Basic Hilbert space theory

Johannes Mueller

November 15, 2017

Characterisation of Hilbert spaces

In the following we will show that all Hilbert spaces are isomorphic to a space of square integrable functions which we can use later as we then can always work in those nice and explicit spaces. Further we will also obtain a generalised Fourier series in Hilbert spaces. In this section it will be usefull to use some uncountable sums if at most countably many summands are non zero. More precisely we will write

$$x = \sum_{i \in I} x_i$$

if $J \coloneqq \{i \in I \mid x_i \neq 0\}$ is countable and for every numeration (j_n) of J we have

$$x = \sum_{n \in \mathbb{N}} x_{j_n} \coloneqq \lim_{N \to \infty} \sum_{n=1}^{N} x_{j_n}.$$

1 Definition (Orthonormal set)**.** A subset $\{e_i\}_{i \in I}$ of a Hilbert space H is called *orthonormal set*, or short *ONS*, if we have $\langle e_i, e_j \rangle = \delta_{ij}$. Further it is called *complete*, or short *CONS*, if the linear span is dense in H.

2 Remark. The concept of complete orthonormal sets takes the role of the classical orthonormal basis in the case of infinite dimensional Hilbert spaces as there is never an orthonormal basis in this case. To see that, let $\{e_i\}_{i \in I}$ be an infinite orthonormal set. Then we can take a sequence $(e_{i_n})_{n \in \mathbb{N}}$ with $i_n \neq i_m$ for $n \neq m$ and now the element $\sum_n \frac{1}{n} e_{i_n} \in H$ can not be expressed as a finite linear combination of elements from the orthonormal set. Thus an orthonormal set can never be a basis in infinite dimensions.

3 Proposition (Existence of CONS)**.** *Every Hilbert space has a complete orthonormal set.*

To prove this fundamental statement we will need two auxiliary results.

4 Proposition (Bessels inequality)**.** *Let $\{e_n\}_{n \in \mathbb{N}}$ be an orthonormal set and let $x \in H$. Then we have*

$$\sum_{n \in \mathbb{N}} |\langle x, e_n \rangle|^2 \leq \|x\|^2.$$

Proof. For $N \in \mathbb{N}$ we set $x_N := \sum_{n=1}^{N} \langle x, e_n \rangle e_n$. Using Pythagoras theorem and $x_N \perp x - x_N$ we get

$$\|x\|^2 = \|x - x_N\|^2 + \|x_N\|^2 \geq \|x_N\|^2 = \sum_{n=1}^{N} |\langle x, e_n \rangle|^2 .$$

As this holds for all N we can pass to the limit. □

5 Corollary. *Let $\{e_i\}_{i\in I}$ be an orthonormal set. Then the set*

$$J_x := \{ i \in I \mid \langle x, e_i \rangle \neq 0 \}$$

is at most countable for all $x \in H$.

Proof. With the usage of Bessels inequality we immediately get that

$$\left\{ i \in I \,\middle|\, |\langle x, e_i \rangle| \geq \frac{1}{n} \right\}$$

is finite for all $n \in \mathbb{N}$ which yields the assertion. □

Proof of proposition 3. Zorns Lemma yields the existence of a wrt inclusion maximal orthonormal set $\{e_i\}_{i\in I}$. If the linear span of the orthonormal set would not be dense, then we claim the existence of a nontrivial, normed element e in the orthogonal complement and therefore the set would not be maximal as we could add e to it. Thus $\{e_i\}_{i\in I}$ has to be a complete orthonormal set. To complete the proof we only have to show the existence such an element e. To do so, set $U := \overline{\operatorname{span}\{e_i\}_{i\in I}}$ and choose $x \notin U$. Further define J_x as above and take a numeration (j_n) of J_x. Then we get using Bessels inequality that

$$e := x - \sum_{n\in\mathbb{N}} \langle x, e_{j_n} \rangle e_{j_n} \in H$$

exists. The element e is non trivial as $e = 0$ would imply

$$x = \sum_{n\in\mathbb{N}} \langle x, e_{j_n} \rangle e_{j_n} \in U$$

which is a contradiction. Thus we can wlog assume e to be normed and further we have $\langle e, e_j \rangle = 0$ for all $j \in J_x$ and $\langle e, e_i \rangle = 0$ for all $i \in I \setminus J_x$ an therefore we can conclude $e \in U^{\perp}$. □

6 Lemma. *Let I be an arbitrary set and let $L^2(I)$ be the Hilbert space induced by the counting measure $\#$. Further χ_i denotes the characteristic function of $\{i\}$. Then $\{\chi_i\}_{i\in I}$ is a complete orthonormal set and we have*

$$v = \sum_{i\in I} v(i)\chi_i \tag{1}$$

in which for a fixed v at most countably many summands are non zero.

Proof. It is obvious that $\{\chi_i\}_{i\in I}$ is an orthonormal set, and if we can prove (1) the completeness follows. A function $v \in L^2(I)$ vanishes outside of a countable set, because if $\{|v| \geq \varepsilon\} \subseteq I$ would be infinite for some $\varepsilon > 0$, v would not be square integrable wrt the counting measure. For a fixed v let $J := \{v \neq 0\} = \{i_n\}_{n\in\mathbb{N}}$, then we have

$$\left\| v - \sum_{n=1}^{N} v(i_n)\chi_{i_n} \right\|_{L^2(I)}^2 = \int \chi_{J\setminus\{i_1,\ldots,i_N\}} |v|^2 \, \mathrm{d}\# \xrightarrow{N\to\infty} 0$$

by dominated convergence, which shows the assertion. □

7 Theorem (Characterisation of Hilbert spaces and Fourier series)**.** *Let H be a Hilbert space and let $\{e_i\}_{i\in I} \subseteq H$ be a complete orthonormal set. Then there is an isometric isomorphism $\Phi\colon L^2(I) \to H$ such that $\chi_i \mapsto e_i$. Further the inverse is given by $\hat{v}(i) := \left(\Phi^{-1}v\right)(i) = \langle v, e_i\rangle$ and we have*

$$v = \sum_{i\in I} \langle v, e_i\rangle \, e_i \tag{2}$$

and for a fixed v at most countably many summands are non zero. Moreover the scalar product can be calculated by

$$\langle v, w\rangle = \sum_{i\in I} \langle v, e_i\rangle \, \overline{\langle w, e_i\rangle} \tag{3}$$

and in particular

$$\|v\|^2 = \sum_{i\in I} |\langle v, e_i\rangle|^2 .$$

Proof. First we convince ourselves that $\chi_i \mapsto e_i$ defines a linear isometry from the span of $\{\chi_i\}_{i\in I}$ onto its range which can uniquely be extended to an isometry Φ defined on the whole space $L^2(I)$. In particular the range is closed in H but further it is also dense, as the span of $\{e_i\}_{i\in I}$ is contained in the range. Therefore Φ is an isometric isomorphism from $L^2(I)$ onto H and thus unitary by using the polarisation identity. This yields

$$\langle v, e_i\rangle = \left\langle \Phi^{-1}v, \Phi^{-1}e_i \right\rangle = \langle \hat{v}, \chi_i\rangle = \hat{v}(i).$$

By applying the previous lemma we get

$$v = \Phi(\hat{v}) = \Phi\left(\sum_{i\in I} \hat{v}(i)\chi_i\right) = \sum_{i\in I} \hat{v}(i)\Phi\chi_i = \sum_{i\in I} \langle v, e_i\rangle \, e_i$$

where we used the continuity of Φ to swap it with the countable sum. Finally we have

$$\langle v, w\rangle = \langle \hat{v}, \hat{w}\rangle = \sum_{i\in I} \hat{v}(i)\overline{\hat{w}(i)} = \sum_{i\in I} \langle v, e_i\rangle \, \overline{\langle w, e_i\rangle}.$$

□

8 Remark. The formula (4) is a direct generalisation of the Fourier expansion in abstract Hilbert spaces. To obtain the classical Fourier expansion one only has to prove that $\left\{(2\pi)^{-\frac{1}{2}} e^{inx}\right\}_{n\in\mathbb{Z}}$ is a complete orthonormal set in $L^2([0, 2\pi])$.

9 Proposition. *For a Hilbert space H the following statements are equivalent.*

(i) *H is separable.*

(ii) *There is an at most countable complete orthonormal set $\{e_n\}$ of H.*

(iii) *Every orthonormal set in H is at most countable.*

Proof. $(iii) \Rightarrow (ii)$ is obvious because every Hilbert space has a complete orthonormal set.

$(ii) \Rightarrow (i)$ is also straight forward as the linear combinations of $\{e_n\}$ with rational coefficients are countable and dense in H.

$(i) \Rightarrow (iii)$: If there would be an uncountable orthonormal set, then it would be an uncountable discrete set in H and therefore H could not be separable. □

10 Corollary. *Every separable Hilbert space is isometrically isomorphic to the space of square summable sequences $l^2(\mathbb{K}) \cong L^2(\mathbb{N})$.*

Proof. Combine Theorem 7 and Proposition 9. □

Orthogonal projection and sesquilinear forms

In this section we will investigate some further properties of the structures of Hilbert spaces. First we will prove that for a closed subspace of a Hilbert space we can find an orthogonal projection in analogy to the finite dimensional case. Then we will show that the dual space of a Hilbert space is isometric to the Hilbert space itself and will characterise all continuous sesquilinear forms.

11 Theorem (Orthogonal projection)**.** *Let $U \subseteq H$ be a closed subspace of a Hilbert space H. Then there exists a linear contraction $P\colon H \to U$ such that $x - Px \in U^\perp$ and further the mapping*

$$H \to U \times U^\perp, \quad x \mapsto (Px, x - Px)$$

is a linear isometric isomorphism. In other words, H is the isometric direct sum of the two closed subspaces U and $U^\perp$, i.e. every element $x \in H$ can be uniquely written as the sum $u + v$ with $u \in U, v \in U^\perp$ and $\|x\|^2 = \|u\|^2 + \|v\|^2$.

Proof. We will only show the existence of the mapping P as the other assertions then follow immediatly.

Let $\{e_j\}_{j\in J} \subseteq U$ be a complete orthonormal set in U. Then there is an complete orthonormal set $\{e_i\}_{i\in I} \subseteq H$ with $J \subseteq I$. Set now

$$Px := \sum_{j\in J} \langle x, e_j\rangle\, e_j \quad \text{for } x \in H.$$

This sum is well defined, as we have $(\langle x, e_j\rangle)_{j\in J} \in L^2(J)$ and as U is closed we get $Px \in U$. With

$$x - Px = \sum_{i\in I\setminus J} \langle x, e_i\rangle\, e_i$$

and (3) we get $x - Px \in U^\perp$. □

12 Remark. With the above notation we get

$$\langle x, u \rangle = \langle x - Px, u \rangle + \langle Px, u \rangle = \langle Px, u \rangle \quad \text{for all } x \in X, u \in U$$

which is sometimes used to define the term projection.

13 Example (Conditional expectation as a projection)**.** We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, that is a measurable space with a normed measure, i.e. $\mathbb{P}(\Omega) = 1$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ field, then $L^2(\mathcal{G})$ is complete and therefore a closed subspace of $L^2(\mathcal{F})$. Let P be the projection onto $L^2(\mathcal{G})$, then we get

$$\langle f, g \rangle_{L^2(\mathcal{F})} = \langle Pf, g \rangle_{L^2(\mathcal{F})} \quad \text{for all } g \in L^2(\mathcal{G}).$$

In particular we have for all $A \in \mathcal{G}$

$$\int f \cdot \chi_A \mathrm{d}\mathbb{P} = \int Pf \cdot \chi_A \mathrm{d}\mathbb{P} \tag{4}$$

which is the just the definition of the *conditional expectation* of f wrt to $\mathcal{G}$. In stochastics P is written as $\mathbb{E}[\cdot|\mathcal{G}]$ and can be extended to a linear contraction from $L^1(\mathcal{F})$ to $L^1(\mathcal{G})$. Indeed we have for $f \in L^2(\mathcal{F})$

$$\|Pf\|_{L^1(\mathcal{G})} = \langle Pf, \operatorname{sign}(Pf) \rangle_{L^2(\mathcal{F})} = \langle f, \operatorname{sign}(Pf) \rangle_{L^2(\mathcal{F})} \leq \|f\|_{L^1(\mathcal{F})}$$

and since $L^2(\mathcal{F}) \subseteq L^1(\mathcal{F})$ is dense the assertion follows and (4) extends for all $f \in L^1(\mathcal{F})$.

14 Theorem (Riesz representation theorem)**.** *Let H be a Hilbert space and $f \in H'$. Then there is a unique element $x_f \in H$ such that $f(y) = \langle y, x_f \rangle$ for all $y \in H$. Further the mapping*

$$\Phi_H \colon H' \to H, \quad f \mapsto x_f$$

is bijective, isometric and conjugate linear, i.e. $\Phi_H(\lambda f) = \overline{\lambda} \Phi_H f$ for all $f \in H'$ and $\lambda \in \mathbb{K}$.

Proof. Assume wlog that $f \neq 0$, then we find a non trivial element $\hat{x} \in \ker(f)^\perp$ such that $f(\hat{x}) = 1$ as $\ker(f)$ is a closed but strict subspace. For $y \in H$ we have $y - f(y)\hat{x} \in \ker(f)$ and therefore

$$\langle y - f(y)\hat{x}, \hat{x} \rangle = 0.$$

This implies $\langle y, \hat{x} \rangle = f(y) \|\hat{x}\|^2$ and therefore $x_f := \frac{\hat{x}}{\|\hat{x}\|^2}$ fulfills the assertion.

The conjugate linearity follows directly from the conjugate linearity of the second component of the scalar product. Further the isometry property follows from

$$\|f\| = \sup_{\|y\|=1} |f(y)| = \sup_{\|y\|=1} |\langle y, x_f \rangle| = \|x_f\|.$$

This computation also shows that $\langle \cdot, x \rangle \in H'$ for all $x \in H$ which directly implies the bijectivity as this is simply the inverse mapping. □

15 Corollary (Adjoint mapping)**.** *Let H_1, H_2 be Hilbert spaces and $A \in \mathcal{L}(H_1, H_2)$. Then there is a unique mapping $A^* \in L(H_2, H_1)$, called the adjoint mapping of A, such that*

$$\langle Ax, y \rangle_{H_2} = \langle x, A^* y \rangle_{H_1} \quad \textit{for all } x \in H_1, y \in H_2.$$

Proof. Set $A^* := \Phi_{H_1} A' \Phi_{H_2}^{-1}$ and calculate the stated properties. The uniqueness is obvious, because A^*y is uniquely determined by the values $\langle x, A^*y\rangle$. $\square$

16 Definition (Bounded sesquilinear forms)**.** Let X, Y be normed spaces and $B\colon X \times Y \to \mathbb{K}$ be a sesquilinear form. Then we call B *bounded* if

$$\|B\| := \sup_{\|x\|_X \leq 1, \|y\|_Y \leq 1} |B(x,y)| < \infty$$

holds.

17 Proposition (Sesquilienar forms)**.** *Let X, Y be normed spaces and let B be a bounded sesquilinear form on $X \times Y$. Then there is a unique conjugate linear bounded mapping $A = A_B\colon Y \to X'$ such that*

$$\big(Ay\big)(x) = B(x,y) \quad \textit{for all } x \in X, y \in Y. \tag{5}$$

Further we have $\|A\| = \|B\|$ and thus there is an isometric bijection

$$\Big\{B \mid B \textit{ sesquilinear form on } X \times Y\Big\} \to L(Y, X'), \quad B \mapsto A_B.$$

Proof. The operator A_B has to satisfy (5), so we can simply define it via (5). The computation

$$\sup_{\|x\| \leq 1} \big(Ay\big)(x) = \sup_{\|x\| \leq 1} B(x,y) \leq \|B\| \, \|y\|$$

shows that Ay indeed is a dual element. The conjugate linearity of A_B follows from the conjugate linearity of the second component of B and the isometry property with

$$\|A_B\| = \sup_{\|y\| \leq 1} \sup_{\|x\| \leq 1} \big|\big(Ay\big)(x)\big| = \sup_{\|y\| \leq 1} \sup_{\|x\| \leq 1} |B(x,y)| = \|B\| \, .$$

Finally the bijectivity of $B \mapsto A_B$ follows as the inverse is given by

$$A \mapsto \Big((x,y) \mapsto \big(Ay\big)(x)\Big).$$

$\square$

18 Theorem (Characterisation of Sesquilinear forms)**.** *Let B be a bounded sesquilinear form on the Hilbert space H. Then there is a unique mapping $A = A_B \in L(H)$ such that*

$$B(x,y) = \langle x, Ay\rangle \quad \textit{for all } x \in H \tag{6}$$

Hence there is a canonical isometric bijection

$$\Big\{B \mid B \textit{ sesquilinear form on } H\Big\} \to L(H), \quad B \mapsto A_B.$$

Further A is self adjoint if and only if $B(x,y) = \overline{B(y,x)}$ for all $x, y \in H$.

Proof. The first part of the assertion follows directly from the previous proposition, if we take the pointwise composition of the conjugate linear mapping A_B with the Riesz isometry Φ_H. Finally A is selfadjoint if and only if we have

$$B(x,y) = \langle x, Ay\rangle = \langle Ax, y\rangle = \overline{\langle y, Ax\rangle} = \overline{B(y,x)} \quad \text{for all } x, y \in H.$$

□

Now we want to give a sufficient condition for B such that the induced operator A_B is an isomorphism.

19 Definition (Coercivity)**.** A sesquilinear form B on a Banach space X is called *coercive* if we have

$$\lfloor B \rfloor := \inf_{\|x\|_X = 1} \operatorname{Re}\big(B(x,x)\big) > 0.$$

20 Corollary (Lax-Milgram)**.** *Let B be a coercive sesquilinear form on the Hilbert space H. Then the operator $A = A_B$ fulfilling*

$$B(x,y) = \langle x, Ay\rangle \quad \textit{for all } x \in H$$

is an isomorphism with $\left\|A^{-1}\right\| \le \lfloor B \rfloor^{-1}$.

Proof. Obviously the coercivity of B implies the injectivity of A. To see that the range of A is closed, take $Ax_n \to y$, then we have

$$\|x_n - x_m\|^2 \le \lfloor B \rfloor \left|B(x_n - x_m, x_n - x_m)\right| \le \lfloor B \rfloor \|Ax_n - Ax_m\| \|x_n - x_m\| \tag{7}$$

and therefore (x_n) is a Cauchy sequence with limit x. Now we have $y = Ax$ as A is continuous and therefore the range of A is closed. If the range would not be dense, we would find an element $x \in \operatorname{ran}(A)^\perp$ which would fulfill

$$\langle x, Ax\rangle = B(x,x) = 0$$

which is absurd. Thus the range is dense and with the closedness we have $\operatorname{ran}(A) = X$. Finally (7) also shows the boundedness of the inverse mapping as well as $\left\|A^{-1}\right\| \le \lfloor B \rfloor^{-1}$. □