y \,|\, e_n) = 0$, so 'no e_n is left in y'.

Proof. Since (x_k) is independent, $x_1 \neq 0$. We may thus define $e_1 := \|x_1\|^{-1} x_1$. Then $\|e_1\| = 1$ and $\operatorname{span}\{x_1\} = \operatorname{span}\{e_1\}$.

Now suppose that we have already constructed orthonormal vectors $e_1,\dots,e_n$ with $\operatorname{span}\{x_1,\dots,x_k\} = \operatorname{span}\{e_1,\dots,e_k\}$ for all $1 \le k \le n$. Put $\tilde{e}_{n+1} := x_{n+1} - \sum_{j=1}^n (x_{n+1} \,|\, e_j) e_j$. Since $x_1,\dots,x_{n+1}$ are linearly independent, $\tilde{e}_{n+1} \neq 0$. Hence, we may define $e_{n+1} := \|\tilde{e}_{n+1}\|^{-1} \tilde{e}_{n+1}$.

By construction, $\|e_{n+1}\| = 1$ and

$$\begin{aligned}(e_{n+1} \,|\, e_k) &= \|\tilde{e}_{n+1}\|^{-1} \Big((x_{n+1} \,|\, e_k) - \sum_{j=1}^n (x_{n+1} \,|\, e_j)(e_j \,|\, e_k) \Big) \\ &= \|\tilde{e}_{n+1}\|^{-1} \big((x_{n+1} \,|\, e_k) - (x_{n+1} \,|\, e_k) \big) = 0.\end{aligned}$$

Moreover, $\operatorname{span}\{x_1,\dots,x_{n+1}\} = \operatorname{span}\{e_1,\dots,e_{n+1}\}$. Proceeding by induction, the claim follows. □

Corollary 4.25. *Every separable Hilbert space has a countable orthonormal basis.*

Proof. If $\overline{\operatorname{span}}\, S = H$, then, by linear algebra, we may pick a linearly independent sequence from $\overline{\operatorname{span}}\, S$ whose span is also dense in H (this sequence may be finite). We then apply the Gram–Schmidt procedure to this sequence to obtain an orthonormal system whose span is dense. □

Example 4.26. Consider the Hilbert space $L^2(-1,1)$. We apply the Gram–Schmidt procedure to the linearly independent monomials $f_j(t) = t^j$ for $j = 1,2,3$.

We have $\|f_0\|^2 := \int_{-1}^{1} 1 \,\mathrm{d}t = 2$. We hence set $e_0(t) = \frac{1}{\sqrt{2}}$. Next observe that

$$(f_1 \,|\, e_0) = \frac{1}{\sqrt{2}} \int_{-1}^{1} t \,\mathrm{d}t = \frac{1}{\sqrt{2}} \Big[\frac{1}{2} t^2\Big|_{-1}^{1} = 0.$$

Hence $\tilde{e}_1 = f_1 - 0e_0$. Since $\|\tilde{e}_1\|^2 = \int_{-1}^{1} t^2 \,\mathrm{d}t = \frac{2}{3}$, we set $e_1(t) = \frac{\sqrt{3}}{\sqrt{2}} t$. As for f_2, we have

$$(f_2 \,|\, e_0) = \int_{-1}^{1} \frac{1}{\sqrt{2}} t^2 \,\mathrm{d}t = \frac{\sqrt{2}}{3}$$

and

$$(f_2 \,|\, e_1) = \int_{-1}^{1} \frac{\sqrt{3}}{\sqrt{2}} t^3 \,\mathrm{d}t = 0.$$

Thus

$$\tilde{e}_2 = f_2 - 0 \cdot e_1 - \frac{\sqrt{2}}{3} e_0 = t^2 - \frac{1}{3}.$$

Since $\|\tilde{e}_2\|^2 = \int_{-1}^{1} (t^2 - \frac{1}{3})^2 \,\mathrm{d}t = \frac{2}{5}$, we obtain $e_2(t) = \sqrt{\frac{5}{2}}(t^2 - \frac{1}{3})$.

Corollary 4.27 (Orthogonal projection onto a finite dimensional space)**.** *Let H be a Hilbert space and $\{e_1, \dots, e_n\}$ be an orthonormal system. Then, for $K :=$ span$\{e_1, \dots, e_n\}$, the orthogonal projection P_K onto K is given by*

$$P_K x = \sum_{j=1}^{n} (x \,|\, e_j) e_j.$$

Proof. In the proof of Theorem 4.24, it was seen that $x - \sum_{j=1}^{n} (x \,|\, e_j) e_j \perp e_k$ for $k = 1, \dots, n$ and thus $x - \sum_{j=1}^{n} (x \,|\, e_j) e_j \in \operatorname{span}\{e_1, \dots, e_n\}^{\perp}$. By Proposition 4.16, $P_K x = \sum_{j=1}^{n} (x \,|\, e_j) e_j$. □

Lemma 4.28 (Bessel's inequality)**.** *Let H be a Hilbert space and $(e_n)_{n \in \mathbb{N}}$ be an orthonormal system. Then, for every $x \in H$,*

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

Proof. For $N \in \mathbb{N}$, let $K_N := \operatorname{span}\{e_1, \ldots, e_N\}$ and P_N be the orthogonal projection onto K_N. By Corollary 4.27, $P_N x = \sum_{k=1}^{N} (x \mid e_k) e_k$. Now Pythagoras' theorem yields

$$\|x\|^2 = \|P_N x\|^2 + \|x - P_N x\|^2 \geq \|P_N x\|^2 = \sum_{k=1}^{N} |(x \mid e_k)|^2.$$

Since N was arbitrary, the claim follows. □

We now can describe also orthogonal projections onto infinite-dimensional subspaces.

Proposition 4.29. *Let H be a Hilbert space and K be a closed linear subspace of H. If $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of K, then the series $\sum_{n=1}^{\infty} (x \mid e_n) e_n$ converges for every $x \in H$ and the orthogonal projection onto K is given by*

$$P_K x := \sum_{n=1}^{\infty} (x \mid e_n) e_n.$$

Proof. For $1 \leq m \leq n$, we have by Pythagoras' theorem

$$\left\| \sum_{k=1}^{n} (x \mid e_k) e_k - \sum_{k=1}^{m} (x \mid e_k) e_k \right\|^2 = \sum_{k=m+1}^{n} |(x \mid e_k)|^2.$$

As a consequence of Bessel's inequality, the latter converges to 0 as $n, m \to \infty$. Thus the elements $\sum_{k=1}^{n} (x \mid e_k) e_k$ form a Cauchy sequence in H in n and is therefore convergent. This proves the first assertion.

For the second assertion, note that

$$y = \sum_{k=1}^{\infty} (x \mid e_k) e_k \in K = \overline{\operatorname{span}} \{e_k : k \in \mathbb{N}\}$$

and $x - y \perp e_j$ for all $j \in \mathbb{N}$, which follows from the fact that $x - \sum_{k=1}^{N} (x \mid e_k) e_k \perp e_j$ (see Corollary 4.27) and the continuity of the inner product. □

Theorem 4.30 gives a useful characterisation of an orthonormal basis in a separable Hilbert space. It actually generalises to non-separable Hilbert spaces, but notation becomes more delicate as a basis will not be countable in that case. In combination with Bessel's inequality, the Parseval identity expresses that an orthonormal system is a basis if and only if it 'exhausts the norm'.

Theorem 4.30. *Let H be a separable Hilbert space, $(e_n)_{n\in\mathbb{N}}$ be an orthonormal system. The following are equivalent.*

(a) *$(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of H.*

(b) *$x \perp e_n$ for all $n \in \mathbb{N}$ implies that $x = 0$.*

(c) *$x = \sum_{k=1}^{\infty} (x \mid e_k) e_k$ for all $x \in H$.*

(d) *$(x \mid y) = \sum_{k=1}^{\infty} (x \mid e_k)(e_k \mid y)$ for all $x, y \in H$.*

(e) *(Parseval's identity) For all $x \in H$ we have*

$$\|x\|^2 = \sum_{k=1}^{\infty} |(x \mid e_k)|^2.$$

Proof. (a) ⇒ (b): If $x \perp e_n$ for all $n \in \mathbb{N}$, then $x \perp \overline{\operatorname{span}}\{e_n : n \in \mathbb{N}\}$. If $\overline{\operatorname{span}}\{e_n : n \in \mathbb{N}\} = H$, it follows that $x \perp x$ and thus $x = 0$ by (IP1).

(b) ⇒ (c): We have $x - \sum_{k=1}^{\infty} (x \,|\, e_k) e_k \perp e_n$ for all $n \in \mathbb{N}$. Thus, by (b) $x = \sum_{k=1}^{\infty} (x \,|\, e_k) e_k$.

(c) ⇒ (d): By (c),

$$\begin{aligned}(x \,|\, y) &= \Big(\sum_{n=1}^{\infty} (x \,|\, e_n) e_n, \sum_{m=1}^{\infty} (y \,|\, e_m) e_n \Big) \\ &= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (x \,|\, e_n) \overline{(y \,|\, e_m)} (e_n \,|\, e_m) \\ &= \sum_{n=1}^{\infty} (x \,|\, e_n)(e_n \,|\, y),\end{aligned}$$

where we used the continuity of the inner product in the second step and the orthonormality of the e_n in the third.

(d) ⇒ (e): Put $x = y$.

(e) ⇒ (a): If $x \in \operatorname{span}\{e_n : n \in \mathbb{N}\}^{\perp}$, then $\|x\|^2 = 0$ by Parseval's identity. Thus $\operatorname{span}\{e_n : n \in \mathbb{N}\}^{\perp} = \{0\}$ and hence $H = \{0\}^{\perp} = (\operatorname{span}\{e_n : n \in \mathbb{N}\}^{\perp})^{\perp} = \overline{\operatorname{span}}\{e_n : n \in \mathbb{N}\}$. □

We rephrase Theorem 4.30 in a slightly different way.

Corollary 4.31. *Let H be a separable Hilbert space and $(e_n)_{n \in \mathbb{N}}$ be an orthonormal basis of H. Then $U\colon H \to \ell^2$, defined by*

$$Ux = ((x \,|\, e_n))_{n \in \mathbb{N}},$$

is an isometric isomorphism.

An isometric isomorphism between normed spaces preserves not only the topological and linear structure, it also preserves properties of the norm, like whether the parallelogram identity is satisfied or not, for example.

Proof. Linearity of U is clear. That it is isometric follows from Theorem 4.30 (e). It thus only remains to show that U is surjective. To that end, given $a \in \ell^2$, put $x = \sum_{k=1}^{\infty} a_k e_k$. This series converges in H. Indeed, putting $x_N := \sum_{k=1}^{N} a_k e_k$, we have, for $N > M$, that $\|x_N - x_M\|_H = \sum_{k=M+1}^{N} |a_k|^2 \to 0$ as $M \to \infty$. Hence (x_N) is a Cauchy sequence and therefore convergent. Clearly, $Ux = a$, proving surjectivity. □

Remark 4.32. Corollary 4.31 shows that by fixing an orthonormal basis of H we can identify H with ℓ^2. Hence we can translate problems in an arbitrary (infinite dimensional) separable Hilbert space H into equivalent problems in ℓ^2, provided that they only depend on the Hilbert space structure. While this is conceptually important, it tends to be not very useful practically since one usually uses specific properties of the elements as well, for example, pointwise or measure related properties in $L^2(\Omega, \Sigma, \mu)$.

Definition 4.33. Let H be an infinite dimensional, separable Hilbert space and (e_n) be an orthonormal basis of H. For $x \in H$, the series

$$\sum_{k=1}^{\infty} (x \,|\, e_k) e_k$$

is called the **Fourier series** of x with respect to (e_n). The vector $((x \,|\, e_k))_{k\in\mathbb{N}} \in \ell^2$ is called the **Fourier coefficients** of x.

Exercise 4.34. Classical Fourier series arise by considering the Hilbert space $L^2(0, 2\pi)$, for convenience endowed with the inner product

$$(f \,|\, g) := \frac{1}{2\pi} \int_0^{2\pi} f(t)\overline{g(t)}\,\mathrm{d}t,$$

and the orthonormal basis $e_k\colon t \mapsto e^{ikt}$ for $k \in \mathbb{Z}$. Then, for $u \in L^2(0, 2\pi)$ the number $\hat{u}(k) := (u \,|\, e_k)$ is called the k-th Fourier coefficient.

(a) Show that if u is continuously differentiable with $u(0) = u(2\pi)$, then $\widehat{u'}(k) = ik\hat{u}(k)$.

(b) Compute the Fourier coefficients of $u\colon t \mapsto \frac{1}{4}(\pi - t)^2$.

(c) Show that

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