Chapter 2
Rudiments of Hilbert Space Theory

As the present work is about Hilbert space quantum mechanics, it is mandatory that the reader has sufficient grounding in Hilbert space theory. This short chapter is designed to indicate what sort of basic equipment one needs in the ensuing more sophisticated chapters. At the same time it can be used as an introduction to elementary Hilbert space theory even for the novice. The material is quite standard and appears of course in numerous works, so we do not explicitly specify any references, though some source material can be found in the bibliography.

2.1 Basic Notions and the Projection Theorem

We begin with a key definition. Unless otherwise stated, all vector spaces in this work have the field $\mathbb{C}$ of complex numbers as their field of scalars. We denote by $\mathbb{N}$ the set of positive integers, i.e. $\mathbb{N} = \{1, 2, 3, \ldots\}$, and let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Definition 2.1 Let E be a (complex) vector space. We say that a mapping $h : E \times E \to \mathbb{C}$ is an *inner product* (in E) and E (equipped with h) is an *inner product space*, if for all $\varphi, \psi, \eta \in E$ and $\alpha, \beta \in \mathbb{C}$ we have

(IP1) $h(\varphi, \alpha\psi + \beta\eta) = \alpha h(\varphi, \psi) + \beta h(\varphi, \eta)$,
(IP2) $h(\varphi, \psi) = \overline{h(\psi, \varphi)}$,
(IP3) $h(\varphi, \varphi) \geq 0$,
(IP4) $h(\varphi, \varphi) > 0$ if $\varphi \neq 0$.

Unless otherwise stated, in the sequel we write $h(\varphi, \psi) = \langle\, \varphi \,|\, \psi \,\rangle$ in any context described by this definition.

In (b) below there is the *Cauchy–Schwarz inequality*.

P. Busch et al., *Quantum Measurement*, Theoretical and Mathematical Physics,
DOI 10.1007/978-3-319-43389-9_2

Theorem 2.1 *Assume that E is an inner product space, $\varphi, \psi, \eta \in E$, $\alpha, \beta \in \mathbb{C}$. Then*

(a) $\langle \alpha\varphi + \beta\psi \,|\, \eta \rangle = \overline{\alpha} \langle \varphi \,|\, \eta \rangle + \overline{\beta} \langle \psi \,|\, \eta \rangle$;
(b) $|\langle \varphi \,|\, \psi \rangle|^2 \le \langle \varphi \,|\, \varphi \rangle \langle \psi \,|\, \psi \rangle$.

Proof Part (a) is an immediate consequence of the definition. To prove (b), note that for any $\alpha \in \mathbb{C}$ we have

$$0 \le \langle \alpha\varphi + \psi \,|\, \alpha\varphi + \psi \rangle = |\alpha|^2 \langle \varphi \,|\, \varphi \rangle + \overline{\alpha} \langle \varphi \,|\, \psi \rangle + \alpha \langle \psi \,|\, \varphi \rangle + \langle \psi \,|\, \psi \rangle .$$

If $\langle \varphi \,|\, \varphi \rangle \neq 0$, choose $\alpha = -\langle \varphi \,|\, \psi \rangle \langle \varphi \,|\, \varphi \rangle^{-1}$, and then

$$|\langle \varphi \,|\, \psi \rangle|^2 \langle \varphi \,|\, \varphi \rangle^{-1} - |\langle \varphi \,|\, \psi \rangle|^2 \langle \varphi \,|\, \varphi \rangle^{-1} - |\langle \varphi \,|\, \psi \rangle|^2 \langle \varphi \,|\, \varphi \rangle^{-1} + \langle \psi \,|\, \psi \rangle \ge 0,$$

which implies the claim. If $\langle \varphi \,|\, \varphi \rangle = 0$, by multiplying φ with a suitable complex number of modulus one we may assume that $\langle \varphi \,|\, \psi \rangle$ is real, and then it is easy to see that the above inequality can be true for all $\alpha \in \mathbb{R}$ only if $\langle \varphi \,|\, \psi \rangle = 0$. □

Remark 2.1 If (IP1) and (a) above hold for h, then the map h is called *sesquilinear*. If, moreover, (IP3) holds, it is a *positive sesquilinear form*. If h is sesquilinear, then $h(\varphi, \psi) = \overline{h(\psi, \varphi)}$ for all $\varphi, \psi \in \mathcal{H}$ if and only if $h(\varphi, \varphi) \in \mathbb{R}$ for all $\varphi \in \mathcal{H}$. Indeed, the "only if" part is obvious, and to prove the "if" part, write

$$h(\alpha\varphi + \beta\psi, \alpha\varphi + \beta\psi) = |\alpha|^2 h(\varphi, \varphi) + \overline{\alpha}\beta h(\varphi, \psi) + \overline{\beta}\alpha h(\psi, \varphi) + |\beta|^2 h(\psi, \psi)$$

and substitute first $\alpha = \beta = 1$ and then $\alpha = 1$, $\beta = i$ to see that $\operatorname{Im} h(\varphi, \psi) = -\operatorname{Im} h(\psi, \varphi)$ and $\operatorname{Re} h(\varphi, \psi) = \operatorname{Re} h(\psi, \varphi)$. In particular, positive sesquilinear forms automatically satisfy (IP2), and the proof we gave for the Cauchy–Schwarz inequality was so formulated that it is valid without assuming (IP4). This generality will be needed later. ◁

The proof of the next result is an easy exercise.

Theorem 2.2 *Let E be an inner product space. Denote $\|\varphi\| = \sqrt{\langle \varphi \,|\, \varphi \rangle}$, when $\varphi \in E$. Then*

(a) $\|\varphi\| \ge 0$ *for all* $\varphi \in E$;
(b) $\|\varphi\| = 0$, *if and only if* $\varphi = 0$;
(c) $\|\alpha\varphi\| = |\alpha| \|\varphi\|$ *for each* $\alpha \in \mathbb{C}$ *and* $\varphi \in E$;
(d) $\|\varphi + \psi\| \le \|\varphi\| + \|\psi\|$ *for all* $\varphi, \psi \in E$.

Remark 2.2 The preceding result means that the map $\varphi \mapsto \sqrt{(\varphi|\varphi)}$ is a *norm* on E. Unless otherwise stated, an inner product space will be equipped with this norm. ◁

Example 2.1 The set $\mathbb{C}^n = \{x = (x_1, \ldots, x_n) \,|\, x_k \in \mathbb{C},\ k = 1, \ldots, n\}$ (where $n \in \mathbb{N}$) is an inner product space with respect to its usual operations $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ and $\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)$ and the inner product $\langle x \mid y \rangle = \sum_{k=1}^{n} \overline{x}_k y_k, x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$. The Cauchy-Schwarz inequality acquires the form

$$\Big| \sum_{k=1}^{n} \overline{x}_k y_k \Big| \le \left(\sum_{k=1}^{n} |x_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} |y_k|^2 \right)^{\frac{1}{2}}.$$

$\triangleleft$

Example 2.2 Denote $\ell^2 = \ell^2_{\mathbb{C}} = \left\{ f : \mathbb{N} \to \mathbb{C} \,\middle|\, \sum_{n=1}^{\infty} |f(n)|^2 < \infty \right\}$ and define as usual $(\alpha f)(n) = \alpha(f(n))$ and $(f + g)(n) = f(n) + g(n)$. Clearly $\alpha f \in \ell^2$ if $\alpha \in \mathbb{C}$ and $f \in \ell^2$. Since $|f(n)g(n)| \le \frac{1}{2}[|f(n)|^2 + |g(n)|^2]$, we have

$$\begin{aligned} |f(n) + g(n)|^2 &\le (|f(n)| + |g(n)|)^2 = |f(n)|^2 + 2|f(n)g(n)| + |g(n)|^2 \\ &\le 2\big[|f(n)|^2 + |g(n)|^2\big], \end{aligned}$$

and so $f + g \in \ell^2$ whenever $f, g \in \ell^2$. Thus ℓ^2 is a vector space. Moreover, the series $\sum_{n=1}^{\infty} \overline{f(n)} g(n)$ converges absolutely if $f, g \in \ell^2$. We define

$$\langle f \mid g \rangle = \sum_{n=1}^{\infty} \overline{f(n)} g(n)$$

for $f, g \in \ell^2$. This defines an inner product in ℓ^2, leading to the norm $\|f\| = \|f\|_2 = \left(\sum_{n=1}^{\infty} |f(n)|^2 \right)^{1/2}$. $\triangleleft$

In the next theorem, the equation in (a) is the inner product space version of the *Pythagorean theorem*. The equation in (b) is called the *parallelogram law* and the one in (c) is the *polarisation identity*. All are proved by straightforward calculations.

Theorem 2.3 *Let E be an inner product space.*

(a) *If $\varphi_1, \ldots, \varphi_n \in E$ are vectors satisfying $\left\langle \varphi_i \mid \varphi_j \right\rangle = 0$, whenever $i \neq j$, then*

$$\left\| \sum_{k=1}^{n} \varphi_k \right\|^2 = \sum_{k=1}^{n} \|\varphi_k\|^2.$$

(b) *For all $\varphi, \psi \in E$,*

$$\|\varphi + \psi\|^2 + \|\varphi - \psi\|^2 = 2\|\varphi\|^2 + 2\|\psi\|^2.$$

(c) *If F is any vector space and $B : F \times F \to \mathbb{C}$ a sesquilinear form, then for any $\varphi, \psi \in F$*

$$B(\varphi, \psi) = \frac{1}{4} \sum_{n=0}^{3} i^n B(\psi + i^n \varphi, \psi + i^n \varphi).$$

The polarisation identity, in particular, shows that the norm of an inner product space completely determines the inner product defining it.

Definition 2.2 If the inner product space E is complete with respect to the norm defined by its inner product, i.e. if every Cauchy sequence converges, then E is called a *Hilbert space*.

Example 2.3 The inner product spaces $\mathbb{C}^n$ and ℓ^2 treated in Examples 2.1 and 2.2 are Hilbert spaces. We omit the standard completeness proofs. $\triangleleft$

Unless otherwise stated, throughout the rest of this work we assume that $\mathcal{H}$ is a Hilbert space whose inner product is the mapping $(\varphi, \psi) \mapsto \langle \varphi \,|\, \psi \rangle$. The parallelogram law has a central role in the study of Hilbert space geometry. For example, let $E \neq \emptyset$ be a closed subset of $\mathcal{H}$ and assume that E is also convex (i.e. $t\varphi + (1-t)\psi \in E$ whenever $\varphi, \psi \in E$ and $t \in [0, 1]$). Let d be the infimum of the set $\{\|\varphi\| \mid \varphi \in E\}$. Then there is a sequence (φ_n) in E with $\lim_{n\to\infty} \|\varphi_n\| = d$. The parallelogram law shows that

$$\begin{aligned}\|\varphi_m - \varphi_n\|^2 &= 2\|\varphi_m\|^2 + 2\|\varphi_n\|^2 - 4\left\|\tfrac{1}{2}(\varphi_m + \varphi_n)\right\|^2 \\ &\leq 2\|\varphi_m\|^2 + 2\|\varphi_n\|^2 - 4d^2 \to 0,\end{aligned}$$

when $m, n \to \infty$. Thus (φ_n) is a Cauchy sequence and hence converges to some $\varphi \in E$ (as E is closed), and $\|\varphi\| = \lim_{n\to\infty} \|\varphi_n\| = d$ by the continuity of the norm. Thus φ is an element of E having the smallest possible norm. The parallelogram law can also be applied analogously to the above proof to show that such a φ is uniquely determined.

Suppose now that M is a closed vector subspace of $\mathcal{H}$, $\varphi \in \mathcal{H}$, and $E = \varphi - M$ $(= \{\varphi - \psi \mid \psi \in M\})$. Then, as shown above, in E there is an element $\xi = \varphi - \psi$ having the smallest possible norm. If $\eta \in M$, $\|\eta\| = 1$, the inner product of $\langle \eta \,|\, \xi \rangle \eta$ and $\xi - \langle \eta \,|\, \xi \rangle \eta$ vanishes, and so the Pythagorean theorem shows that

$$|\langle \eta \,|\, \xi \rangle|^2 + \|\xi - \langle \eta \,|\, \xi \rangle \eta\|^2 = \|\xi\|^2.$$

But $\xi - \langle \eta \,|\, \xi \rangle \eta \in E$, so that $\|\xi - \langle \eta \,|\, \xi \rangle \eta\|^2 \geq \|\xi\|^2$, implying $\langle \eta \,|\, \xi \rangle = 0$. Thus ξ belongs to the *orthogonal complement* $M^\perp = \{\theta \in \mathcal{H} \mid \langle \theta \,|\, \eta \rangle = 0 \text{ for all } \eta \in M\}$ of M and $\varphi = \psi + \xi$ where $\psi \in M$, $\xi \in M^\perp$. Since $M \cap M^\perp = \{0\}$ and

$$M^\perp = \cap_{\eta \in M} \{\theta \in \mathcal{H} \mid \langle \theta \,|\, \eta \rangle = 0\}$$

is a closed subspace of $\mathcal{H}$, we have proved the following *projection theorem*:

Theorem 2.4 *If M is a closed subspace of $\mathcal{H}$, then $\mathcal{H}$ is the direct sum of M and the closed subspace $M^\perp$, that is, $\mathcal{H} = M \oplus M^\perp$.*

The statement $\mathcal{H} = M \oplus M^\perp$ means that every $\varphi \in \mathcal{H}$ can be uniquely expressed as $\varphi = \psi + \xi$ with $\psi \in M$ and $\xi \in M^\perp$. Denoting $P_M \varphi = \psi$, we thus obtain a mapping $P_M : \mathcal{H} \to M$, which we call the *(orthogonal) projection* of $\mathcal{H}$ onto M. The definition immediately shows that P_M is linear. Since $\|\varphi\|^2 = \|\psi\|^2 + \|\xi\|^2$ by the Pythagorean theorem, we have $\|P_M\varphi\| = \|\psi\| \leq \|\varphi\|$.

2.2 The Fréchet–Riesz Theorem and Bounded Linear Operators

Let $\mathcal{H}$ be a Hilbert space. A linear map $T : \mathcal{H} \to \mathcal{H}$ is called a *bounded (linear) operator* or a *bounded linear map* if there is a constant $C \in [0, \infty)$ satisfying $\|T\varphi\| \leq C\,\|\varphi\|$ for all $\varphi \in \mathcal{H}$. The existence of such a constant C is equivalent to the norm continuity of T. We let $\mathcal{L}(\mathcal{H})$ denote the set of all bounded linear maps $T : \mathcal{H} \to \mathcal{H}$. If $T \in \mathcal{L}(\mathcal{H})$ we write $\|T\| = \sup\{\|T\varphi\| \mid \varphi \in \mathcal{H},\ \|\varphi\| \leq 1\}$. It is easy to see that $\mathcal{L}(\mathcal{H})$ is a vector space, and $\|\cdot\|$ is a norm. Moreover, $\|T\varphi\| \leq \|T\|\,\|\varphi\|$, which implies that $\|ST\| \leq \|S\|\,\|T\|$ for all $S,\ T \in \mathcal{L}(\mathcal{H})$.

In general, we denote by G^* the *dual* of a normed space G (over $\mathbb{C}$), i.e. G^* is the space of continuous linear functionals $f : G \to \mathbb{C}$. Here continuity is equivalent to the condition

$$\|f\| = \sup\{|f(x)| \mid \|x\| \leq 1\} < \infty,$$

and the function $\|\cdot\|$ is a norm on G^*. The following key result is called the *Fréchet–Riesz representation theorem*.

Theorem 2.5 *For each $f \in \mathcal{H}^*$ there is a unique $\psi \in \mathcal{H}$ satisfying $f(\varphi) = \langle\, \psi \mid \varphi \,\rangle$ for all $\varphi \in \mathcal{H}$. Moreover, $\|\psi\| = \|f\|$.*

Proof Let $f \in \mathcal{H}^*$. We may assume that $f \neq 0$, so that $M = \{\varphi \mid f(\varphi) = 0\}$ is a proper closed subspace of $\mathcal{H}$. It follows from Theorem 2.4 that there is a $\xi \in M^\perp$ such that $\|\xi\| = 1$. If $\varphi \in \mathcal{H}$, then

$$\varphi - \frac{f(\varphi)}{f(\xi)}\xi \in M,$$

since $f\big(\varphi - \big(f(\varphi)/f(\xi)\big)\,\xi\big) = f(\varphi) - \big(f(\varphi)/f(\xi)\big)\,f(\xi) = 0$. This means that $\big\langle\, \xi \mid \varphi - \big(f(\varphi)/f(\xi)\big)\,\xi \,\big\rangle = 0$, implying $\langle\, \xi \mid \varphi \,\rangle = f(\varphi)/f(\xi)\ \langle\, \xi \mid \xi \,\rangle = f(\varphi)/f(\xi)$. Therefore we may choose $\psi = \overline{f(\xi)}\xi$. The uniqueness part is clear, since if $\big\langle\, \psi \mid \psi - \psi' \,\big\rangle = \big\langle\, \psi' \mid \psi - \psi' \,\big\rangle$, then $\big\|\psi - \psi'\big\|^2 = 0$. As $|f(\varphi)| \leq \|\psi\|\,\|\varphi\|$, we have $\|f\| \leq \|\psi\|$, and on the other hand $\|\psi\|^2 = f(\psi) \leq \|f\|\,\|\psi\|$, so that $\|\psi\| \leq \|f\|$. □

A straightforward consequence is that the mapping $\psi \mapsto f_\psi$ where $f_\psi(\varphi) = \langle\, \psi \mid \varphi \,\rangle$ for all $\varphi \in \mathcal{H}$ is a conjugate-linear isometric bijection from $\mathcal{H}$ onto $\mathcal{H}^*$. Another consequence is the following result.

Proposition 2.1 *Let $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a* bounded sesquilinear form, *i.e. a mapping satisfying the following conditions:*

(i) $B(\alpha\varphi + \beta\psi, \xi) = \overline{\alpha}B(\varphi, \xi) + \overline{\beta}B(\psi, \xi)$ *and*
(ii) $B(\varphi, \alpha\psi + \beta\xi) = \alpha B(\varphi, \psi) + \beta B(\varphi, \xi)$ *for all* $\alpha,\ \beta \in \mathbb{C},\ \varphi,\ \psi,\ \xi \in \mathcal{H}$;
(iii) $\sup\left\{|B(\varphi, \psi)| \,\middle|\, \|\varphi\| \le 1,\ \|\psi\| \le 1\right\} < \infty$.

Then there is a unique $S \in \mathcal{L}(\mathcal{H})$ such that $B(\varphi, \psi) = \langle\, S\varphi \,|\, \psi \,\rangle$ for all $\varphi,\ \psi \in \mathcal{H}$. Moreover, $\|S\| = \sup\left\{|B(\varphi, \psi)| \,\middle|\, \|\varphi\| \le 1,\ \|\psi\| \le 1\right\}$.

Proof Let C denote the supremum in (iii). If $\varphi \in \mathcal{H}$, we get a linear functional f_φ on $\mathcal{H}$ by setting $f_\varphi(\psi) = B(\varphi, \psi)$, and since $|f_\varphi(\psi)| \le C\,\|\varphi\|\,\|\psi\|$, f_φ is continuous. Theorem 2.5 yields a unique $\xi_\varphi \in \mathcal{H}$ such that $f_\varphi(\psi) = \left\langle \xi_\varphi \,\middle|\, \psi \right\rangle$ for all $\psi \in \mathcal{H}$. We define $S\varphi = \xi_\varphi$. Since $B(\alpha\varphi_1 + \beta\varphi_2, \psi) = \overline{\alpha}B(\varphi_1, \psi) + \overline{\beta}(\varphi_2, \psi) = \overline{\alpha}\,\langle\, S\varphi_1 \,|\, \psi \,\rangle + \overline{\beta}\,\langle\, S\varphi_2 \,|\, \psi \,\rangle = \langle\, \alpha S\varphi_1 + \beta S\varphi_2 \,|\, \psi \,\rangle$, S is linear. Since

$$\|S\varphi\|^2 = \langle\, S\varphi \,|\, S\varphi \,\rangle = B(\varphi, S\varphi) \le C\,\|\varphi\|\,\|S\varphi\|$$

we have $\|S\varphi\| \le C\,\|\varphi\|$, and so S is bounded. The uniqueness of S follows from that of ξ_φ. The proof of the norm equality is an easy exercise. □

The above result can be used to define for each $T \in \mathcal{L}(\mathcal{H})$ its *adjoint* as the map $T^* \in \mathcal{L}(\mathcal{H})$ which is characterised by the equation $\langle\, \varphi \,|\, T\psi \,\rangle = \langle\, T^*\varphi \,|\, \psi \,\rangle$ for all $\varphi,\ \psi \in \mathcal{L}(\mathcal{H})$: we simply take $B(\varphi, \psi) = \langle\, \varphi \,|\, T\psi \,\rangle$ in Proposition 2.1. Since $\|T^*\varphi\|^2 \le \langle\, \varphi \,|\, TT^*\varphi \,\rangle \le \|\varphi\|\,\|T\|\,\|T^*\varphi\|$, it is clear that $\|T^*\| \le \|T\|$. Using (a) in the next theorem, we see that on the other hand $\|T\| = \|T^{**}\| \le \|T^*\|$, and so $\|T^*\| = \|T\|$.

Theorem 2.6 *If $S,\ T \in \mathcal{L}(\mathcal{H})$ and $\alpha \in \mathbb{C}$, then*

(a) $T^{**} = T$;
(b) $(S + T)^* = S^* + T^*$;
(c) $(\alpha T)^* = \overline{\alpha}T^*$;
(d) $(ST)^* = T^*S^*$;
(e) $\|T^*T\| = \|T\|^2$.

We omit the simple proof. We still mention some notions defined in terms of the adjoint of $T \in \mathcal{L}(\mathcal{H})$. If $T^* = T$, T is *selfadjoint*. If $T^*T = TT^*$, T is *normal*. If $T^*T = TT^* = I$, where I (or $I_\mathcal{H}$) is the identity map of $\mathcal{H}$, T is *unitary*. If $\|T\varphi\| = \|\varphi\|$ for all $\varphi \in \mathcal{H}$, T is *isometric*. Using the polarisation identity it is easy to see that T is unitary if and only if it is an isometric surjection.

The norm of a selfadjoint operator has the following property.

Proposition 2.2 *If $T \in \mathcal{L}(\mathcal{H})$ is selfadjoint, then*

$$\|T\| = \sup_{\|\varphi\| \le 1} |\,\langle\, \varphi \,|\, T\varphi \,\rangle\,|.$$

Proof Using the polarisation identity and the parallelogram law we obtain for $\varphi,\ \psi \in \mathcal{H}$ with $\|\varphi\| \le 1$, $\|\psi\| \le 1$,

$$\begin{aligned}|\,\mathrm{Re}\,\langle\,\varphi\,|\,T\psi\,\rangle\,| &= t\frac{1}{4}|\,\langle\,\psi+\varphi\,|\,T(\psi+\varphi)\,\rangle - \langle\,\psi-\varphi\,|\,T(\psi-\varphi)\,\rangle\,| \\ &\le \tfrac{1}{4}M(\|\psi+\varphi\|^2+\|\psi-\varphi\|^2) = \tfrac{1}{2}M(\|\varphi\|^2+\|\psi\|^2) \le M,\end{aligned}$$

where $M = \sup_{\|\varphi\|\le 1} |\,\langle\,\varphi\,|\,T\varphi\,\rangle\,|$. (Note that, e.g., $\langle\,\psi+\varphi\,|\,T(\psi+\varphi)\,\rangle \in \mathbb{R}$.) Suppose that $\|\varphi\| \le 1$ and $\|\psi\| \le 1$. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $|\,\langle\,\varphi\,|\,T\psi\,\rangle\,| = \alpha\,\langle\,\varphi\,|\,T\psi\,\rangle = \langle\,\varphi\,|\,T\alpha\psi\,\rangle$. The first part of the proof (applied to the vectors $\alpha\psi$ and φ) yields $|\,\langle\,\varphi\,|\,T\psi\,\rangle\,| = \mathrm{Re}\,\langle\,\varphi\,|\,T\alpha\psi\,\rangle \le M$, so that

$$\|T\| = \sup\left\{|\,\langle\,\varphi\,|\,T\psi\,\rangle\,|\;\middle|\;\|\varphi\| \le 1,\ \|\psi\| \le 1\right\} \le M.$$

Conversely, $|\,\langle\,\varphi\,|\,T\varphi\,\rangle\,| \le \|T\|$ when $\|\varphi\| \le 1$, and so $M \le \|T\|$. □

We end this section with two useful decomposition results.

Proposition 2.3 *If $T \in \mathcal{L}(\mathcal{H})$ then T can be written in a unique way as $T = A + iB$ where $A,\ B \in \mathcal{L}(\mathcal{H})$ are selfadjoint. The operator T is normal if and only if $AB = BA$.*

Proof If we have $T = A + iB$, then necessarily $A = \frac{1}{2}(T + T^*)$ and $B = \frac{1}{2i}(T - T^*)$, and conversely. A simple calculation proves the second claim. □

For the set of the selfadjoint operators in $\mathcal{L}(\mathcal{H})$ we use the notation $\mathcal{L}_s(\mathcal{H})$. We now consider a natural partial order in $\mathcal{L}_s(\mathcal{H})$. We say that $T \in \mathcal{L}(\mathcal{H})$ is *positive*, and write $T \ge 0$, if $\langle\,\varphi\,|\,T\varphi\,\rangle \ge 0$ for all $\varphi \in \mathcal{H}$. We denote the set of the positive operators $T \in \mathcal{L}(\mathcal{H})$ by $\mathcal{L}_s(\mathcal{H})^+$. It follows from Remark 2.1 that for $T \in \mathcal{L}(\mathcal{H})$ we have $T \in \mathcal{L}_s(\mathcal{H})$ if and only if $\langle\,\varphi\,|\,T\varphi\,\rangle \in \mathbb{R}$ for all $\varphi \in \mathcal{H}$. In particular, $\mathcal{L}_s(\mathcal{H})^+ \subset \mathcal{L}_s(\mathcal{H})$. For $S,\ T \in \mathcal{L}_s(\mathcal{H})$ we write $S \le T$ if and only if $T - S \ge 0$. Clearly $T \le T$, and if $R \le S$ and $S \le T$, then $R \le T$. Moreover, the conditions $S \le T$ and $T \le S$ together imply $\langle\,\varphi\,|\,(T - S)\varphi\,\rangle = 0$ for all $\varphi \in \mathcal{H}$, and so by the polarisation identity (or Proposition 2.2) $S = T$. Thus we have a partial order in $\mathcal{L}_s(\mathcal{H})$. In the next decomposition result there is no uniqueness claim.

Proposition 2.4 *If $A \in \mathcal{L}_s(\mathcal{H})$ then A can be written as $A = A_1 - A_2$ where $A_1,\ A_2 \in \mathcal{L}_s(\mathcal{H})^+$.*

Proof We may choose $A_1 = \frac{1}{2}(\|A\|\,I + A)$ and $A_2 = \frac{1}{2}(\|A\|\,I - A)$. □

2.3 Strong, Weak, and Monotone Convergence of Nets of Operators

The usual norm on the space $\mathcal{L}(\mathcal{H})$ of bounded linear operators on $\mathcal{H}$ determines its canonical Banach space structure. There are several other important locally convex topologies on $\mathcal{L}(\mathcal{H})$. We postpone the discussion of some of them to later chapters. In this section two notions of convergence in $\mathcal{L}(\mathcal{H})$ are introduced: strong and weak. They are related to the so-called strong and weak operator topologies, but here we avoid the explicit use of these topologies.

Let $(\mathcal{I}, \geq)$ be a *directed set*. This means that "$\geq$" is a binary relation on the set $\mathcal{I}$ satisfying the following conditions:

(D1) $m \geq p$ whenever $m \geq n$ and $n \geq p$;
(D2) $m \geq m$ whenever $m \in \mathcal{I}$;
(D3) whenever $m, n \in \mathcal{I}$ there is some $p \in \mathcal{I}$ satisfying $p \geq m$ and $p \geq n$.

A mapping $i \mapsto x_i$ from $\mathcal{I}$ into a set X is then called a *net* or a *generalised sequence* (in X). Such a net is often denoted by $(x_i)_{i \in \mathcal{I}}$, generalising the notation for a sequence. If X here is a topological space, the net $(x_i)_{i \in \mathcal{I}}$ *converges* to a point $x \in X$ if for every neighbourhood U of x there is some $i_0 \in \mathcal{I}$ such that $x_i \in U$ whenever $i \geq i_0$. If X is a Hausdorff space, e.g., a metrisable space, condition (D3) implies that x, the *limit* of the net $(x_i)_{i \in \mathcal{I}}$, is uniquely determined. We use the notations $\lim_{i \in \mathcal{I}} x_i$, $\lim_i x_i$, $\lim x_i$ for this limit.

Definition 2.3 Let $(\mathcal{I}, \geq)$ be a directed set and $T_i \in \mathcal{L}(\mathcal{H})$ for all $i \in \mathcal{I}$.

(a) The net $(T_i)_{i \in \mathcal{I}}$ *converges strongly* to an operator $T \in \mathcal{L}(\mathcal{H})$ if $\lim T_i \varphi = T\varphi$ for all $\varphi \in \mathcal{H}$. We then denote $T_i \to^{s} T$ or $T = \text{s-lim}\, T_i$.
(b) The net $(T_i)_{i \in \mathcal{I}}$ *converges weakly* to an operator $T \in \mathcal{L}(\mathcal{H})$ if $\lim \langle \varphi \mid T_i \psi \rangle = \langle \varphi \mid T\psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$. We then denote $T_i \to^{w} T$ or $T = \text{w-lim}\, T_i$.

Since the inner products $\langle \varphi \mid T\psi \rangle$ completely determine T, the limit operator in (b) is also completely determined. The polarisation identity of Theorem 2.3 (c) shows that the condition $\lim \langle \varphi \mid T_i \varphi \rangle = \langle \varphi \mid T\varphi \rangle$ for all $\varphi \in \mathcal{H}$ already guarantees that $T_i \to^{w} T$. It is easy to see that a norm convergent net converges strongly and a strongly convergent net converges weakly. If the dimension of $\mathcal{H}$ is infinite, then in general neither implication can be reversed.

Let $\mathcal{F} \subset \mathcal{L}_s(\mathcal{H})$. The set $\mathcal{F}$ is *bounded above* if it has an *upper bound*, say $T \in \mathcal{L}_s(\mathcal{H})$, satisfying $S \leq T$ for all $S \in \mathcal{F}$. If T_0 is an upper bound of $\mathcal{F}$ satisfying $T_0 \leq T$ for every upper bound T of $\mathcal{F}$, then T_0 is the (clearly uniquely determined) *least upper bound (supremum)* of $\mathcal{F}$, and we denote $T_0 = \sup \mathcal{F}$. A *lower bound* and the *greatest lower bound (infimum)* $\inf \mathcal{F}$ are analogously defined. The same terminology is used for any partially ordered set.

Theorem 2.7 *Let $(\mathcal{I}, \geq)$ be a directed set and $(T_i)_{i \in \mathcal{I}}$ an increasing net in $\mathcal{L}_s(\mathcal{H})$ (i.e. $T_i \geq T_j$ whenever $i \geq j$). If the set $\{T_i \mid i \in \mathcal{I}\}$ is bounded above, then it has the least upper bound, say T. Moreover, $T_i \to^{s} T$ and $T_i \to^{w} T$. The similar statement concerning the greatest lower bounds of decreasing nets bounded below is also valid.*

Proof For each $\varphi \in \mathcal{H}$, the net $(\langle\, \varphi \mid T_i\varphi \,\rangle)_{i\in\mathcal{I}}$ in $\mathbb{R}$ is increasing and bounded above by $\langle\, \varphi \mid S_0\varphi \,\rangle$ where $S_0 \in \mathcal{L}_s(\mathcal{H})$ is some upper bound of $\left\{T_i \,\middle|\, i \in \mathcal{I}\right\}$ and so it has a limit which we denote by $f(\varphi)$. The polarisation identity shows that we can also define $B(\varphi, \psi) = \lim_{i\in\mathcal{I}} \langle\, \varphi \mid T_i\psi \,\rangle = \frac{1}{4}\sum_{n=0}^{3} i^n f(\psi + i^n\varphi)$ for all $\varphi,\ \psi \in \mathcal{H}$. The usual limit rules (valid also for nets) show that B satisfies the conditions (i) and (ii) in Proposition 2.1. We show that its boundedness condition (iii) also holds. Without loss of generality we may assume that $\mathcal{I}$ has a smallest element i_0, and since $\left\langle\, \xi \mid T_{i_0}\xi \,\right\rangle \le f(\xi) \le \langle\, \xi \mid S_0\xi \,\rangle$ for all $\xi \in \mathcal{H}$, we get $|B(\varphi, \psi)| \le |f(\psi + i^n\varphi)| \le \|\psi + i^n\varphi\|^2 \max\{\left\|T_{i_0}\right\|, \|S_0\|\} \le 4\max\{\left\|T_{i_0}\right\|, \|S_0\|\}$ whenever $\|\varphi\| \le 1$ and $\|\psi\| \le 1$. Using Proposition 2.1 we thus get a unique $T \in \mathcal{L}(\mathcal{H})$ such that $B(\varphi, \psi) = \langle\, T\varphi \mid \psi \,\rangle$ for all $\varphi,\ \psi \in \mathcal{H}$. One immediately verifies that $T \in \mathcal{L}_s(\mathcal{H})$ and $\langle\, \varphi \mid T\psi \,\rangle = \lim_{i\in\mathcal{I}} \langle\, \varphi \mid T_i\psi \,\rangle$ for all $\varphi,\ \psi \in \mathcal{H}$. By definition, $T_i \le T$, and if $S \in \mathcal{L}_s(\mathcal{H})$ satisfies $T_i \le S$ for all $i \in \mathcal{I}$, then $\langle\, \varphi \mid T\varphi \,\rangle = \lim_{i\in\mathcal{I}} \langle\, \varphi \mid T_i\varphi \,\rangle \le \langle\, \varphi \mid S\varphi \,\rangle$. Thus $T = \sup_{i\in\mathcal{I}} T_i$. We have also seen that $T = \text{w-lim}\, T_i$. We still show that $T = \text{s-lim}\, T_i$. The mapping $(\xi, \eta) \mapsto \langle\, \xi \mid (T - T_i)\eta \,\rangle$ is a positive sesquilinear form, and so it satisfies the Cauchy–Schwarz inequality (see Remark 2.1). Therefore, if $\varphi \in \mathcal{H}$, then

$$
\begin{aligned}
|\langle\, \xi \mid (T - T_i)\varphi \,\rangle|^2 &\le \langle\, \xi \mid (T - T_i)\xi \,\rangle \langle\, \varphi \mid (T - T_i)\varphi \,\rangle \\
&\le \left\langle\, \xi \mid (T - T_{i_0})\xi \,\right\rangle \langle\, \varphi \mid (T - T_i)\varphi \,\rangle \\
&\le \left\|T - T_{i_0}\right\| \langle\, \varphi \mid (T - T_i)\varphi \,\rangle,
\end{aligned}
$$

whenever $i \in \mathcal{I}$ and $\|\xi\| \le 1$, and so

$$
\|(T - T_i)\varphi\| = \sup_{\|\xi\|\le 1} |\langle\, \xi \mid (T - T_i)\varphi \,\rangle| \le (\left\|T - T_{i_0}\right\| \langle\, \varphi \mid (T - T_i)\varphi \,\rangle)^{\frac{1}{2}} \longrightarrow 0.
$$

When the operators above are multiplied by -1 we get the claim concerning decreasing nets. □

The following observation will be used later.

Theorem 2.8 (a) *Let $(T_i)_{i\in\mathcal{I}}$ be a net in $\mathcal{L}(\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H})$ such that $T_i \xrightarrow{\text{w}} T$. Then $T_i^* \xrightarrow{\text{w}} T^*$ and $T_i S \xrightarrow{\text{w}} TS$, $ST_i \xrightarrow{\text{w}} ST$ for all $S \in \mathcal{L}(\mathcal{H})$.*

(b) *If $(T_i)_{i\in\mathcal{I}}$ is a net in $\mathcal{L}_s(\mathcal{H})^+$ which is increasing and bounded above or decreasing and bounded below, and $T = \text{w-lim}\, T_i$, then $T^2 = \text{w-lim}\, T_i^2$.*

Proof (a) A straightforward calculation yields this.

(b) From Theorem 2.7 it follows that $T = \text{s-lim}\, T_i$. Hence for every φ we get $\left\langle\, \varphi \mid T_i^2\varphi \,\right\rangle = \langle\, T_i\varphi \mid T_i\varphi \,\rangle = \|T_i\varphi\|^2 \to \|T\varphi\|^2 = \left\langle\, \varphi \mid T^2\varphi \,\right\rangle$ implying the claim. □

2.4 The Projection Lattice $\mathcal{P}(\mathcal{H})$

We have defined the (orthogonal) projection of $\mathcal{H}$ onto a closed subspace $M \subset \mathcal{H}$ as the map P_M ($\in \mathcal{L}(\mathcal{H})$) for which $P_M\varphi = \psi$ when $\varphi = \psi + \xi$ with $\psi \in M, \xi \in M^\perp$. We say that P is a *projection* if $P = P_M$ for some closed subspace M of $\mathcal{H}$. We omit the proof of the following list of elementary properties of projections. We generally denote $I = I_\mathcal{H} = \mathrm{id}_\mathcal{H} = P_\mathcal{H}$.

Theorem 2.9 *Let M be a closed subspace of $\mathcal{H}$ and $P = P_M$. Then*

(a) $M = P(\mathcal{H}) = \{\varphi \,|\, P\varphi = \varphi\} = \{\varphi \,|\, \|P\varphi\| = \|\varphi\|\}$;
(b) $M^\perp = \{\varphi \,|\, P\varphi = 0\}$;
(c) $P = P^2 = P^*$;
(d) $P_M + P_{M^\perp} = I$;
(e) $M^{\perp\perp} = M$.

The following characterisation is basic.

Theorem 2.10 *For a linear map $P : \mathcal{H} \to \mathcal{H}$, the following conditions are equivalent:*

(i) *P is a projection;*
(ii) *$P = P^2$ and $\langle \varphi \,|\, P\psi \rangle = \langle P\varphi \,|\, \psi \rangle$ for all $\varphi, \psi \in \mathcal{H}$;*
(iii) *$P = P^2$ and $\|P\| \le 1$ (in particular, $P \in \mathcal{L}(\mathcal{H})$).*

Proof By Theorem 2.9 (i) implies (ii), and (ii) implies (iii) because $\|P\varphi\|^2 = \langle P\varphi \,|\, P\varphi \rangle = \langle \varphi \,|\, P^2\varphi \rangle = \langle \varphi \,|\, P\varphi \rangle \le \|\varphi\| \, \|P\varphi\|$ so that $\|P\varphi\| \le \|\varphi\|$. Now assume (iii). As $P = P^2$, one gets the direct sum representation $\mathcal{H} = M \oplus N$ with the notation $M = P(\mathcal{H})$, $N = (I - P)(\mathcal{H})$. Moreover, N is the kernel of P and M that of $I - P$, and so both are closed. Since also $\mathcal{H} = M \oplus M^\perp$, to see that $P = P_M$ it will suffice to show, e.g., that $M^\perp \subset N$. If $\varphi \in N^\perp$, we have $P\varphi = \varphi + \psi$ where $\psi = P\varphi - \varphi \in N$, and so $\langle \psi \,|\, \varphi \rangle = 0$, implying $\|\varphi\|^2 \ge \|P\varphi\|^2 = \|\varphi\|^2 + \|\psi\|^2$ so that $\psi = 0$ and $\varphi = P\varphi \in M$. Thus $N^\perp \subset M$, implying $M^\perp \subset N^{\perp\perp} = N$. □

In the rest of this section we use the notation $\mathcal{M}(\mathcal{H})$ for the set of the closed subspaces of $\mathcal{H}$. Theorems 2.9 and 2.10 show that the mapping $P \mapsto P(\mathcal{H})$ is a bijection from the set $\mathcal{P}(\mathcal{H}) = \{P \in \mathcal{L}(\mathcal{H}) \,|\, P = P^2 = P^*\}$ onto $\mathcal{M}(\mathcal{H})$. We mention some results related to this bijection. The proofs are straightforward and we omit them. In the following three theorems we assume that $P, Q \in \mathcal{P}(\mathcal{H})$ and denote $M = P(\mathcal{H})$, $N = Q(\mathcal{H})$.

Theorem 2.11 *The following conditions are equivalent:*

(i) $M \subset N$;
(ii) $QP = P$;
(iii) $PQ = P$;
(iv) *$\|P\varphi\| \le \|Q\varphi\|$ for all $\varphi \in \mathcal{H}$;*
(v) *$P \le Q$ (i.e. $\langle \varphi \,|\, P\varphi \rangle \le \langle \varphi \,|\, Q\varphi \rangle$ for all $\varphi \in \mathcal{H}$);*

(vi) $Q - P \in \mathcal{P}(\mathcal{H})$.

Theorem 2.12 *The following conditions are equivalent:*

(i) $M \perp N$ *(i.e.* $\langle \varphi \mid \psi \rangle = 0$ *for all* $\varphi \in M$*,* $\psi \in N$*);*
(ii) $PQ = 0$*;*
(iii) $QP = 0$*;*
(iv) $Q + P \in \mathcal{P}(\mathcal{H})$.

Theorem 2.13 *The following conditions are equivalent:*

(i) $PQ = QP$*;*
(ii) $PQ \in \mathcal{P}(\mathcal{H})$*;*
(iii) $QP \in \mathcal{P}(\mathcal{H})$.

If these conditions hold, then $PQ(\mathcal{H}) = P(\mathcal{H}) \cap Q(\mathcal{H})$.

The next result generalises parts (i) and (v) of Theorem 2.11.

Proposition 2.5 *If* $P \in \mathcal{P}(\mathcal{H})$*,* $M = P(\mathcal{H})$ *and* $T \in \mathcal{L}_s(\mathcal{H})^+$*, then the following conditions are equivalent:*

(i) $T(\mathcal{H}) \subset M$ *and* $\|T\| \le 1$*;*
(ii) $T \le P$.

Proof First assume (i). For all $\varphi \in \mathcal{H}$ we get

$$\langle \varphi \mid T\varphi \rangle = \langle \varphi \mid PT\varphi \rangle = \langle P\varphi \mid TP\varphi \rangle \le \|P\varphi\|\,\|T\|\,\|P\varphi\| \le \|P\varphi\|^2 = \langle \varphi \mid P\varphi \rangle\,,$$

since $PT = T$, so that $T = T^* = T^*P^* = TP$.

Next assume (ii). Since the map $(\varphi, \psi) \mapsto \langle \varphi \mid T\psi \rangle$ is a positive sesquilinear form, the Cauchy–Schwarz inequality gives $|\langle \xi \mid T\varphi \rangle|^2 \le \langle \xi \mid T\xi \rangle \langle \varphi \mid T\varphi \rangle \le \langle \xi \mid P\xi \rangle \langle \varphi \mid P\varphi \rangle$ for all $\xi, \varphi \in \mathcal{H}$. In particular, $\|T\varphi\|^2 = \sup_{\|\xi\| \le 1} |\langle \xi \mid T\varphi \rangle|^2 \le 1$ when $\|\varphi\| \le 1$, and so $\|T\| \le 1$. Moreover, $T\varphi = 0$ if $P\varphi = 0$, and so $\langle \varphi \mid T\psi \rangle = \langle T\varphi \mid \psi \rangle = 0$ whenever $\psi \in \mathcal{H}$ and $\varphi \in M^\perp$, implying $T\psi \in M^{\perp\perp} = M$. □

In the partially ordered set $(\mathcal{L}_s(\mathcal{H}), \le)$ it is common that even two-element sets fail to have least upper bounds and greatest lower bounds. Thus $(\mathcal{L}_s(\mathcal{H}), \le)$ is not a lattice. (By definition, a *lattice* is a partially ordered set such that any two-element subset has supremum and infimum.) In fact, it is an *antilattice*, that is, any two elements S and T have greatest lower bound exactly when they are comparable, i.e. $S \le T$ or $T \le S$ (and the same is true of least upper bounds). For a proof of this result due to R. Kadison we refer to [1], p. 417. The subset $\mathcal{P}(\mathcal{H})$ of $(\mathcal{L}_s(\mathcal{H}), \le)$ is, however, even a complete lattice. The next result gives even more information: if $\mathcal{P}(\mathcal{H})$ is regarded as a subset of the larger partially ordered set $\left\{T \in \mathcal{L}_s(\mathcal{H}) \,\middle|\, 0 \le T \le I\right\}$ (which we denote by $[0, I]$), the greatest lower bound and the least upper bound of a subset of $\mathcal{P}(\mathcal{H})$ are not affected.

Theorem 2.14 *Let $(P_i)_{i\in\mathcal{I}}$ be any family in $\mathcal{P}(\mathcal{H})$ and denote $M_i = P_i(\mathcal{H})$ for $i \in \mathcal{I}$. Write $M = \cap_{i\in\mathcal{I}} M_i$ and let N be the intersection of all the closed subspaces of $\mathcal{H}$ containing every M_i, $i \in \mathcal{I}$. Then P_M is the greatest lower bound and P_N the least upper bound of the set $\{P_i \mid i \in \mathcal{I}\}$ with respect to the set* $[0, I] = \{T \in \mathcal{L}_s(\mathcal{H}) \mid 0 \le T \le I\}$.

Proof According to Theorem 2.11, $P_M \le P_i$ for all $i \in \mathcal{I}$. On the other hand, if $T \in [0, I]$ is a lower bound of the set $\{P_i \mid i \in \mathcal{I}\}$, Proposition 2.5 shows us that $\|T\| \le 1$ and $T(\mathcal{H}) \subset M_i$ for all $i \in \mathcal{I}$, i.e. $T(\mathcal{H}) \subset M$, so that by the same theorem $T \le P_M$. This proves the claim about M.

As for N, note that by the above argument, the set of the complementary projections, $\{I - P_i \mid i \in \mathcal{I}\}$, has a greatest lower bound S in the set $[0, I]$. The mapping $T \mapsto I - T$ keeps $[0, I]$ invariant and inverts the order of operators, which implies that $I - S$ is the least upper bound of the set $\{P_i \mid i \in \mathcal{I}\}$ in $[0, I]$. Now S and hence also $I - S$ is a projection. We show that $I - S = P_N$. We have $M_i \subset (I - S)(\mathcal{H})$ for all $i \in \mathcal{I}$, so that by the definition of N we have $I - S \ge P_N$. On the other hand, P_N is an upper bound of the set $\{P_i \mid i \in \mathcal{I}\}$, and so $I - S \le P_N$. □

In the situation of the above theorem we denote $P_M = \wedge_{i\in\mathcal{I}} P_i$ and $P_N = \vee_{i\in\mathcal{I}} P_i$. In the case of finitely many projections $P_1, \ldots, P_n$ we may also write $\wedge_{i\in\{1,\ldots,n\}} P_i = P_1 \wedge \cdots \wedge P_n$ and $\vee_{i\in\{1,\ldots,n\}} P_i = P_1 \vee \cdots \vee P_n$.

Remark 2.3 Consider on $\mathcal{P}(\mathcal{H})$ the map $P \mapsto P^\perp = I - P$. It is an order-reversing involution and maps each element to a complement, i.e. $P \vee P^\perp = I$ and $P \wedge P^\perp = 0$ (see the above theorem). Such a map is called an *orthocomplementation*. We say that $P, R \in \mathcal{P}(\mathcal{H})$ are *orthogonal* if $P \le R^\perp$, equivalently, $R \le P^\perp$, and we denote $P \perp R$. If P and R are orthogonal, they are also *disjoint*, that is, $P \wedge R = 0$. If $\dim(\mathcal{H}) \ge 2$, it is easy to give examples of projections which are disjoint but not orthogonal. ◁

2.5 The Square Root of a Positive Operator

We prove that any positive operator has a unique positive square root.

Theorem 2.15 *Suppose that $A \in \mathcal{L}_s(\mathcal{H})^+$. There is a uniquely determined operator $B \in \mathcal{L}_s(\mathcal{H})^+$ (usually denoted by $A^{\frac{1}{2}}$ or $\sqrt{A}$ and called the* square root of *A) satisfying $B^2 = A$.*

Proof We may assume that $A \in \mathcal{L}_s(\mathcal{H})$ is such that $0 \le A \le I$. Define a sequence (B_n) of operators recursively by setting $B_1 = 0$,

$$B_{n+1} = \tfrac{1}{2}[(I - A) + B_n^2],$$

$n = 1, 2, \ldots$. Then

$$B_{n+1} - B_n = \tfrac{1}{2}[B_n^2 - B_{n-1}^2] = \tfrac{1}{2}(B_n + B_{n-1})(B_n - B_{n-1}),$$

since the operators B_n ja B_{n-1}, being polynomials of $I - A$, commute. By induction one sees that in fact each B_n is a polynomial of $I - A$ with positive coefficients. After this observation, from the equation obtained before we see by induction that each difference is a polynomial in $I - A$ with positive coefficients, so that this difference is a positive operator, as every power $(I - A)^k \geq 0$. (Observe that $\langle (I-A)^{2m}\varphi \,|\, \varphi\rangle = \langle (I-A)^m\varphi \,|\, (I-A)^m\varphi\rangle \geq 0$ and $\langle (I-A)^{2m+1}\varphi \,|\, \varphi\rangle = \langle (I-A)(I-A)^m\varphi \,|\, (I-A)^m\varphi\rangle \geq 0$.) By induction it is seen that $\|B_n\| \leq 1$. Hence the increasing sequence (B_n) has a positive bounded (strong) limit operator $\tilde{B} \leq I$ by Theorem 2.7, and the equality $\tilde{B} = \frac{1}{2}[I - A + \tilde{B}^2]$ follows from Theorem 2.8 (b). Denoting $B = I - \tilde{B}$, we thus have $B^2 = A$.

We next prove the uniqueness claim. Consider the above situation where $0 \leq A \leq I$. Let B be the operator constructed above, satisfying $B^2 = A$. Let also $C \in \mathcal{L}_s(\mathcal{H})^+$ satisfy $C^2 = A$. Then $CA = C^3 = AC$, so that C commutes with every polynomial in A, implying that $CB = BC$. We now use the method described in the first part of the proof to obtain two operators $B_1, C_1 \in \mathcal{L}_s(\mathcal{H})^+$ such that $B_1^2 = B$ and $C_1^2 = C$. Let $\varphi \in \mathcal{H}$ and denote $\psi = (B - C)\varphi$. Then

$$\begin{aligned}\|B_1\psi\|^2 + \|C_1\psi\|^2 &= \langle B_1^2\psi \,|\, \psi\rangle + \langle C_1^2\psi \,|\, \psi\rangle = \langle B\psi \,|\, \psi\rangle + \langle C\psi \,|\, \psi\rangle \\ &= \langle B(B-C)\varphi \,|\, \psi\rangle + \langle C(B-C)\varphi \,|\, \psi\rangle \\ &= \langle (B^2 - C^2)\varphi \,|\, \psi\rangle = 0.\end{aligned}$$

It follows that $B_1\psi = C_1\psi = 0$, so that $B\psi = B_1B_1\psi = 0$ and $C\psi = C_1C_1\psi = 0$, implying $\|(B-C)\varphi\|^2 = \langle (B-C)^2\varphi \,|\, \varphi\rangle = \langle (B-C)\psi \,|\, \varphi\rangle = 0$. Thus $B\varphi = C\varphi$ for all $\varphi \in \mathcal{H}$, i.e. $B = C$. □

The square root gives rise to the following definition.

Definition 2.4 If $T \in \mathcal{L}(\mathcal{H})$, the positive operator $\sqrt{T^*T}$ is denoted by $|T|$ and called the *absolute value* of T.

The absolute value $|T|$ may be characterised as the only positive operator A satisfying $\|A\varphi\| = \|T\varphi\|$ for all $\varphi \in \mathcal{H}$ (exercise).

Remark 2.4 Using the square root we get a quick proof for the implication (ii) $\Longrightarrow$ (i) in Proposition 2.5: $\left\|T^{\frac{1}{2}}\varphi\right\|^2 = \langle\, \varphi \,|\, T\varphi\,\rangle \leq \langle\, \varphi \,|\, P\varphi\,\rangle \leq 1$ if $\|\varphi\| \leq 1$, and so $\|T\| \leq \left\|T^{\frac{1}{2}}\right\|^2 \leq 1$. Moreover, $P\varphi = 0$ implies $T\varphi = (T^{\frac{1}{2}})^2\varphi = 0$, and the proof is completed as originally. ◁

We conclude with another application of the square root.

Proposition 2.6 *Any operator $T \in \mathcal{L}(\mathcal{H})$ can be written as a linear combination of four unitary operators.*

Proof We first write $T = A + iB$ with A and B selfadjoint. We may assume that $\|A\| \leq 1$ and $\|B\| \leq 1$. Define $U = A + i\sqrt{I - A^2}$. Then $U^* = A - i\sqrt{I - A^2}$ and we have $U^*U = UU^* = I$, $A = \frac{1}{2}(U + U^*)$. Similarly, B is the linear combination of two unitary operators. □

2.6 The Polar Decomposition of a Bounded Operator

For any $T \in \mathcal{L}(\mathcal{H})$, we denote $\ker(T) = \{\varphi \in \mathcal{H} \mid T\varphi = 0\}$.

Lemma 2.1 *If $T \in \mathcal{L}(\mathcal{H})$, then $\overline{T(\mathcal{H})} = \ker(T^*)^\perp$.*

Proof As $T(\mathcal{H})^\perp = \overline{T(\mathcal{H})}^\perp$, it is enough to show that $\ker(T^*) = T(\mathcal{H})^\perp$. The following conditions are equivalent for a vector $\varphi \in \mathcal{H} : \varphi \in \ker(T^*)$, $\langle T^*\varphi|\psi\rangle = 0$ for all $\psi \in \mathcal{H}$, $\langle\varphi|T\psi\rangle = 0$ for all $\psi \in \mathcal{H}$. □

For every operator $T \in \mathcal{L}(\mathcal{H})$ we denote by $\mathrm{supp}(T)$ the (orthogonal) projection of $\mathcal{H}$ onto the closed subspace $\ker(T)^\perp$. We say that an operator $V \in \mathcal{L}(\mathcal{H})$ is a *partial isometry* or *partially isometric* if $\|V\mathrm{supp}(V)\varphi\| = \|\mathrm{supp}(V)\varphi\|$ for all $\varphi \in \mathcal{H}$. We then say that $\mathrm{supp}(V)$ is the *initial projection* of V, and the projection P_M onto the (closed since $\mathrm{supp}(V)(\mathcal{H})$ is closed) subspace $M = \{V\varphi \mid \varphi \in \mathcal{H}\} = \{V\mathrm{supp}(V)\varphi \mid \varphi \in \mathcal{H}\}$ of $\mathcal{H}$ is the *final projection* of V.

Theorem 2.16 *Let $V \in \mathcal{L}(\mathcal{H})$ be partially isometric, P the initial projection of V and Q the final projection of V. Then*

(a) *V^* is partially isometric, Q is the initial projection of V^*, P is the final projection of V^*, and*
(b) *$V^*V = P$, $VV^* = Q$.*

Proof Let $\varphi \in P(\mathcal{H})$ and $\psi = V\varphi \in Q(\mathcal{H})$. The polarisation identity implies that restricted to the subspace $P(\mathcal{H})$, V preserves the inner product, so that for each $\xi \in \mathcal{H}$ we get

$$\langle\xi|\varphi\rangle = \langle P\xi|\varphi\rangle + \langle(I - P)\xi|\varphi\rangle = \langle P\xi|\varphi\rangle = \langle VP\xi|V\varphi\rangle = \langle V\xi|\psi\rangle = \langle\xi|V^*\psi\rangle.$$

Thus $V^*\psi = \varphi$. It follows that $V^*|Q(\mathcal{H})$ is the inverse of the map $V|P(\mathcal{H}) : P(\mathcal{H}) \to Q(\mathcal{H})$ and hence isometric. In view of Lemma 2.1, $Q = \mathrm{supp}(V^*)$ and $P(\mathcal{H}) = \ker(V)^\perp = \overline{V^*(\mathcal{H})} = V^*(\mathcal{H})$, and (a) is proved. Clearly also (b) is true, since $V^*|Q(\mathcal{H})$ and $V|P(\mathcal{H})$ are the inverses of each other. □

Theorem 2.17 *If $V \in \mathcal{L}(\mathcal{H})$ is such that V^*V is a projection, then V is partially isometric.*

Proof Let $V^*V = P$ be a projection. If $\varphi \in \mathcal{H}$, then $\|V\varphi\|^2 = \langle\varphi|V^*V\varphi\rangle = \langle\varphi|P\varphi\rangle = \|P\varphi\|^2$, and so $V|P(\mathcal{H})$ is isometric and $V|(P(\mathcal{H}))^\perp = 0$. Here $P = \mathrm{supp}(V)$, because $P(\mathcal{H})^\perp = \ker(V)$. □

Theorem 2.18 *Let $T \in \mathcal{L}(\mathcal{H})$. There is a uniquely determined pair of operators $V, A \in \mathcal{L}(\mathcal{H})$ such that*

(i) $T = VA$,
(ii) $A \geq 0$, *and*
(iii) V *is a partially isometric operator whose initial projection is* $\operatorname{supp}(A)$.

Then $A = |T|$ and $\operatorname{supp}(V) = \operatorname{supp}(|T|) = \operatorname{supp}(T)$. *Moreover, $|T| = V^*T$.*

Proof Let $P = \operatorname{supp}(T)$, $Q = \operatorname{supp}(T^*)$. If $\varphi \in \mathcal{H}$, we have $\|T\varphi\|^2 = \langle\varphi|T^*T\varphi\rangle = \langle\varphi|\,|T|^2\varphi\rangle = \|\,|T|\varphi\,\|^2$, so that $\operatorname{supp}(|T|) = P$ and hence in view of Lemma 2.1, $\overline{|T|(\mathcal{H})} = P(\mathcal{H})$, as $|T|^* = |T|$. This shows, moreover, that the mapping $|T|\varphi \mapsto T\varphi$ is a well-defined (for $T\varphi - T\psi = 0$, when $|T|\varphi - |T|\psi = 0$) linear isometry from $|T|(\mathcal{H})$ onto $T(\mathcal{H})$ and can hence be uniquely extended to a linear isometry $V_0 : P(\mathcal{H}) \to Q(\mathcal{H})$ (for $Q(\mathcal{H}) = \overline{T(\mathcal{H})}$ by Lemma 2.1). Let now V be a partially isometric operator such that $\operatorname{supp}(V) = P$ and $V|P(\mathcal{H}) = V_0$, i.e. V is the map $\varphi \mapsto V_0P\varphi$. When we choose $A = |T|$, the requirements (i)–(iii) are fulfilled. From the construction it is clear that $\operatorname{supp}(V) = \operatorname{supp}(|T|) = \operatorname{supp}(T)$ and $|T| = P|T| = V^*V|T| = V^*T$ (see Theorem 2.16 (b)). We next prove the uniqueness part. Let V and A be such that the conditions (i)–(iii) are valid. Then $T^*T = AV^*VA = A\operatorname{supp}(V)A = A^2$ (see Theorem 2.16 (b)), and so by the uniqueness of the positive square root (Theorem 2.15) we get $A = |T|$. From this also the values of V are uniquely determined. □

Definition 2.5 The representation $T = V|T|$ mentioned in Theorem 2.18 is called the *polar decomposition* of T.

2.7 Orthonormal Sets

If $\varphi, \psi \in \mathcal{H}$ and $\langle\varphi|\psi\rangle = 0$, we say that φ and ψ are *mutually orthogonal* and we write $\varphi \perp \psi$. If $K \subset \mathcal{H}$, $\varphi \in \mathcal{H}$ and $\varphi \perp \psi$ for all $\psi \in K$, we denote $\varphi \perp K$. The set $K \subset \mathcal{H}$ is *orthogonal*, if $\varphi \perp \psi$ whenever $\varphi, \psi \in K$, $\varphi \neq \psi$. The family $(\varphi_i)_{i\in\mathcal{I}}$ (and especially a sequence) in $\mathcal{H}$ is orthogonal if $\varphi_i \perp \psi_j$ whenever $i \neq j$. A set or a family is *orthonormal*, if it is orthogonal and in addition every vector in it has norm 1.

A family $(c_i)_{i\in\mathcal{I}}$ of complex numbers is *summable*, if there is a constant $M \in [0, \infty)$ such that $\sum_{i\in F} |c_i| \leq M$ for every finite subset F of $\mathcal{I}$. Then for all $n \in \mathbb{N}$ $|c_i| \geq \frac{1}{n}$ for at most a finite number of $i \in \mathcal{I}$, so that the set $\{i \in \mathcal{I}|\ |c_i| > 0\}$ is at most countable: $\{i_1, i_2, \ldots\}$, and the series $\sum_{k=1}^{\infty} c_{i_k}$ is absolutely convergent (in the case of an infinite set), so that independently of the numbering we may define the *sum* of the family $(c_i)_{i\in\mathcal{I}}$ as $\sum_{i\in\mathcal{I}} c_i = \sum_{k=1}^{\infty} c_{i_k}$. If $\{i \in \mathcal{I}|\ |c_i| > 0\}$ is finite, the definition of the sum is obvious (in the case of the empty set the sum is 0). We denote $\sum_{i\in\mathcal{I}} |c_i| = \infty$ if the family is not summable.

Lemma 2.2 *Let $\varphi \in \mathcal{H}$.*

(a) *If $(\varphi_k)_{k=1}^n$ is a finite orthonormal family in $\mathcal{H}$, then for all $c_1, \dots, c_n \in \mathbb{C}$ we have*

$$\Big\| \varphi - \sum_{k=1}^n c_k \varphi_k \Big\| \geq \Big\| \varphi - \sum_{k=1}^n \langle \varphi_k | \varphi \rangle \varphi_k \Big\|,$$

and equality holds only if $c_k = \langle \varphi_k | \varphi \rangle$ for all $k = 1, \dots, n$. Moreover,

$$\sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2 \leq \|\varphi\|^2. \tag{2.1}$$

(b) *If $K \subset \mathcal{H}$ is an orthonormal set, then the family $(|\langle \psi | \varphi \rangle|^2)_{\psi \in K}$ is summable and*

$$\sum_{\psi \in K} |\langle \psi | \varphi \rangle|^2 \leq \|\varphi\|^2 \qquad \textit{(Bessel's inequality).}$$

Proof (a) As

$$\begin{aligned}
\|\varphi - \sum_{k=1}^n c_k \varphi_k\|^2 &= \|\varphi\|^2 - \sum_{k=1}^n \overline{c_k} \langle \varphi_k | \varphi \rangle - \sum_{k=1}^n c_k \langle \varphi | \varphi_k \rangle + \sum_{k=1}^n |c_k|^2 \\
&= \|\varphi\|^2 + \sum_{k=1}^n \big[|\langle \varphi_k | \varphi \rangle|^2 - \overline{c_k} \langle \varphi_k | \varphi \rangle - c_k \overline{\langle \varphi_k | \varphi \rangle} + |c_k|^2 \big] - \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2 \\
&= \|\varphi\|^2 + \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle - c_k|^2 - \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2,
\end{aligned}$$

the choice $c_k = \langle \varphi_k | \varphi \rangle$ yields (2.1). Moreover, $\|\varphi - \sum_{k=1}^n c_k \varphi_k\|^2$ also gets its smallest possible value $\|\varphi\|^2 - \sum_{k=1}^n |\langle \varphi_k | \varphi \rangle|^2$ with this and only this choice.

(b) The claim is an immediate consequence of (a). □

Theorem 2.19 *Let (φ_n) be an orthogonal sequence in $\mathcal{H}$.*

(a) *The series $\sum_{n=1}^\infty \varphi_n$ converges if and only if $\sum_{n=1}^\infty \|\varphi_n\|^2 < \infty$. If $\varphi = \sum_{n=1}^\infty \varphi_n$, then $\|\varphi\|^2 = \sum_{n=1}^\infty \|\varphi_n\|^2$.*

(b) *If (φ_n) is orthonormal, then for every $\varphi \in \mathcal{H}$ the series $\sum_{k=1}^\infty \langle \varphi_k | \varphi \rangle \varphi_k$ converges, and if ψ is its sum, then $(\varphi - \psi) \perp \varphi_n$ for all $n \in \mathbb{N}$.*

Proof (a) If $n > m$, then by orthogonality

$$\begin{aligned}
\Big\| \sum_{k=1}^n \varphi_k - \sum_{k=1}^m \varphi_k \Big\|^2 = \Big\| \sum_{k=m+1}^n \varphi_k \Big\|^2 &= \sum_{k=m+1}^n \|\varphi_k\|^2 \\
&= \sum_{k=1}^n \|\varphi_k\|^2 - \sum_{k=1}^m \|\varphi_k\|^2.
\end{aligned}$$

This implies that the sequence of the partial sums $s_n = \sum_{k=1}^{n} \varphi_k$ is a Cauchy sequence in $\mathcal{H}$ if and only if the sequence of the partial sums $\sum_{k=1}^{n} \|\varphi_k\|^2$ is a Cauchy sequence in $\mathbb{R}$. This proves the first claim as $\mathcal{H}$ and $\mathbb{R}$ are complete. Since $\|s_n\|^2 = \sum_{k=1}^{n} \|\varphi\|^2$, the continuity of the norm shows that

$$\|\varphi\|^2 = \lim_{n\to\infty} \|s_n\|^2 = \lim_{n\to\infty} \sum_{k=1}^{n} \|\varphi_k\|^2, \text{ when } \varphi = \sum_{k=1}^{\infty} \varphi_k.$$

(b) The convergence of the series follows form (a) and Lemma 2.2 (b), since $\|\langle\varphi_k|\varphi\rangle\varphi_k\|^2 = |\langle\varphi_k|\varphi\rangle|^2$. Since the map $\xi \mapsto \langle\varphi_n|\xi\rangle$ is continuous, we obtain

$$\langle\varphi - \psi|\varphi_n\rangle = \langle\varphi|\varphi_n\rangle - \sum_{k=1}^{\infty}\langle\varphi|\varphi_k\rangle\langle\varphi_k|\varphi_n\rangle = 0.$$ □

In the next theorem the notation $\varphi = \sum_{\xi\in K}\langle\xi|\varphi\rangle\xi$ means that $\langle\xi|\varphi\rangle\xi \neq 0$ for at most a countable number of vectors ξ: $\xi_1, \xi_2, \ldots$, and that independently of the numbering, $\varphi = \sum_n \langle\varphi_n|\varphi\rangle\varphi_n$ (a convergent series or a finite or "empty" sum ($= 0$)). In (iii) and (iv) there is a summable family.

Theorem 2.20 *Let $K \subset \mathcal{H}$ be an orthonormal set. The following conditions are equivalent:*

(i) *if $\varphi \perp K$, then $\varphi = 0$;*
(ii) $\varphi = \sum_{\xi\in K}\langle\xi|\varphi\rangle\xi$ *for all* $\varphi \in \mathcal{H}$;
(iii) $\langle\psi|\varphi\rangle = \sum_{\xi\in K}\langle\xi|\varphi\rangle\overline{\langle\xi|\psi\rangle}$ *for all* $\varphi,\ \psi \in \mathcal{H}$;
(iv) $\|\varphi\|^2 = \sum_{\xi\in K}|\langle\xi|\varphi\rangle|^2$ *for all* $\varphi \in \mathcal{H}$;
(v) *the vector subspace M of $\mathcal{H}$ generated by the set K is dense in $\mathcal{H}$.*

Proof Assume first (i). Let $\varphi \in \mathcal{H}$. According to Lemma 2.2 (b) $\langle\xi|\varphi\rangle \neq 0$ for at most a countable number of the vectors $\xi \in K$; let $\xi_1, \xi_2, \ldots$ be their arbitrary numbering. By Theorem 2.19 (b) there is $\psi = \sum_n\langle\xi_n|\varphi\rangle\xi_n$, and $\langle\xi_n|\varphi - \psi\rangle = 0$ for all $n = 1, 2, \ldots$. Also, if $\xi \in K \setminus \{\xi_1, \xi_2, \ldots\}$ then $\langle\xi|\varphi - \psi\rangle = \langle\xi|\varphi\rangle - \sum_n\langle\xi_n|\varphi\rangle\langle\xi|\xi_n\rangle = 0$. Hence $\varphi - \psi \perp K$, so that by (i) $\varphi = \psi$ and so (ii) holds.

Assume now (ii). By the Cauchy–Schwarz inequality and Lemma 2.2, for every finite set $F \subset K$ we have

$$\sum_{\xi\in F}|\langle\xi|\varphi\rangle\overline{\langle\xi|\psi\rangle}| \leq \Big(\sum_{\xi\in F}|\langle\xi|\varphi\rangle|^2\Big)^{\frac{1}{2}}\Big(\sum_{\xi\in F}|\langle\xi|\psi\rangle|^2\Big)^{\frac{1}{2}} \leq \|\varphi\|\,\|\psi\|,$$

so that the right hand side of (iii) is defined. By Lemma 2.2 (b) the set $\{\xi \in K\,|\,\langle\xi|\varphi\rangle \neq 0 \text{ or } \langle\xi|\psi\rangle \neq 0\}$ can be enumerated: $\xi_1, \xi_2, \ldots$. We may assume that this is an infinite sequence. Denote

$$\varphi_n = \sum_{k=1}^{n} \langle \xi_k | \varphi \rangle \xi_k \text{ and } \psi_n = \sum_{k=1}^{n} \langle \xi_k | \psi \rangle \xi_k, \text{ so that}$$

$$\langle \psi_n | \varphi_n \rangle = \sum_{k=1}^{n} \sum_{p=1}^{n} \langle \langle \xi_p | \psi \rangle \xi_p | \langle \xi_k | \varphi \rangle \xi_k \rangle = \sum_{k=1}^{n} \overline{\langle \xi_k | \psi \rangle} \langle \xi_k | \varphi \rangle.$$

When $n \to \infty$, by the continuity of the inner product we get (iii). Choosing $\varphi = \psi$ we see at once that (iii) implies (iv). One also sees immediately that (iv) implies (i) so that the conditions (i)–(iv) are equivalent. On the other hand, (ii) implies (v), and (v) implies (i), for if $\varphi \perp K$ then $\varphi \perp M$. □

Definition 2.6 We say that an orthonormal set $K \subset \mathcal{H}$ satisfying the equivalent conditions in the preceding theorem is an *orthonormal basis* or a *complete orthonormal system* in $\mathcal{H}$. The equations appearing in conditions (iii) and (iv) in Theorem 2.20 are called the *Parseval identities*. The numbers $\langle \xi | \varphi \rangle$, $\xi \in K$, are called the *Fourier coefficients* of φ with respect to K.

Every Hilbert space has an orthonormal basis. We prove this using the Zorn lemma. If the space is *separable* (i.e. if it has a countable dense subset), the use of Zorn's lemma (and other methods equivalent to the axiom of choice) could be avoided, but we will not elaborate this question any further.

Theorem 2.21 *If $L \subset \mathcal{H}$ is an orthonormal set, there is an orthonormal basis of $\mathcal{H}$ containing L.*

Proof We equip the set $\mathcal{F}$ of the orthonormal subsets of $\mathcal{H}$ containing L with the inclusion order. Every linearly ordered subset $\mathcal{F}_0$ of $\mathcal{F}$ has an upper bound, namely $\cup_{F \in \mathcal{F}_0} F$. According to Zorn's lemma $\mathcal{F}$ has a maximal element K. The maximality means that K satisfies condition (i) in Theorem 2.20. □

Theorem 2.22 *Let K be an orthonormal basis of $\mathcal{H}$. The following conditions are equivalent:*

(i) *the set K is at most countable;*
(ii) *the space $\mathcal{H}$ is separable.*

Proof Assume first (i). The set of the linear combinations of the elements of K with coefficients whose real and imaginary parts are rational, is countable. Using condition (v) in Theorem 2.20 and the density of the set of rational numbers in $\mathbb{R}$ we easily see that this set is dense in $\mathcal{H}$, and so (ii) holds. Next assume (ii). If $\varphi, \psi \in K$ and $\varphi \neq \psi$, then $\|\varphi - \psi\|^2 = 2$, so that the open balls $\mathrm{B}(\varphi, \frac{1}{2}\sqrt{2})$ and $\mathrm{B}(\psi, \frac{1}{2}\sqrt{2})$ are disjoint. Each one of these meets a certain at most countable set, and so (i) holds. □

Remark 2.5 If K and L are orthonormal bases of the same Hilbert space $\mathcal{H}$ then they have the same cardinality (exercise). This cardinality is called the *Hilbert dimension* of $\mathcal{H}$. If M is a closed subspace of $\mathcal{H}$ we call the Hilbert dimension of its orthogonal complement the *Hilbert codimension* of M. If $\mathcal{H}$ is finite-dimensional, then its Hilbert dimension is just its dimension in the usual algebraic sense. (This will follow, e.g., from Theorem 3.3.) However, if $\mathcal{H} = \ell^2$, its Hilbert dimension is the cardinality of $\mathbb{N}$, but it does have an uncountable linearly independent subset. Such a set is, for instance, the set $\left\{ \sum_{n=1}^{\infty} c^n e_n \,\middle|\, 0 < c < 1 \right\}$ where, for each $n \in \mathbb{N}$, e_n is the function taking the value 1 at n and zero elsewhere (exercise). $\triangleleft$

2.8 Direct Sums of Hilbert Spaces

In this section we assume that $\mathcal{I}$ is a set and $\mathcal{H}_i$ is a Hilbert space for every $i \in \mathcal{I}$. The Cartesian product $\prod_{i\in\mathcal{I}} \mathcal{H}_i$ consists of all vector families $(\varphi_i) = (\varphi_i)_{i\in\mathcal{I}}$ in $\cup_{i\in\mathcal{I}}\mathcal{H}_i$ having the property that $\varphi_i \in \mathcal{H}_i$ for all $i \in \mathcal{I}$. Clearly it is a vector space with respect to the pointwise operations: $\alpha(\varphi_i) = (\alpha\varphi_i)$, $(\varphi_i) + (\psi_i) = (\varphi_i + \psi_i)$. We now consider an important subspace.

Proposition 2.7 *The set* $E = \left\{(\varphi_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} \mathcal{H}_i \,\middle|\, \sum_{i\in\mathcal{I}} \|\varphi_i\|^2 < \infty\right\}$ *is a vector subspace of the Cartesian product* $\prod_{i\in\mathcal{I}} \mathcal{H}_i$*. If* (φ_i), $(\psi_i) \in E$*, then the family* $(\langle\, \varphi_i \mid \psi_i \,\rangle)_{i\in\mathcal{I}}$ *is summable. Moreover, the mapping* $h : E \times E \to \mathbb{C}$ *defined by the formula* $h((\varphi_i), (\psi_i)) = \sum_{i\in\mathcal{I}} \langle\, \varphi_i \mid \psi_i \,\rangle$ *is an inner product and equipped with this inner product* E *is a Hilbert space.*

We leave the rather straightforward proof as an exercise; the central ideas are already present in the case of the scalar sequence space ℓ^2. The Hilbert space E in the above proposition is called the *direct sum* or *Hilbert sum* of the family $(\mathcal{H}_i)_{i\in\mathcal{I}}$ of Hilbert spaces. We use the notations $\bigoplus_{i\in\mathcal{I}} \mathcal{H}_i$ and $\sum_{i\in\mathcal{I}}^{\oplus} \mathcal{H}_i$ for it.

Example 2.4 If we take $\mathcal{H}_i = \mathbb{C}$ for all $i \in \mathcal{I}$, then we denote $\bigoplus_{i\in\mathcal{I}} \mathcal{H}_i = \ell^2(\mathcal{I})$. If $\mathcal{I}$ happens to be an orthonormal basis of a Hilbert space $\mathcal{H}$, then the mapping $\varphi \mapsto \chi_{\{\varphi\}}$ from $\mathcal{I}$ to $\ell^2(\mathcal{I})$ (where $\chi_{\{\varphi\}}$ is the function taking the value 1 at φ and zero elsewhere) extends to an isometric isomorphism from $\mathcal{H}$ onto $\ell^2(\mathcal{I})$. This follows readily from Theorem 2.20 and Proposition 2.7. $\triangleleft$

Example 2.5 Let $\mathcal{H}$ be a Hilbert space and $(\mathcal{H}_i)_{i\in\mathcal{I}}$ a family of closed subspaces of $\mathcal{H}$ such that $\mathcal{H}_i \perp \mathcal{H}_j$ whenever $i \neq j$ and $\mathcal{H}$ itself is the only closed subspace containing every $\mathcal{H}_i$. For each $\varphi \in \mathcal{H}$ let $\Phi\varphi$ be the family $(P_i\varphi)_{i\in\mathcal{I}}$ where P_i is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_i$. Then Φ is an isometric isomorphism from $\mathcal{H}$ onto $\bigoplus_{i\in\mathcal{I}} \mathcal{H}_i$ (exercise). For this reason we may sometimes use the identification $\mathcal{H} = \bigoplus_{i\in\mathcal{I}} \mathcal{H}_i$. $\triangleleft$

2.9 Tensor Products of Hilbert Spaces

One can define the general algebraic tensor product of vector spaces or modules, but here we only consider the case of Hilbert spaces, and the result will be a Hilbert space which could be viewed as the completion of the algebraic tensor with respect to a natural inner product. We are not, however, going to develop this latter point of view now. We begin with a lemma showing the existence of the (Hilbert space) tensor product of two Hilbert spaces. To avoid trivialities we assume that our Hilbert spaces contain non-zero elements.

Lemma 2.3 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. There is a Hilbert space $\mathcal{H}_\otimes$ with a bilinear map $f : \mathcal{H} \times \mathcal{K} \to \mathcal{H}_\otimes$ such that*

(i) *the subspace generated by the set $\{ f(\varphi, \psi) \mid \varphi \in \mathcal{H},\ \psi \in \mathcal{K}\}$ is dense in $\mathcal{H}_\otimes$ and*
(ii) *$\langle f(\varphi_1, \psi_1) \mid f(\varphi_2, \psi_2) \rangle = \langle \varphi_1 \mid \varphi_2 \rangle \langle \psi_1 \mid \psi_2 \rangle$ for all $\varphi_1, \varphi_2 \in \mathcal{H}$ and $\psi_1, \psi_2 \in \mathcal{K}$.*

Proof Choose an orthonormal basis H (resp. K) for $\mathcal{H}$ (resp. $\mathcal{K}$). Consider the Hilbert space $\mathcal{H} = \ell^2(H \times K)$ of square-summable families $(c_{\xi,\eta})_{(\xi,\eta)\in H\times K}$. Define a map $f : \mathcal{H} \times \mathcal{K} \to \ell^2(H \times K)$ by setting

$$f(\varphi, \psi) = \big(\langle \xi \mid \varphi \rangle \langle \eta \mid \psi \rangle \big)_{(\xi,\eta)\in H\times K}$$

for all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$. The map f is clearly bilinear and it satisfies the condition (i) since $f(\xi, \eta) = \chi_{\{(\xi,\eta)\}}$ whenever $\xi \in H$ and $\eta \in K$. Using Theorem 2.20 we also get

$$\begin{aligned}
\langle f(\varphi_1, \psi_1) \mid f(\varphi_2, \psi_2) \rangle &= \sum_{(\xi,\eta)\in H\times K} \overline{\langle \xi \mid \varphi_1 \rangle \langle \eta \mid \psi_1 \rangle} \langle \xi \mid \varphi_2 \rangle \langle \eta \mid \psi_2 \rangle \\
&= \sum_{\xi\in H} \langle \varphi_1 \mid \xi \rangle \langle \xi \mid \varphi_2 \rangle \sum_{\eta\in K} \langle \psi_1 \mid \eta \rangle \langle \eta \mid \psi_2 \rangle \\
&= \langle \varphi_1 \mid \varphi_2 \rangle \langle \psi_1 \mid \psi_2 \rangle
\end{aligned}$$

for all $\varphi_1, \varphi_2 \in \mathcal{H}$ and $\psi_1, \psi_2 \in \mathcal{K}$, so that the condition (ii) also holds. □

Next we show that the above pair $(\mathcal{H}_\otimes, f)$ is essentially unique.

Lemma 2.4 *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and suppose that $\mathcal{H}_\otimes$ and f are as in Lemma 2.3. Suppose that also the pair $(\mathcal{H}', f')$ satisfies the conditions (i) and (ii) of Lemma 2.3. Then there is an isometric isomorphism $g : \mathcal{H}_\otimes \to \mathcal{H}'$ such that $g \circ f = f'$.*

Proof From the conditions (i) and (ii) of Lemma 2.3 it follows that

$$\Big\|\sum_{j=1}^{n} f(\varphi_j,\psi_j)\Big\|^2 = \sum_{j,k=1}^{n} \big\langle f(\varphi_j,\psi_j) \,|\, f(\varphi_k,\psi_k)\big\rangle$$
$$= \sum_{j,k=1}^{n} \big\langle \varphi_j \,|\, \varphi_k\big\rangle\big\langle \psi_j \,|\, \psi_k\big\rangle = \Big\|\sum_{j=1}^{n} f'(\varphi_j,\psi_j)\Big\|^2$$

for all $\varphi_j \in \mathcal{H}$ and $\psi_j \in \mathcal{K}$, $j = 1, \ldots, n$. Thus the map

$$\sum_{j=1}^{n} f(\varphi_j,\psi_j) \mapsto \sum_{j=1}^{n} f'(\varphi_j,\psi_j)$$

is a well-defined isometric isomorphism from a dense subspace of $\mathcal{H}_\otimes$ onto a dense subspace of $\mathcal{H}'$. We obtain the isomorphism g by extending the above map by continuity to the whole of $\mathcal{H}_\otimes$. □

Definition 2.7 Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. If $\mathcal{H}_\otimes$ is a Hilbert space and $f : \mathcal{H} \times \mathcal{K} \to \mathcal{H}_\otimes$ is a map satisfying the conditions (i) and (ii) of Lemma 2.3 we say that the pair $(\mathcal{H}_\otimes, f)$ or simply $\mathcal{H}_\otimes$ is a *Hilbert tensor product* of the spaces $\mathcal{H}$ and $\mathcal{K}$.

If there seems to be no danger of confusion, we may call the Hilbert tensor product $(\mathcal{H}_\otimes, f)$ (or $\mathcal{H}_\otimes$) simply the *tensor product* of $\mathcal{H}$ and $\mathcal{K}$. According to Lemma 2.4 the tensor product $(\mathcal{H}_\otimes, f)$ of $\mathcal{H}$ and $\mathcal{K}$ is uniquely defined up to an isometric isomorphism. We thus use the notations $\mathcal{H}_\otimes = \mathcal{H} \otimes \mathcal{K}$ and $f(\varphi,\psi) = \varphi \otimes \psi$ in the sequel. From the construction of the tensor product in Lemma 2.3 we also see that there is an isometric isomorphism $g : \mathcal{H} \otimes \mathcal{K} \to \mathcal{K} \otimes \mathcal{H}$ such that $g(\varphi \otimes \psi) = \psi \otimes \varphi$ for all $\varphi \in \mathcal{H}$ and $\psi \in \mathcal{K}$. If $K \subset \mathcal{H}$ and $L \subset \mathcal{K}$ are orthonormal bases, then according to the conditions (i) and (ii) of Lemma 2.3 the set $\{\xi \otimes \eta \,|\, \xi \in K,\ \eta \in L\}$ is an orthonormal basis of the tensor product $\mathcal{H} \otimes \mathcal{K}$. Note that although the so-called *simple tensors* $\varphi \otimes \psi$ (topologically) generate the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$, a vector $\zeta \in \mathcal{H} \otimes \mathcal{K}$ need not be a simple tensor. Any vector $\zeta \in \mathcal{H} \otimes \mathcal{K}$ can be represented as a linear combination of simple tensors or as a limit of such linear combinations. As was already mentioned, the Hilbert tensor product $\mathcal{H} \otimes \mathcal{K}$ is in fact the Hilbert space completion of the general algebraic tensor product of the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$.

One can also form, with essentially the same process as for two Hilbert spaces, Hilbert tensor products $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ for more than two Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$. It is also possible to define $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$ by induction by using the notion of the tensor product of two Hilbert space and the easily verified associativity result $(\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_3 = \mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3)$.

Remark 2.6 Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. Let K be an orthonormal basis of $\mathcal{K}$ and choose $\mathcal{H}_\mu = \mathcal{H}$ for every $\mu \in K$. There is an isometric isomorphism

$\Phi : \sum_{\mu\in K}^{\oplus} \mathcal{H}_\mu \to \mathcal{H}\otimes\mathcal{K}$ satisfying $\Phi(\xi^{(\nu)}) = \xi\otimes\nu$ for all $\nu\in K$ and $\xi\in\mathcal{H}$ where $\xi^{(\nu)}\in\sum_{\mu\in K}^{\oplus}\mathcal{H}_\mu$ is the family taking the value ξ at ν and zero elsewhere. We leave the verification as an exercise. Accordingly, we sometimes identify the tensor product $\mathcal{H}\otimes\mathcal{K}$ with the direct sum $\sum_{\mu\in K}^{\oplus}\mathcal{H}_\mu$. ◁

2.10 Exercises

Unless otherwise stated, $\mathcal{H}$ is an arbitrary Hilbert space.

1. Prove Theorem 2.2.
2. Prove Theorem 2.3.
3. Prove Theorem 2.6.
4. Prove Theorem 2.9.
5. Prove Theorem 2.11.
6. Prove Theorem 2.12.
7. Prove Theorem 2.13.
8. Prove by using the polarisation identity that a linear map $T : \mathcal{H}\to\mathcal{H}$ is isometric if and only if it preserves the inner products (i.e $\langle T\varphi \mid T\psi\rangle = \langle\varphi\mid\psi\rangle$ for all $\varphi,\ \psi\in\mathcal{H}$).
9. Show that an operator $T\in\mathcal{L}(\mathcal{H})$ is unitary if and only if T is isometric and surjective.
10. Let $T\in\mathcal{L}(\mathcal{H})$. Show that
$$\|T\| = \sup\left\{|\langle\varphi|T\psi\rangle| \;\middle|\; \|\varphi\|\le 1,\ \|\psi\|\le 1\right\}.$$
11. Let M be a closed subspace of $\mathcal{H}$. Let $f : M\to\mathbb{C}$ be a continuous linear functional. Show that there is one and only one continuous linear functional $\tilde{f} : \mathcal{H}\to\mathbb{C}$ such that $\tilde{f}|M = f$ and $\|\tilde{f}\| = \|f\|$. Show that $\tilde{f}(\varphi) = 0$ for all $\varphi\in M^\perp$.
12. Let $(T_i)_{i\in\mathcal{I}}$ be a net in $\mathcal{L}(\mathcal{H})$ and $T\in\mathcal{L}(\mathcal{H})$. Prove: If the net $(T_i)_{i\in\mathcal{I}}$ converges strongly to T, then it converges weakly to T.
13. Let $(\mathcal{I},\ge)$ be a directed set, let $S_i\in\mathcal{L}(\mathcal{H})$ and $T_i\in\mathcal{L}(\mathcal{H})$ for all $i\in\mathcal{I}$, and let $S,\ T\in\mathcal{L}(\mathcal{H})$. Prove: If $S_i\to^s S$, $T_i\to^s T$ and $\sup\left\{\|S_i\| \;\middle|\; i\in\mathcal{I}\right\}<\infty$, then $S_iT_i\to^s ST$.
14. Prove: If $T\in\mathcal{L}_s(\mathcal{H})^+$ then $T^n\in\mathcal{L}_s(\mathcal{H})^+$ for all $n\in\mathbb{N}$.
15. Consider the Hilbert space $\ell^2 = \left\{f : \mathbb{N}\to\mathbb{C} \;\middle|\; \sum_{k=1}^\infty |f(k)|^2 < \infty\right\}$ (equipped with the inner product $\langle f\mid g\rangle = \sum_{k=1}^\infty \overline{f(k)}g(k)$ as usual). Define $T : \ell^2\to\ell^2$ by setting $Tf(1) = 0$ and $Tf(k) = f(k-1)$ whenever $k\ge 2$.
(a) Show that T is isometric.
(b) Is T unitary?

16. (a) What is the adjoint T^* of the operator T defined in the previous exercise?
(b) Let T be as in the previous exercise. Does the sequence $(T^n)_{n\in\mathbb{N}}$ converge strongly? Does it converge weakly?
17. Define the mapping $T : \ell^2 \to \ell^2$ via the formula $(Tf)(n) = f(n+1)$. Show that the adjoint of T satisfies the formula $(T^*f)(n) = f(n-1)$ when $n \geq 2$, and $(T^*f)(1) = 0$.
18. We saw in Exercise 15 that the operator T defined there is isometric. Is T^* isometric? Is any one of the formulas $T^*T = I$, $TT^* = I$ true (when I is the identity operator of ℓ^2)? Is T normal?
19. Let P, $Q \in \mathcal{P}(\mathcal{H})$ (i.e. P and Q are projections). Denote $A_n = (PQP)^n$ for all $n \in \mathbb{N}$. Show that (A_n) is a decreasing sequence in $\mathcal{L}_s(\mathcal{H})^+$ and converges weakly and strongly to some $B \in \mathcal{L}_s(\mathcal{H})^+$.
20. We retain the notations and assumptions of the preceding exercise. Prove the following claims:
(a) $B \in \mathcal{P}(\mathcal{H})$;
(b) $BQP = B$;
(c) $B \leq Q$; (hint: show that $B(I-Q)[B(I-Q)]^* = 0$.)
(d) $PB = P$, i.e. $B \leq P$;
(e) $B = P \wedge Q$.
21. Assume that $(T_i)_{i\in\mathcal{I}}$ is a net of operators in $\mathcal{L}(\mathcal{H})$ and w-lim $T_i = T \in \mathcal{L}(\mathcal{H})$. Is T necessarily selfadjoint if each T_i is selfadjoint?
22. Let $g : \mathbb{N} \to \mathbb{C}$ be a bounded function and $T_g : \ell^2 \to \ell^2$ the operator defined by the formula $T_g f = gf$. Show that T_g is continuous and $\|T\| = \|g\|_\infty$ where $\|g\|_\infty = \sup_{n\in\mathbb{N}} |g(n)|$.
23. In the situation of the preceding exercise, give necessary and sufficient conditions for T_g to be
(a) normal,
(b) selfadjoint,
(c) unitary,
(d) a projection.
24. We use the notation of Exercise 22. Let $(\mathcal{I}, \geq)$ be a directed set and $g_i : \mathbb{N} \to \mathbb{C}$ a bounded function for each $i \in \mathcal{I}$. Assume that $\sup_{i\in\mathcal{I}} \|g_i\|_\infty < \infty$. Show that for a bounded function $g : \mathbb{N} \to \mathbb{C}$ we have s-lim $T_{g_i} = T_g$, if and only if lim $g_i(n) = g(n)$ for all $n \in \mathbb{N}$.
25. Is the claim in the previous exercise true without the assumption that $\sup_{i\in\mathcal{I}} \|g_i\|_\infty < \infty$?
26. Show that the absolute value $|T|$ of an operator $T \in \mathcal{L}(\mathcal{H})$ is the unique positive operator A satisfying $\|A\varphi\| = \|T\varphi\|$ for all $\varphi \in \mathcal{H}$.
27. Let $g : \mathbb{N} \to \mathbb{C}$ be a bounded function and $T_g \in \mathcal{L}(\mathcal{H})$ as in Exercise 22. Describe the polar decomposition of the operator T_g.
28. Describe the polar decomposition of the operator T defined in Exercise 17. Describe also the polar decomposition of T^*.
29. Let $T = VA$ be the polar decomposition of the operator $T \in \mathcal{L}(\mathcal{H})$. Give a necessary and sufficient condition concerning T

(a) for V to be isometric,
(b) for V to be unitary.

30. Show that there is an isometric bijection $J : H \to H$ which is conjugate linear, i.e. $J(\alpha\varphi + \beta\psi) = \overline{\alpha}J\varphi + \overline{\beta}J\psi$ for all $\varphi, \psi \in H$ and $\alpha, \beta \in \mathbb{C}$. (Hint: "practise" with ℓ^2.)
31. Show using the preceding exercise that there is an isometric linear bijection $\Phi : \mathcal{H} \to \mathcal{H}^*$, where $\mathcal{H}^*$ is the dual of $\mathcal{H}$, i.e. the Banach space of continuous linear functionals $f : \mathcal{H} \to \mathbb{C}$.
32. Prove Proposition 2.7 except for the completeness of E.
33. Prove that the inner product space in Proposition 2.7 is complete.
34. Prove the claim made in Example 2.4.
35. Prove the statement left as an exercise in Example 2.5.
36. Complete the details of Remark 2.6.
37. Show that if K and L are orthonormal bases of the same Hilbert space $\mathcal{H}$, then they have the same cardinality.
38. Prove the last statement in Remark 2.5.

Reference

1. Luxemburg, W.A.J., Zaanen, A.C.: Riesz Spaces, vol. I. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., New York (1971) (North-Holland Mathematical Library)

http://www.springer.com/978-3-319-43387-5

Quantum Measurement
Busch, P.; Lahti, P.J.; Pellonpää, J.-P.; Ylinen, K.
2016, XII, 542 p. 7 illus., 2 illus. in color., Hardcover
ISBN: 978-3-319-43387-5