14

Hilbert Spaces Basics

(BRUCE: Perhaps this should be move to between Chapters 7 & 8?)

Definition 14.1. *Let H be a complex vector space. An inner product on H is a function, $\langle\cdot|\cdot\rangle : H \times H \to \mathbb{C}$, such that*

1. $\langle ax + by|z\rangle = a\langle x|z\rangle + b\langle y|z\rangle$ *i.e.* $x \to \langle x|z\rangle$ *is linear.*
2. $\overline{\langle x|y\rangle} = \langle y|x\rangle$.
3. $\|x\|^2 := \langle x|x\rangle \geq 0$ *with equality* $\|x\|^2 = 0$ *iff* $x = 0$.

Notice that combining properties (1) and (2) that $x \to \langle z|x\rangle$ is anti-linear for fixed $z \in H$, i.e.

$$\langle z|ax + by\rangle = \bar{a}\langle z|x\rangle + \bar{b}\langle z|y\rangle.$$

The following identity will be used frequently in the sequel without further mention,

$$\begin{aligned}\|x+y\|^2 &= \langle x+y|x+y\rangle = \|x\|^2 + \|y\|^2 + \langle x|y\rangle + \langle y|x\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x|y\rangle. \end{aligned} \tag{14.1}$$

Theorem 14.2 (Schwarz Inequality). *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space, then for all $x, y \in H$*

$$|\langle x|y\rangle| \leq \|x\|\|y\|$$

and equality holds iff x and y are linearly dependent.

Proof. If $y = 0$, the result holds trivially. So assume that $y \neq 0$ and observe; if $x = \alpha y$ for some $\alpha \in \mathbb{C}$, then $\langle x|y\rangle = \alpha\|y\|^2$ and hence

$$|\langle x|y\rangle| = |\alpha|\,\|y\|^2 = \|x\|\|y\|.$$

Now suppose that $x \in H$ is arbitrary, let $z := x - \|y\|^{-2}\langle x|y\rangle y$. (So z is the "orthogonal projection" of x onto y, see Figure 14.1.) Then

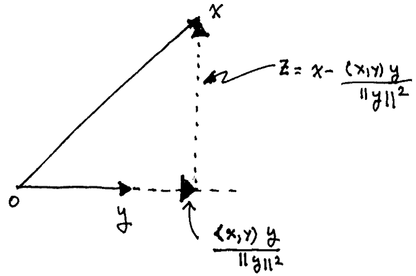

Fig. 14.1. The picture behind the proof of the Schwarz inequality.

$$\begin{aligned}0 \leq \|z\|^2 &= \left\| x - \frac{\langle x|y\rangle}{\|y\|^2}y \right\|^2 = \|x\|^2 + \frac{|\langle x|y\rangle|^2}{\|y\|^4}\|y\|^2 - 2\mathrm{Re}\langle x|\frac{\langle x|y\rangle}{\|y\|^2}y\rangle \\ &= \|x\|^2 - \frac{|\langle x|y\rangle|^2}{\|y\|^2}\end{aligned}$$

from which it follows that $0 \leq \|y\|^2\|x\|^2 - |\langle x|y\rangle|^2$ with equality iff $z = 0$ or equivalently iff $x = \|y\|^{-2}\langle x|y\rangle y$. ■

Corollary 14.3. *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space and $\|x\| := \sqrt{\langle x|x\rangle}$. Then the* ***Hilbertian norm****, $\|\cdot\|$, is a norm on H. Moreover $\langle\cdot|\cdot\rangle$ is continuous on $H \times H$, where H is viewed as the normed space $(H, \|\cdot\|)$.*

Proof. If $x, y \in H$, then, using the Schwarz's inequality,

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

Taking the square root of this inequality shows $\|\cdot\|$ satisfies the triangle inequality. Checking that $\|\cdot\|$ satisfies the remaining axioms of a norm is not routine and will be left to the reader. If $x, x'y, y' \in H$, then

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

from which it follows that $\langle\cdot|\cdot\rangle$ is continuous. ■

Definition 14.4. *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space, we say $x, y \in H$ are* ***orthogonal*** *and write $x \perp y$ iff $\langle x|y\rangle = 0$. More generally if $A \subset H$ is a set, $x \in H$ is* ***orthogonal to*** *A (write $x \perp A$) iff $\langle x|y\rangle = 0$ for all $y \in A$. Let $A^\perp = \{x \in H : x \perp A\}$ be the set of vectors orthogonal to A. A subset $S \subset H$ is an* ***orthogonal set*** *if $x \perp y$ for all distinct elements $x, y \in S$. If S further satisfies, $\|x\| = 1$ for all $x \in S$, then S is said to be* ***orthonormal set.***

Proposition 14.5. *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space then*

1. ***(Parallelogram Law)***
$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{14.2}$$
for all $x, y \in H$.
2. ***(Pythagorean Theorem)*** *If $S \subset H$ is a finite orthogonal set, then*
$$\left\| \sum_{x\in S} x \right\|^2 = \sum_{x\in S} \|x\|^2. \tag{14.3}$$
3. *If $A \subset H$ is a set, then $A^\perp$ is a* ***closed*** *linear subspace of H.*

Remark 14.6. See Proposition 14.54 for the "converse" of the parallelogram law.

Proof. I will assume that H is a complex Hilbert space, the real case being easier. Items 1. and 2. are proved by the following elementary computations;

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

and

$$\begin{aligned}
\left\| \sum_{x\in S} x \right\|^2 &= \langle \sum_{x\in S} x | \sum_{y\in S} y\rangle = \sum_{x,y\in S} \langle x|y\rangle \\
&= \sum_{x\in S} \langle x|x\rangle = \sum_{x\in S} \|x\|^2.
\end{aligned}$$

Item 3. is a consequence of the continuity of $\langle\cdot|\cdot\rangle$ and the fact that

$$A^\perp = \cap_{x\in A} \mathrm{Nul}(\langle\cdot|x\rangle)$$

where $\mathrm{Nul}(\langle\cdot|x\rangle) = \{y \in H : \langle y|x\rangle = 0\}$ – a closed subspace of H. ■

Definition 14.7. *A* ***Hilbert space*** *is an inner product space $(H, \langle\cdot|\cdot\rangle)$ such that the induced Hilbertian norm is complete.*

Example 14.8. Suppose X is a set and $\mu : X \to (0, \infty)$, then $H := \ell^2(\mu)$ is a Hilbert space when equipped with the inner product,

$$\langle f|g\rangle := \sum_{x\in X} f(x)\, \bar{g}(x)\, \mu(x).$$

In Exercise 14.6 you will show every Hilbert space H is "equivalent" to a Hilbert space of this form with $\mu \equiv 1$.

More example of Hilbert spaces will be given later after we develop the Lebesgue integral, see Example 23.1 below.

Definition 14.9. *A subset C of a vector space X is said to be convex if for all $x, y \in C$ the line segment $[x, y] := \{tx + (1-t)y : 0 \le t \le 1\}$ joining x to y is contained in C as well. (Notice that any vector subspace of X is convex.)*

Theorem 14.10. *Suppose that H is a Hilbert space and $M \subset H$ is a closed convex subset of H. Then for any $x \in H$ there exists a unique $y \in M$ such that*

$$\|x-y\| = d(x, M) = \inf_{z\in M} \|x-z\|.$$

Moreover, if M is a vector subspace of H, then the point y may also be characterized as the unique point in M such that $(x-y) \perp M$.

Proof. Uniqueness. By replacing M by $M - x := \{m - x : m \in M\}$ we may assume $x = 0$. Let $\delta := d(0, M) = \inf_{m\in M} \|m\|$ and $y, z \in M$, see Figure 14.2.

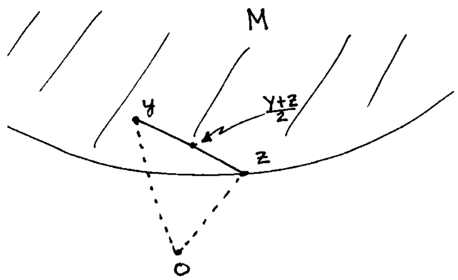

Fig. 14.2. The geometry of convex sets.

By the parallelogram law and the convexity of M,

$$\begin{aligned}
2\|y\|^2 + 2\|z\|^2 &= \|y+z\|^2 + \|y-z\|^2 \\
&= 4\|\frac{y+z}{2}\|^2 + \|y-z\|^2 \ge 4\delta^2 + \|y-z\|^2.
\end{aligned} \tag{14.4}$$

Hence if $\|y\| = \|z\| = \delta$, then $2\delta^2 + 2\delta^2 \ge 4\delta^2 + \|y-z\|^2$, so that $\|y-z\|^2 = 0$. Therefore, if a minimizer for $d(0,\cdot)|_M$ exists, it is unique.

Existence. Let $y_n \in M$ be chosen such that $\|y_n\| = \delta_n \to \delta \equiv d(0, M)$. Taking $y = y_m$ and $z = y_n$ in Eq. (14.4) shows

$$2\delta_m^2 + 2\delta_n^2 \ge 4\delta^2 + \|y_n - y_m\|^2.$$

Passing to the limit $m, n \to \infty$ in this equation implies,

$$2\delta^2 + 2\delta^2 \ge 4\delta^2 + \limsup_{m,n\to\infty} \|y_n - y_m\|^2,$$

i.e. $\limsup_{m,n\to\infty} \|y_n - y_m\|^2 = 0$. Therefore, by completeness of H, $\{y_n\}_{n=1}^\infty$ is convergent. Because M is closed, $y := \lim_{n\to\infty} y_n \in M$ and because the norm is continuous,

$$\|y\| = \lim_{n\to\infty} \|y_n\| = \delta = d(0, M).$$

So y is the desired point in M which is closest to 0.

Now suppose M is a closed subspace of H and $x \in H$. Let $y \in M$ be the closest point in M to x. Then for $w \in M$, the function

$$g(t) := \|x - (y + tw)\|^2 = \|x - y\|^2 - 2t\mathrm{Re}\langle x - y|w\rangle + t^2\|w\|^2$$

has a minimum at $t = 0$ and therefore $0 = g'(0) = -2\mathrm{Re}\langle x - y|w\rangle$. Since $w \in M$ is arbitrary, this implies that $(x - y) \perp M$.

Finally suppose $y \in M$ is any point such that $(x - y) \perp M$. Then for $z \in M$, by Pythagorean's theorem,

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

which shows $d(x, M)^2 \ge \|x - y\|^2$. That is to say y is the point in M closest to x. ∎

Definition 14.11. *Suppose that $A : H \to H$ is a bounded operator. The* ***adjoint*** *of A, denote A^*, is the unique operator $A^* : H \to H$ such that $\langle Ax|y\rangle = \langle x|A^*y\rangle$. (The proof that A^* exists and is unique will be given in Proposition 14.16 below.) A bounded operator $A : H \to H$ is* ***self - adjoint*** *or* ***Hermitian*** *if $A = A^*$.*

Definition 14.12. *Let H be a Hilbert space and $M \subset H$ be a closed subspace. The orthogonal projection of H onto M is the function $P_M : H \to H$ such that for $x \in H$, $P_M(x)$ is the unique element in M such that $(x - P_M(x)) \perp M$.*

Theorem 14.13 (Projection Theorem). *Let H be a Hilbert space and $M \subset H$ be a closed subspace. The orthogonal projection P_M satisfies:*

1. *P_M is linear and hence we will write $P_M x$ rather than $P_M(x)$.*
2. *$P_M^2 = P_M$ (P_M is a projection).*
3. *$P_M^* = P_M$, (P_M is self-adjoint).*
4. *$\mathrm{Ran}(P_M) = M$ and $\mathrm{Nul}(P_M) = M^\perp$.*

Proof.

1. Let $x_1, x_2 \in H$ and $\alpha \in \mathbb{F}$, then $P_M x_1 + \alpha P_M x_2 \in M$ and

$$P_M x_1 + \alpha P_M x_2 - (x_1 + \alpha x_2) = [P_M x_1 - x_1 + \alpha(P_M x_2 - x_2)] \in M^\perp$$

showing $P_M x_1 + \alpha P_M x_2 = P_M(x_1 + \alpha x_2)$, i.e. P_M is linear.

2. Obviously $\mathrm{Ran}(P_M) = M$ and $P_M x = x$ for all $x \in M$. Therefore $P_M^2 = P_M$.
3. Let $x, y \in H$, then since $(x - P_M x)$ and $(y - P_M y)$ are in $M^\perp$,

$$\begin{aligned}\langle P_M x|y\rangle &= \langle P_M x|P_M y + y - P_M y\rangle = \langle P_M x|P_M y\rangle \\ &= \langle P_M x + (x - P_M x)|P_M y\rangle = \langle x|P_M y\rangle.\end{aligned}$$

4. We have already seen, $\mathrm{Ran}(P_M) \subset M$ and $P_M x = 0$ iff $x = x - 0 \in M^\perp$, i.e. $\mathrm{Nul}(P_M) = M^\perp$.

∎

Corollary 14.14. *If $M \subset H$ is a proper closed subspace of a Hilbert space H, then $H = M \oplus M^\perp$.*

Proof. Given $x \in H$, let $y = P_M x$ so that $x - y \in M^\perp$. Then $x = y + (x - y) \in M + M^\perp$. If $x \in M \cap M^\perp$, then $x \perp x$, i.e. $\|x\|^2 = \langle x|x\rangle = 0$. So $M \cap M^\perp = \{0\}$. ∎

Exercise 14.1. Suppose M is a subset of H, then $M^{\perp\perp} = \overline{\mathrm{span}(M)}$.

Theorem 14.15 (Riesz Theorem). *Let H^* be the dual space of H (Notation 7.9). The map*

$$z \in H \xrightarrow{j} \langle \cdot|z\rangle \in H^* \tag{14.5}$$

is a conjugate linear[1] *isometric isomorphism.*

Proof. The map j is conjugate linear by the axioms of the inner products. Moreover, for $x, z \in H$,

$$|\langle x|z\rangle| \le \|x\| \, \|z\| \text{ for all } x \in H$$

with equality when $x = z$. This implies that $\|jz\|_{H^*} = \|\langle\cdot|z\rangle\|_{H^*} = \|z\|$. Therefore j is isometric and this implies j is injective. To finish the proof we must show that j is surjective. So let $f \in H^*$ which we assume, with out loss of generality, is non-zero. Then $M = \mathrm{Nul}(f)$ – a closed proper subspace of H. Since, by Corollary 14.14, $H = M \oplus M^\perp$, $f : H/M \cong M^\perp \to \mathbb{F}$ is a linear isomorphism. This shows that $\dim(M^\perp) = 1$ and hence $H = M \oplus \mathbb{F}x_0$ where $x_0 \in M^\perp \setminus \{0\}$.[2] Choose $z = \lambda x_0 \in M^\perp$ such that $f(x_0) = \langle x_0|z\rangle$, i.e. $\lambda = \bar{f}(x_0)/\|x_0\|^2$. Then for $x = m + \lambda x_0$ with $m \in M$ and $\lambda \in \mathbb{F}$,

$$f(x) = \lambda f(x_0) = \lambda\langle x_0|z\rangle = \langle \lambda x_0|z\rangle = \langle m + \lambda x_0|z\rangle = \langle x|z\rangle$$

which shows that $f = jz$. ∎

[1] Recall that j is conjugate linear if

$$j(z_1 + \alpha z_2) = jz_1 + \bar{\alpha} jz_2$$

for all $z_1, z_2 \in H$ and $\alpha \in \mathbb{C}$.

[2] Alternatively, choose $x_0 \in M^\perp \setminus \{0\}$ such that $f(x_0) = 1$. For $x \in M^\perp$ we have $f(x - \lambda x_0) = 0$ provided that $\lambda := f(x)$. Therefore $x - \lambda x_0 \in M \cap M^\perp = \{0\}$, i.e. $x = \lambda x_0$. This again shows that $M^\perp$ is spanned by x_0.

Proposition 14.16 (Adjoints). *Let H and K be Hilbert spaces and $A : H \to K$ be a bounded operator. Then there exists a unique bounded operator $A^* : K \to H$ such that*

$$\langle Ax|y\rangle_K = \langle x|A^*y\rangle_H \text{ for all } x \in H \text{ and } y \in K. \tag{14.6}$$

Moreover, for all $A, B \in L(H,K)$ and $\lambda \in \mathbb{C}$,

1. $(A+\lambda B)^* = A^* + \bar{\lambda}B^*$,
2. $A^{**} := (A^*)^* = A$,
3. $\|A^*\| = \|A\|$ *and*
4. $\|A^*A\| = \|A\|^2$.
5. *If $K = H$, then $(AB)^* = B^*A^*$. In particular $A \in L(H)$ has a bounded inverse iff A^* has a bounded inverse and $(A^*)^{-1} = \left(A^{-1}\right)^*$.*

Proof. For each $y \in K$, the map $x \to \langle Ax|y\rangle_K$ is in H^* and therefore there exists, by Theorem 14.15, a unique vector $z \in H$ such that

$$\langle Ax|y\rangle_K = \langle x|z\rangle_H \text{ for all } x \in H.$$

This shows there is a unique map $A^* : K \to H$ such that $\langle Ax|y\rangle_K = \langle x|A^*(y)\rangle_H$ for all $x \in H$ and $y \in K$.

To see A^* is linear, let $y_1, y_2 \in K$ and $\lambda \in \mathbb{C}$, then for any $x \in H$,

$$\begin{aligned}\langle Ax|y_1+\lambda y_2\rangle_K &= \langle Ax|y_1\rangle_K + \bar{\lambda}\langle Ax|y_2\rangle_K \\ &= \langle x|A^*(y_1)\rangle_K + \bar{\lambda}\langle x|A^*(y_2)\rangle_K \\ &= \langle x|A^*(y_1)+\lambda A^*(y_2)\rangle_K\end{aligned}$$

and by the uniqueness of $A^*(y_1+\lambda y_2)$ we find

$$A^*(y_1+\lambda y_2) = A^*(y_1) + \lambda A^*(y_2).$$

This shows A^* is linear and so we will now write A^*y instead of $A^*(y)$.

Since

$$\langle A^*y|x\rangle_H = \overline{\langle x|A^*y\rangle_H} = \overline{\langle Ax|y\rangle_K} = \langle y|Ax\rangle_K$$

it follows that $A^{**} = A$. The assertion that $(A+\lambda B)^* = A^* + \bar{\lambda}B^*$ is Exercise 14.2.

Making use of the Schwarz inequality (Theorem 14.2), we have

$$\begin{aligned}\|A^*\| &= \sup_{k\in K:\|k\|=1}\|A^*k\| \\ &= \sup_{k\in K:\|k\|=1}\sup_{h\in H:\|h\|=1}|\langle A^*k|h\rangle| \\ &= \sup_{h\in H:\|h\|=1}\sup_{k\in K:\|k\|=1}|\langle k|Ah\rangle| = \sup_{h\in H:\|h\|=1}\|Ah\| = \|A\|.\end{aligned}$$

The last item is a consequence of the following two inequalities;

$$\|A^*A\| \le \|A^*\|\,\|A\| = \|A\|^2$$

and

$$\begin{aligned}\|A\|^2 &= \sup_{h\in H:\|h\|=1}\|Ah\|^2 = \sup_{h\in H:\|h\|=1}|\langle Ah|Ah\rangle| \\ &= \sup_{h\in H:\|h\|=1}|\langle h|A^*Ah\rangle| \le \sup_{h\in H:\|h\|=1}\|A^*Ah\| = \|A^*A\|.\end{aligned}$$

Now suppose that $K = H$. Then

$$\langle ABh|k\rangle = \langle Bh|A^*k\rangle = \langle h|B^*A^*k\rangle$$

which shows $(AB)^* = B^*A^*$. If A^{-1} exists then

$$\begin{aligned}\left(A^{-1}\right)^* A^* &= \left(AA^{-1}\right)^* = I^* = I \text{ and} \\ A^*\left(A^{-1}\right)^* &= \left(A^{-1}A\right)^* = I^* = I.\end{aligned}$$

This shows that A^* is invertible and $(A^*)^{-1} = \left(A^{-1}\right)^*$. Similarly if A^* is invertible then so is $A = A^{**}$. ∎

Exercise 14.2. Let H, K, M be Hilbert spaces, $A, B \in L(H,K)$, $C \in L(K,M)$ and $\lambda \in \mathbb{C}$. Show $(A+\lambda B)^* = A^* + \bar{\lambda}B^*$ and $(CA)^* = A^*C^* \in L(M,H)$.

Exercise 14.3. Let $H = \mathbb{C}^n$ and $K = \mathbb{C}^m$ equipped with the usual inner products, i.e. $\langle z|w\rangle_H = z\cdot\bar{w}$ for $z, w \in H$. Let A be an $m\times n$ matrix thought of as a linear operator from H to K. Show the matrix associated to $A^* : K \to H$ is the conjugate transpose of A.

Lemma 14.17. *Suppose $A : H \to K$ is a bounded operator, then:*

1. $\mathrm{Nul}(A^*) = \mathrm{Ran}(A)^\perp$.
2. $\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^\perp$.
3. *if $K = H$ and $V \subset H$ is an A – invariant subspace (i.e. $A(V) \subset V$), then $V^\perp$ is A^* – invariant.*

Proof. An element $y \in K$ is in $\mathrm{Nul}(A^*)$ iff $0 = \langle A^*y|x\rangle = \langle y|Ax\rangle$ for all $x \in H$ which happens iff $y \in \mathrm{Ran}(A)^\perp$. Because, by Exercise 14.1, $\overline{\mathrm{Ran}(A)} = \mathrm{Ran}(A)^{\perp\perp}$, and so by the first item, $\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^\perp$. Now suppose $A(V) \subset V$ and $y \in V^\perp$, then

$$\langle A^*y|x\rangle = \langle y|Ax\rangle = 0 \text{ for all } x \in V$$

which shows $A^*y \in V^\perp$. ∎

14.1 Hilbert Space Basis

Proposition 14.18 (Bessel's Inequality). *Let T be an orthonormal set, then for any $x \in H$,*

$$\sum_{v\in T} |\langle x|v\rangle|^2 \le \|x\|^2 \text{ for all } x \in H. \tag{14.7}$$

In particular the set $T_x := \{v \in T : \langle x|v\rangle \neq 0\}$ is at most countable for all $x \in H$.

Proof. Let $\Gamma \subset T$ be any finite set. Then

$$0 \le \|x - \sum_{v\in\Gamma}\langle x|v\rangle v\|^2 = \|x\|^2 - 2\mathrm{Re}\sum_{v\in\Gamma}\langle x|v\rangle\ \langle v|x\rangle + \sum_{v\in\Gamma}|\langle x|v\rangle|^2$$
$$= \|x\|^2 - \sum_{v\in\Gamma}|\langle x|v\rangle|^2$$

showing that $\sum_{v\in\Gamma}|\langle x|v\rangle|^2 \le \|x\|^2$. Taking the supremum of this inequality over $\Gamma \subset\subset T$ then proves Eq. (14.7). ∎

Proposition 14.19. *Suppose $T \subset H$ is an orthogonal set. Then $s = \sum_{v\in T} v$ exists in H (see Definition 7.15) iff $\sum_{v\in T}\|v\|^2 < \infty$. (In particular T must be at most a countable set.) Moreover, if $\sum_{v\in T}\|v\|^2 < \infty$, then*

1. *$\|s\|^2 = \sum_{v\in T}\|v\|^2$ and*
2. *$\langle s|x\rangle = \sum_{v\in T}\langle v|x\rangle$ for all $x \in H$.*

Similarly if $\{v_n\}_{n=1}^\infty$ is an orthogonal set, then $s = \sum_{n=1}^\infty v_n$ exists in H iff $\sum_{n=1}^\infty \|v_n\|^2 < \infty$. In particular if $\sum_{n=1}^\infty v_n$ exists, then it is independent of rearrangements of $\{v_n\}_{n=1}^\infty$.

Proof. Suppose $s = \sum_{v\in T} v$ exists. Then there exists $\Gamma \subset\subset T$ such that

$$\sum_{v\in\Lambda}\|v\|^2 = \left\|\sum_{v\in\Lambda} v\right\|^2 \le 1$$

for all $\Lambda \subset\subset T \setminus \Gamma$, wherein the first inequality we have used Pythagorean's theorem. Taking the supremum over such Λ shows that $\sum_{v\in T\setminus\Gamma}\|v\|^2 \le 1$ and therefore

$$\sum_{v\in T}\|v\|^2 \le 1 + \sum_{v\in\Gamma}\|v\|^2 < \infty.$$

Conversely, suppose that $\sum_{v\in T}\|v\|^2 < \infty$. Then for all $\varepsilon > 0$ there exists $\Gamma_\varepsilon \subset\subset T$ such that if $\Lambda \subset\subset T \setminus \Gamma_\varepsilon$,

$$\left\|\sum_{v\in\Lambda} v\right\|^2 = \sum_{v\in\Lambda}\|v\|^2 < \varepsilon^2. \tag{14.8}$$

Hence by Lemma 7.16, $\sum_{v\in T} v$ exists.

For item 1, let Γ_ε be as above and set $s_\varepsilon := \sum_{v\in\Gamma_\varepsilon} v$. Then

$$|\|s\| - \|s_\varepsilon\|| \le \|s - s_\varepsilon\| < \varepsilon$$

and by Eq. (14.8),

$$0 \le \sum_{v\in T}\|v\|^2 - \|s_\varepsilon\|^2 = \sum_{v\notin\Gamma_\varepsilon}\|v\|^2 \le \varepsilon^2.$$

Letting $\varepsilon \downarrow 0$ we deduce from the previous two equations that $\|s_\varepsilon\| \to \|s\|$ and $\|s_\varepsilon\|^2 \to \sum_{v\in T}\|v\|^2$ as $\varepsilon \downarrow 0$ and therefore $\|s\|^2 = \sum_{v\in T}\|v\|^2$.

Item 2. is a special case of Lemma 7.16. For the final assertion, let $s_N := \sum_{n=1}^N v_n$ and suppose that $\lim_{N\to\infty} s_N = s$ exists in H and in particular $\{s_N\}_{N=1}^\infty$ is Cauchy. So for $N > M$.

$$\sum_{n=M+1}^N \|v_n\|^2 = \|s_N - s_M\|^2 \to 0 \text{ as } M, N \to \infty$$

which shows that $\sum_{n=1}^\infty \|v_n\|^2$ is convergent, i.e. $\sum_{n=1}^\infty \|v_n\|^2 < \infty$.

Alternative proof of item 1. We could use the last result to prove Item 1. Indeed, if $\sum_{v\in T}\|v\|^2 < \infty$, then T is countable and so we may writer $T = \{v_n\}_{n=1}^\infty$. Then $s = \lim_{N\to\infty} s_N$ with s_N as above. Since the norm, $\|\cdot\|$, is continuous on H,

$$\|s\|^2 = \lim_{N\to\infty}\|s_N\|^2 = \lim_{N\to\infty}\left\|\sum_{n=1}^N v_n\right\|^2 = \lim_{N\to\infty}\sum_{n=1}^N\|v_n\|^2$$
$$= \sum_{n=1}^\infty \|v_n\|^2 = \sum_{v\in T}\|v\|^2.$$

∎

Corollary 14.20. *Suppose H is a Hilbert space, $\beta \subset H$ is an orthonormal set and $M = \overline{\mathrm{span}\ \beta}$. Then*

$$P_M x = \sum_{u\in\beta}\langle x|u\rangle u, \tag{14.9}$$

$$\sum_{u\in\beta}|\langle x|u\rangle|^2 = \|P_M x\|^2 \text{ and} \tag{14.10}$$

$$\sum_{u\in\beta}\langle x|u\rangle\langle u|y\rangle = \langle P_M x|y\rangle \tag{14.11}$$

for all $x, y \in H$.

Proof. By Bessel's inequality, $\sum_{u\in\beta}|\langle x|u\rangle|^2 \le \|x\|^2$ for all $x\in H$ and hence by Proposition 14.18, $Px := \sum_{u\in\beta}\langle x|u\rangle u$ exists in H and for all $x, y \in H$,

$$\langle Px|y\rangle = \sum_{u\in\beta}\langle\langle x|u\rangle u|y\rangle = \sum_{u\in\beta}\langle x|u\rangle\langle u|y\rangle. \tag{14.12}$$

Taking $y\in\beta$ in Eq. (14.12) gives $\langle Px|y\rangle = \langle x|y\rangle$, i.e. that $\langle x - Px|y\rangle = 0$ for all $y \in \beta$. So $(x - Px) \perp \operatorname{span}\beta$ and by continuity we also have $(x - Px) \perp M = \overline{\operatorname{span}\beta}$. Since Px is also in M, it follows from the definition of P_M that $Px = P_M x$ proving Eq. (14.9). Equations (14.10) and (14.11) now follow from (14.12), Proposition 14.19 and the fact that $\langle P_M x|y\rangle = \langle P_M^2 x|y\rangle = \langle P_M x|P_M y\rangle$ for all $x, y \in H$. ∎

Definition 14.21 (Basis). *Let H be a Hilbert space. A* ***basis*** *β of H is a maximal orthonormal subset $\beta\subset H$.*

Proposition 14.22. *Every Hilbert space has an orthonormal basis.*

Proof. Let $\mathcal{F}$ be the collection of all orthonormal subsets of H ordered by inclusion. If $\Phi\subset\mathcal{F}$ is linearly ordered then $\cup\Phi$ is an upper bound. By Zorn's Lemma (see Theorem B.7) there exists a maximal element $\beta\in\mathcal{F}$. ∎

An orthonormal set $\beta\subset H$ is said to be **complete** if $\beta^\perp = \{0\}$. That is to say if $\langle x|u\rangle = 0$ for all $u\in\beta$ then $x = 0$.

Lemma 14.23. *Let β be an orthonormal subset of H then the following are equivalent:*

1. *β is a basis,*
2. *β is complete and*
3. *$\overline{\operatorname{span}\beta} = H$.*

Proof. If β is not complete, then there exists a unit vector $x\in\beta^\perp\setminus\{0\}$. The set $\beta\cup\{x\}$ is an orthonormal set properly containing β, so β is not maximal. Conversely, if β is not maximal, there exists an orthonormal set $\beta_1\subset H$ such that $\beta \subsetneqq \beta_1$. Then if $x\in\beta_1\setminus\beta$, we have $\langle x|u\rangle = 0$ for all $u\in\beta$ showing β is not complete. This proves the equivalence of (1) and (2). If β is not complete and $x\in\beta^\perp\setminus\{0\}$, then $\overline{\operatorname{span}\beta}\subset x^\perp$ which is a proper subspace of H. Conversely if $\overline{\operatorname{span}\beta}$ is a proper subspace of H, $\beta^\perp = \overline{\operatorname{span}\beta}^\perp$ is a non-trivial subspace by Corollary 14.14 and β is not complete. This shows that (2) and (3) are equivalent. ∎

Theorem 14.24. *Let $\beta\subset H$ be an orthonormal set. Then the following are equivalent:*

1. *β is complete, i.e. β is an orthonormal basis for H.*
2. *$x = \sum_{u\in\beta}\langle x|u\rangle u$ for all $x\in H$.*
3. *$\langle x|y\rangle = \sum_{u\in\beta}\langle x|u\rangle\,\langle u|y\rangle$ for all $x, y\in H$.*
4. *$\|x\|^2 = \sum_{u\in\beta}|\langle x|u\rangle|^2$ for all $x\in H$.*

Proof. Let $M = \overline{\operatorname{span}\beta}$ and $P = P_M$.

(1) ⇒ (2) By Corollary 14.20, $\sum_{u\in\beta}\langle x|u\rangle u = P_M x$. Therefore

$$x - \sum_{u\in\beta}\langle x|u\rangle u = x - P_M x \in M^\perp = \beta^\perp = \{0\}.$$

(2) ⇒ (3) is a consequence of Proposition 14.19.

(3) ⇒ (4) is obvious, just take $y = x$.

(4) ⇒ (1) If $x\in\beta^\perp$, then by 4), $\|x\| = 0$, i.e. $x = 0$. This shows that β is complete. ∎

Suppose $\Gamma := \{u_n\}_{n=1}^\infty$ is a collection of vectors in an inner product space $(H, \langle\cdot|\cdot\rangle)$. The standard **Gram-Schmidt** process produces from Γ an orthonormal subset, $\beta = \{v_n\}_{n=1}^\infty$, such that every element $u_n\in\Gamma$ is a finite linear combination of elements from β. Recall the procedure is to define v_n inductively by setting

$$\tilde v_{n+1} := v_{n+1} - \sum_{j=1}^n\langle u_{n+1}|v_j\rangle v_j = v_{n+1} - P_n v_{n+1}$$

where P_n is orthogonal projection onto $M_n := \operatorname{span}(\{v_k\}_{k=1}^n)$. If $v_{n+1} := 0$, let $\tilde v_{n+1} = 0$, otherwise set $v_{n+1} := \|\tilde v_{n+1}\|^{-1}\tilde v_{n+1}$. Finally re-index the resulting sequence so as to throw out those v_n with $v_n = 0$. The result is an orthonormal subset, $\beta\subset H$, with the desired properties.

Definition 14.25. *As subset, Γ, of a normed space X is said to be* ***total*** *if $\operatorname{span}(\Gamma)$ is a dense in X.*

Remark 14.26. Suppose that $\{u_n\}_{n=1}^\infty$ is a **total** subset of H. Let $\{v_n\}_{n=1}^\infty$ be the vectors found by performing Gram-Schmidt on the set $\{u_n\}_{n=1}^\infty$. Then $\beta := \{v_n\}_{n=1}^\infty$ is an orthonormal basis for H. Indeed, if $h\in H$ is orthogonal to β then h is orthogonal to $\{u_n\}_{n=1}^\infty$ and hence also $\overline{\operatorname{span}\{u_n\}_{n=1}^\infty} = H$. In particular h is orthogonal to itself and so $h = 0$.

Proposition 14.27. *A Hilbert space H is separable iff H has a countable orthonormal basis $\beta\subset H$. Moreover, if H is separable, all orthonormal bases of H are countable. (See Proposition 4.14 in Conway's, "A Course in Functional Analysis," for a more general version of this proposition.)*

Proof. Let $\mathbb{D}\subset H$ be a countable dense set $\mathbb{D} = \{u_n\}_{n=1}^\infty$. By Gram-Schmidt process there exists $\beta = \{v_n\}_{n=1}^\infty$ an orthonormal set such that $\operatorname{span}\{v_n : n = 1, 2\dots, N\}\supseteq\operatorname{span}\{u_n : n = 1, 2\dots, N\}$. So if $\langle x|v_n\rangle = 0$ for all n then $\langle x|u_n\rangle = 0$ for all n. Since $\mathbb{D}\subset H$ is dense we may choose $\{w_k\}\subset\mathbb{D}$ such that $x = \lim_{k\to\infty} w_k$ and therefore $\langle x|x\rangle = \lim_{k\to\infty}\langle x|w_k\rangle = 0$. That is to say $x=0$ and β is complete. Conversely if $\beta\subset H$ is a countable orthonormal basis, then the countable set

$$\mathbb{D}=\left\{\sum_{u\in\beta}a_u u : a_u\in\mathbb{Q}+i\mathbb{Q} : \#\{u : a_u\neq 0\}<\infty\right\}$$

is dense in H. Finally let $\beta=\{u_n\}_{n=1}^\infty$ be an orthonormal basis and $\beta_1\subset H$ be another orthonormal basis. Then the sets

$$B_n=\{v\in\beta_1 : \langle v|u_n\rangle\neq 0\}$$

are countable for each $n\in\mathbb{N}$ and hence $B:=\bigcup_{n=1}^\infty B_n$ is a countable subset of β_1. Suppose there exists $v\in\beta_1\setminus B$, then $\langle v|u_n\rangle=0$ for all n and since $\beta=\{u_n\}_{n=1}^\infty$ is an orthonormal basis, this implies $v=0$ which is impossible since $\|v\|=1$. Therefore $\beta_1\setminus B=\emptyset$ and hence $\beta_1=B$ is countable. ■

Proposition 14.28. *Suppose X and Y are sets and $\mu : X\to(0,\infty)$ and $\nu : Y\to(0,\infty)$ are give weight functions. For functions $f : X\to\mathbb{C}$ and $g : Y\to\mathbb{C}$ let $f\otimes g : X\times Y\to\mathbb{C}$ be defined by $f\otimes g(x,y):=f(x)\,g(y)$. If $\beta\subset\ell^2(\mu)$ and $\gamma\subset\ell^2(\nu)$ are orthonormal bases, then*

$$\beta\otimes\gamma:=\{f\otimes g : f\in\beta \text{ and } g\in\gamma\}$$

is an orthonormal basis for $\ell^2(\mu\otimes\nu)$.

Proof. Let $f,f'\in\ell^2(\mu)$ and $g,g'\in\ell^2(\nu)$, then by the Tonelli's Theorem 4.22 for sums and Hölder's inequality,

$$\begin{aligned}\sum_{X\times Y}|f\otimes g\cdot f'\otimes g'|\,\mu\otimes\nu&=\sum_X|ff'|\,\mu\cdot\sum_Y|gg'|\,\nu\\&\leq\|f\|_{\ell^2(\mu)}\,\|f'\|_{\ell^2(\mu)}\,\|g\|_{\ell^2(\nu)}\,\|g'\|_{\ell^2(\nu)}=1<\infty.\end{aligned}$$

So by Fubini's Theorem 4.23 for sums,

$$\begin{aligned}\langle f\otimes g|f'\otimes g'\rangle_{\ell^2(\mu\otimes\nu)}&=\sum_X f\bar{f}'\mu\cdot\sum_Y g\bar{g}'\nu\\&=\langle f|f'\rangle_{\ell^2(\mu)}\langle g|g'\rangle_{\ell^2(\nu)}=\delta_{f,f'}\delta_{g,g'}.\end{aligned}$$

Therefore, $\beta\otimes\gamma$ is an orthonormal subset of $\ell^2(\mu\otimes\nu)$. So it only remains to show $\beta\otimes\gamma$ is complete. We will give two proofs of this fact. Let $F\in\ell^2(\mu\otimes\nu)$. In the first proof we will verify item 4. of Theorem 14.24 while in the second we will verify item 1 of Theorem 14.24.

First Proof. By Tonelli's Theorem,

$$\sum_{x\in X}\mu(x)\sum_{y\in Y}\nu(y)\,|F(x,y)|^2=\|F\|^2_{\ell^2(\mu\otimes\nu)}<\infty$$

and since $\mu>0$, it follows that

$$\sum_{y\in Y}|F(x,y)|^2\,\nu(y)<\infty \text{ for all } x\in X,$$

i.e. $F(x,\cdot)\in\ell^2(\nu)$ for all $x\in X$. By the completeness of γ,

$$\sum_Y|F(x,y)|^2\,\nu(y)=\langle F(x,\cdot)\,|F(x,\cdot)\rangle_{\ell^2(\nu)}=\sum_{g\in\gamma}\left|\langle F(x,\cdot)\,|g\rangle_{\ell^2(\nu)}\right|^2$$

and therefore,

$$\begin{aligned}\|F\|^2_{\ell^2(\mu\otimes\nu)}&=\sum_{x\in X}\mu(x)\sum_{y\in Y}\nu(y)\,|F(x,y)|^2\\&=\sum_{x\in X}\sum_{g\in\gamma}\left|\langle F(x,\cdot)\,|g\rangle_{\ell^2(\nu)}\right|^2\mu(x).\end{aligned}\tag{14.13}$$

and in particular, $x\to\langle F(x,\cdot)\,|g\rangle_{\ell^2(\nu)}$ is in $\ell^2(\mu)$. So by the completeness of β and the Fubini and Tonelli theorems, we find

$$\begin{aligned}\sum_X\left|\langle F(x,\cdot)\,|g\rangle_{\ell^2(\nu)}\right|^2\mu(x)&=\sum_{f\in\beta}\left|\sum_X\langle F(x,\cdot)\,|g\rangle_{\ell^2(\nu)}f(x)\,\mu(x)\right|^2\\&=\sum_{f\in\beta}\left|\sum_X\left(\sum_Y F(x,y)\,g(y)\,\nu(y)\right)f(x)\,\mu(x)\right|^2\\&=\sum_{f\in\beta}\left|\sum_{X\times Y}F(x,y)\,\overline{f\otimes g}(x,y)\,\mu\otimes\nu(x,y)\right|^2\\&=\sum_{f\in\beta}\left|\langle F|f\otimes g\rangle_{\ell^2(\mu\otimes\nu)}\right|^2.\end{aligned}$$

Combining this result with Eq. (14.13) shows

$$\|F\|^2_{\ell^2(\mu\otimes\nu)}=\sum_{f\in\beta,\ g\in\gamma}\left|\langle F|f\otimes g\rangle_{\ell^2(\mu\otimes\nu)}\right|^2$$

as desired.

Second Proof. Suppose, for all $f\in\beta$ and $g\in\gamma$ that $\langle F|f\otimes g\rangle=0$, i.e.

$$\begin{aligned}0=\langle F|f\otimes g\rangle_{\ell^2(\mu\otimes\nu)}&=\sum_{x\in X}\mu(x)\sum_{y\in Y}\nu(y)\,F(x,y)\bar{f}(x)g(y)\\&=\sum_{x\in X}\mu(x)\,\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\bar{f}(x).\end{aligned}\tag{14.14}$$

Since

$$\sum_{x\in X} \left|\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\right|^2 \mu(x) \leq \sum_{x\in X}\mu(x)\sum_{y\in Y}|F(x,y)|^2\nu(y) < \infty, \tag{14.15}$$

it follows from Eq. (14.14) and the completeness of β that $\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)} = 0$ for all $x \in X$. By the completeness of γ we conclude that $F(x,y) = 0$ for all $(x,y) \in X \times Y$. ∎

Definition 14.29. *A linear map $U : H \to K$ is an **isometry** if $\|Ux\|_K = \|x\|_H$ for all $x \in H$ and U is **unitary if** U is also surjective.*

Exercise 14.4. Let $U : H \to K$ be a linear map, show the following are equivalent:

1. $U : H \to K$ is an isometry,
2. $\langle Ux|Ux'\rangle_K = \langle x|x'\rangle_H$ for all $x, x' \in H$, (see Eq. (14.31) below)
3. $U^*U = id_H$.

Exercise 14.5. Let $U : H \to K$ be a linear map, show the following are equivalent:

1. $U : H \to K$ is unitary
2. $U^*U = id_H$ and $UU^* = id_K$.
3. U is invertible and $U^{-1} = U^*$.

Exercise 14.6. Let H be a Hilbert space. Use Theorem 14.24 to show there exists a set X and a unitary map $U : H \to \ell^2(X)$. Moreover, if H is separable and $\dim(H) = \infty$, then X can be taken to be $\mathbb{N}$ so that H is unitarily equivalent to $\ell^2 = \ell^2(\mathbb{N})$.

14.2 Some Spectral Theory

For this section let H and K be two Hilbert space over $\mathbb{C}$.

Exercise 14.7. Suppose $A : H \to H$ is a bounded self-adjoint operator. Show:

1. If λ is an eigenvalue of A, i.e. $Ax = \lambda x$ for some $x \in H \setminus \{0\}$, then $\lambda \in \mathbb{R}$.
2. If λ and μ are two distinct eigenvalues of A with eigenvectors x and y respectively, then $x \perp y$.

Unlike in finite dimensions, it is possible that an operator on a complex Hilbert space may have no eigenvalues, see Example 14.35 and Lemma 14.36 below for a couple of examples. For this reason it is useful to generalize the notion of an eigenvalue as follows.

Definition 14.30. *Suppose X is a Banach space over $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) and $A \in L(X)$. We say $\lambda \in \mathbb{F}$ is in the **spectrum** of A if $A - \lambda I$ does **not** have a bounded[3] inverse. The **spectrum** will be denoted by $\sigma(A) \subset \mathbb{F}$. The **resolvent set** for A is $\rho(A) := \mathbb{F}\setminus\sigma(A)$.*

Remark 14.31. If λ is an eigenvalue of A, then $A - \lambda I$ is not injective and hence not invertible. Therefore any eigenvalue of A is in the spectrum of A. If H is a Hilbert space ant $A \in L(H)$, it follows from item 5. of Proposition 14.16 that $\lambda \in \sigma(A)$ iff $\bar\lambda \in \sigma(A^*)$, i.e.

$$\sigma(A^*) = \left\{\bar\lambda : \lambda \in \sigma(A)\right\}.$$

Exercise 14.8. Suppose X is a Banach space and $A \in L(X)$. Use Corollary 7.20 to show $\sigma(A)$ is a closed subset of $\left\{\lambda \in \mathbb{F} : |\lambda| \leq \|A\| := \|A\|_{L(X)}\right\}$.

Lemma 14.32. *Suppose that $A \in L(H)$ is a normal operator, i.e. $[A, A^*] = 0$. Then $\lambda \in \sigma(A)$ iff*

$$\inf_{\|\psi\|=1}\|(A - \lambda 1)\psi\| = 0. \tag{14.16}$$

In other words, $\lambda \in \sigma(A)$ iff there is an "approximate sequence of eigenvectors" for (A, λ), i.e. there exists $\psi_n \in H$ such that $\|\psi_n\| = 1$ and $A\psi_n - \lambda\psi_n \to 0$ as $n \to \infty$.

Proof. By replacing A by $A - \lambda I$ we may assume that $\lambda = 0$. If $0 \notin \sigma(A)$, then

$$\inf_{\|\psi\|=1}\|A\psi\| = \inf\frac{\|A\psi\|}{\|\psi\|} = \inf\frac{\|\psi\|}{\|A^{-1}\psi\|} = 1/\left\|A^{-1}\right\| > 0.$$

Now suppose that $\inf_{\|\psi\|=1}\|A\psi\| = \varepsilon > 0$ or equivalently we have

$$\|A\psi\| \geq \varepsilon\|\psi\|$$

for all $\psi \in H$. Because A is normal,

$$\|A\psi\|^2 = \langle A^*A\psi|\psi\rangle = \langle AA^*\psi|\psi\rangle = \langle A^*\psi|A^*\psi\rangle = \|A^*\psi\|^2.$$

Therefore we also have

$$\|A^*\psi\| = \|A\psi\| \geq \varepsilon\|\psi\| \ \forall\, \psi \in H. \tag{14.17}$$

This shows in particular that A and A^* are injective, $\mathrm{Ran}(A)$ is closed and hence by Lemma 14.17

$$\mathrm{Ran}(A) = \overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^\perp = \{0\}^\perp = H.$$

Therefore A is algebraically invertible and the inverse is bounded by Eq. (14.17). ∎

[3] It will follow by the open mapping Theorem 35.1 or the closed graph Theorem 35.3 that the word bounded may be omitted from this definition.

Lemma 14.33. *Suppose that $A \in L(H)$ is self-adjoint (i.e. $A = A^*$) then*

$$\sigma(A) \subset \left[-\|A\|_{op}, \|A\|_{op}\right] \subset \mathbb{R}.$$

Proof. Writting $\lambda = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned} \|(A+\alpha+i\beta)\psi\|^2 &= \|(A+\alpha)\psi\|^2 + |\beta|^2 \|\psi\|^2 + 2\operatorname{Re}((A+\alpha)\psi, i\beta\psi) \\ &= \|(A+\alpha)\psi\|^2 + |\beta|^2 \|\psi\|^2 \end{aligned} \tag{14.18}$$

wherein we have used

$$\operatorname{Re}[i\beta((A+\alpha)\psi,\psi)] = \beta \operatorname{Im}((A+\alpha)\psi,\psi) = 0$$

since

$$((A+\alpha)\psi,\psi) = (\psi,(A+\alpha)\psi) = \overline{((A+\alpha)\psi,\psi)}.$$

Eq. (14.18) along with Lemma 14.32 shows that $\lambda \notin \sigma(A)$ if $\beta \neq 0$, i.e. $\sigma(A) \subset \mathbb{R}$. The fact that $\sigma(A)$ is now contained in $\left[-\|A\|_{op}, \|A\|_{op}\right]$ is a consequence of Exercise 14.8. ■

Remark 14.34. It is not true that $\sigma(A) \subset \mathbb{R}$ implies $A = A^*$. For example let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on $H = \mathbb{C}^2$, then $\sigma(A) = \{0\}$ yet $A \neq A^*$.

Example 14.35. Let $S \in L(H)$ be a (not necessarily) normal operator. The proof of Lemma 14.32 gives $\lambda \in \sigma(S)$ if Eq. (14.16) holds. However the converse is not always valid unless S is normal. For example, let $S : \ell^2 \to \ell^2$ be the shift, $S(\omega_1, \omega_2, \dots) = (0, \omega_1, \omega_2, \dots)$. Then for any $\lambda \in D := \{z \in \mathbb{C} : |z| < 1\}$,

$$\|(S-\lambda)\psi\| = \|S\psi - \lambda\psi\| \geq |\|S\psi\| - |\lambda|\,\|\psi\|| = (1-|\lambda|)\|\psi\|$$

and so there does not exists an approximate sequence of eigenvectors for (S,λ). However, as we will now show, $\sigma(S) = \bar{D}$.

To prove this it suffices to show by Remark 14.31 and Exercise 14.8 that $D \subset \sigma(S^*)$. For if this is the cae then $\bar{D} \subset \sigma(S^*) \subset \bar{D}$ and hence $\sigma(S) = \bar{D}$ since $\bar{D}$ is invariant under complex conjugation.

A simple computation shows,

$$S^*(\omega_1, \omega_2, \dots) = (\omega_2, \omega_3, \dots)$$

and $\omega = (\omega_1, \omega_2, \dots)$ is an eigenvector for S^* with eigenvalue $\lambda \in \mathbb{C}$ iff

$$0 = (S^* - \lambda I)(\omega_1, \omega_2, \dots) = (\omega_2 - \lambda\omega_1, \omega_3 - \lambda\omega_2, \dots).$$

Solving these equation shows

$$\omega_2 = \lambda\omega_1,\ \omega_3 = \lambda\omega_2 = \lambda^2\omega_1\ , \dots,\ \omega_n = \lambda^{n-1}\omega_1.$$

Hence if $\lambda \in D$, we may let $\omega_1 = 1$ above to find

$$S^*(1, \lambda, \lambda^2, \dots) = \lambda(1, \lambda, \lambda^2, \dots)$$

where $(1, \lambda, \lambda^2, \dots) \in \ell^2$. Thus we have shown λ is an eigenvalue for S^* for all $\lambda \in D$ and hence $D \subset \sigma(S^*)$.

Lemma 14.36. *Let $H = \ell^2(\mathbb{Z})$ and let $A : H \to H$ be defined by*

$$Af(k) = i(f(k+1) - f(k-1)) \text{ for all } k \in \mathbb{Z}.$$

Then:

1. *A is a bounded self-adjoint operator.*
2. *A has no eigenvalues.*
3. *$\sigma(A) = [-2, 2]$.*

Proof. For another (simpler) proof of this lemma, see Exercise 23.8 below.
1. Since

$$\|Af\|_2 \leq \|f(\cdot+1)\|_2 + \|f(\cdot-1)\|_2 = 2\|f\|_2,$$

$\|A\|_{op} \leq 2 < \infty$. Moreover, for $f, g \in \ell^2(\mathbb{Z})$,

$$\begin{aligned} \langle Af|g\rangle &= \sum_k i(f(k+1) - f(k-1))\bar{g}(k) \\ &= \sum_k if(k)\bar{g}(k-1) - \sum_k if(k)\bar{g}(k+1) \\ &= \sum_k f(k)\overline{Ag(k)} = \langle f|Ag\rangle, \end{aligned}$$

which shows $A = A^*$.

2. From Lemma 14.33, we know that $\sigma(A) \subset [-2, 2]$. If $\lambda \in [-2, 2]$ and $f \in H$ satisfies $Af = \lambda f$, then

$$f(k+1) = -i\lambda f(k) + f(k-1) \text{ for all } k \in \mathbb{Z}. \tag{14.19}$$

This is a second order difference equations which can be solved analogously to second order ordinary differential equations. The idea is to start by looking for a solution of the form $f(k) = \alpha^k$. Then Eq. (14.19) beocmes, $\alpha^{k+1} = -i\lambda\alpha^k + \alpha^{k-1}$ or equivalently that

$$\alpha^2 + i\lambda\alpha - 1 = 0.$$

So we will have a solution if $\alpha \in \{\alpha_\pm\}$ where

$$\alpha_\pm = \frac{-i\lambda \pm \sqrt{4 - \lambda^2}}{2}.$$

For $|\lambda| \neq 2$, there are two distinct roots and the general solution to Eq. (14.19) is of the form

$$f(k) = c_+\alpha_+^k + c_-\alpha_-^k \tag{14.20}$$

for some constants $c_\pm \in \mathbb{C}$ and $|\lambda| = 2$, the general solution has the form

$$f(k) = c\alpha_+^k + dk\alpha_+^k \tag{14.21}$$

Since in all cases, $|\alpha_\pm| = \frac{1}{4}\left(\lambda^2 + 4 - \lambda^2\right) = 1$, it follows that neither of these functions, f, will be in $\ell^2(\mathbb{Z})$ unless they are identically zero. This shows that A has no eigenvalues.

3. The above argument suggest a method for constructing approximate eigenfucntions. Namely, let $\lambda \in [-2,2]$ and define $f_n(k) := 1_{|k|\leq n}\alpha^k$ where $\alpha = \alpha_+$. Then a simple computation shows

$$\lim_{n\to\infty} \frac{\|(A-\lambda I) f_n\|_2}{\|f_n\|_2} = 0 \tag{14.22}$$

and therefore $\lambda \in \sigma(A)$. ■

Exercise 14.9. Verify Eq. (14.22). Also show by explicit computations that

$$\lim_{n\to\infty} \frac{\|(A-\lambda I) f_n\|_2}{\|f_n\|_2} \neq 0$$

if $\lambda \notin [-2,2]$.

The next couple of results will be needed for the next section.

Theorem 14.37 (Rayleigh quotient). *Suppose $T \in L(H) := L(H,H)$ is a bounded self-adjoint operator, then*

$$\|T\| = \sup_{f\neq 0} \frac{|\langle f|Tf\rangle|}{\|f\|^2}.$$

Moreover if there exists a non-zero element $g \in H$ such that

$$\frac{|\langle Tg|g\rangle|}{\|g\|^2} = \|T\|,$$

then g is an eigenvector of T with $Tg = \lambda g$ and $\lambda \in \{\pm\|T\|\}$.

Proof. Let

$$M := \sup_{f\neq 0} \frac{|\langle f|Tf\rangle|}{\|f\|^2}.$$

We wish to show $M = \|T\|$. Since

$$|\langle f|Tf\rangle| \leq \|f\|\|Tf\| \leq \|T\|\|f\|^2,$$

we see $M \leq \|T\|$. Conversely let $f,g \in H$ and compute

$$\begin{aligned}\langle f+g|T(f+g)\rangle &- \langle f-g|T(f-g)\rangle \\ &= \langle f|Tg\rangle + \langle g|Tf\rangle + \langle f|Tg\rangle + \langle g|Tf\rangle \\ &= 2[\langle f|Tg\rangle + \langle Tg|f\rangle] = 2[\langle f|Tg\rangle + \overline{\langle f|Tg\rangle}] \\ &= 4\mathrm{Re}\langle f|Tg\rangle.\end{aligned}$$

Therefore, if $\|f\| = \|g\| = 1$, it follows that

$$|\mathrm{Re}\langle f|Tg\rangle| \leq \frac{M}{4}\left\{\|f+g\|^2 + \|f-g\|^2\right\} = \frac{M}{4}\left\{2\|f\|^2 + 2\|g\|^2\right\} = M.$$

By replacing f be $e^{i\theta}f$ where θ is chosen so that $e^{i\theta}\langle f|Tg\rangle$ is real, we find

$$|\langle f|Tg\rangle| \leq M \text{ for all } \|f\| = \|g\| = 1.$$

Hence

$$\|T\| = \sup_{\|f\|=\|g\|=1} |\langle f|Tg\rangle| \leq M.$$

If $g \in H \setminus \{0\}$ and $\|T\| = |\langle Tg|g\rangle|/\|g\|^2$ then, using the Cauchy Schwarz inequality,

$$\|T\| = \frac{|\langle Tg|g\rangle|}{\|g\|^2} \leq \frac{\|Tg\|}{\|g\|} \leq \|T\|. \tag{14.23}$$

This implies $|\langle Tg|g\rangle| = \|Tg\|\|g\|$ and forces equality in the Cauchy Schwarz inequality. So by Theorem 14.2, Tg and g are linearly dependent, i.e. $Tg = \lambda g$ for some $\lambda \in \mathbb{C}$. Substituting this into (14.23) shows that $|\lambda| = \|T\|$. Since T is self-adjoint,

$$\lambda\|g\|^2 = \langle \lambda g|g\rangle = \langle Tg|g\rangle = \langle g|Tg\rangle = \langle g|\lambda g\rangle = \bar{\lambda}\langle g|g\rangle,$$

which implies that $\lambda \in \mathbb{R}$ and therefore, $\lambda \in \{\pm\|T\|\}$. ■

Lemma 14.38 (Invariant subspaces). *Let $T : H \to H$ be a self-adjoint operator and M be a T – invariant subspace of H, i.e. $T(M) \subset M$. Then $M^\perp$ is also a T – invariant subspace, i.e. $T(M^\perp) \subset M^\perp$.*

Proof. Let $x \in M$ and $y \in M^\perp$, then $Tx \in M$ and hence

$$0 = \langle Tx|y\rangle = \langle x|Ty\rangle \text{ for all } x \in M.$$

Thus $Ty \in M^\perp$. ■

14.3 Compact Operators on a Hilbert Space

In this section let H and B be Hilbert spaces and $U := \{x \in H : \|x\| < 1\}$ be the **unit ball** in H. Recall from Definition 11.16 that a bounded operator, $K : H \to B$, is compact iff $\overline{K(U)}$ is compact in B. Equivalently, for all bounded sequences $\{x_n\}_{n=1}^\infty \subset H$, the sequence $\{Kx_n\}_{n=1}^\infty$ has a convergent subsequence in B. Because of Theorem 11.15, if $\dim(H) = \infty$ and $K : H \to B$ is invertible, then K is **not** compact.

Definition 14.39. *$K : H \to B$ is said to have* ***finite rank*** *if* $\operatorname{Ran}(K) \subset B$ *is finite dimensional.*

The following result is a simple consequence of Corollaries 11.13 and 11.14.

Corollary 14.40. *If $K : H \to B$ is a finite rank operator, then K is compact. In particular if either* $\dim(H) < \infty$ *or* $\dim(B) < \infty$ *then any bounded operator $K : H \to B$ is finite rank and hence compact.*

Lemma 14.41. *Let $\mathcal{K} := \mathcal{K}(H,B)$ denote the compact operators from H to B. Then $\mathcal{K}(H,B)$ is a norm closed subspace of $L(H,B)$.*

Proof. The fact that $\mathcal{K}$ is a vector subspace of $L(H,B)$ will be left to the reader. To finish the proof, we must show that $K \in L(H,B)$ is compact if there exists $K_n \in \mathcal{K}(H,B)$ such that $\lim_{n\to\infty} \|K_n - K\|_{op} = 0$.

First Proof. Given $\varepsilon > 0$, choose $N = N(\varepsilon)$ such that $\|K_N - K\| < \varepsilon$. Using the fact that $K_N U$ is precompact, choose a finite subset $\Lambda \subset U$ such that $\min_{x\in\Lambda} \|y - K_N x\| < \varepsilon$ for all $y \in K_N(U)$. Then for $z = Kx_0 \in K(U)$ and $x \in \Lambda$,

$$\begin{aligned}\|z - Kx\| &= \|(K - K_N)x_0 + K_N(x_0 - x) + (K_N - K)x\| \\ &\le 2\varepsilon + \|K_N x_0 - K_N x\|.\end{aligned}$$

Therefore $\min_{x\in\Lambda} \|z - K_N x\| < 3\varepsilon$, which shows $K(U)$ is 3ε bounded for all $\varepsilon > 0$, $K(U)$ is totally bounded and hence precompact.

Second Proof. Suppose $\{x_n\}_{n=1}^\infty$ is a bounded sequence in H. By compactness, there is a subsequence $\{x_n^1\}_{n=1}^\infty$ of $\{x_n\}_{n=1}^\infty$ such that $\{K_1 x_n^1\}_{n=1}^\infty$ is convergent in B. Working inductively, we may construct subsequences

$$\{x_n\}_{n=1}^\infty \supset \{x_n^1\}_{n=1}^\infty \supset \{x_n^2\}_{n=1}^\infty \cdots \supset \{x_n^m\}_{n=1}^\infty \supset \ldots$$

such that $\{K_m x_n^m\}_{n=1}^\infty$ is convergent in B for each m. By the usual Cantor's diagonalization procedure, let $y_n := x_n^n$, then $\{y_n\}_{n=1}^\infty$ is a subsequence of $\{x_n\}_{n=1}^\infty$ such that $\{K_m y_n\}_{n=1}^\infty$ is convergent for all m. Since

$$\begin{aligned}\|Ky_n - Ky_l\| &\le \|(K - K_m)\, y_n\| + \|K_m(y_n - y_l)\| + \|(K_m - K)\, y_l\| \\ &\le 2\,\|K - K_m\| + \|K_m(y_n - y_l)\|\,,\end{aligned}$$

$$\lim \sup_{n,l\to\infty} \|Ky_n - Ky_l\| \le 2\,\|K - K_m\| \to 0 \text{ as } m \to \infty,$$

which shows $\{Ky_n\}_{n=1}^\infty$ is Cauchy and hence convergent. ∎

Proposition 14.42. *A bounded operator $K : H \to B$ is compact iff there exists finite rank operators, $K_n : H \to B$, such that $\|K - K_n\| \to 0$ as $n \to \infty$.*

Proof. Since $\overline{K(U)}$ is compact it contains a countable dense subset and from this it follows that $\overline{K(H)}$ is a separable subspace of B. Let $\{\phi_n\}$ be an orthonormal basis for $\overline{K(H)} \subset B$ and

$$P_N y = \sum_{n=1}^N \langle y|\phi_n\rangle \phi_n$$

be the orthogonal projection of y onto $\operatorname{span}\{\phi_n\}_{n=1}^N$. Then $\lim_{N\to\infty} \|P_N y - y\| = 0$ for all $y \in K(H)$. Define $K_n := P_n K$ – a finite rank operator on H. For sake of contradiction suppose that

$$\lim\sup_{n\to\infty} \|K - K_n\| = \varepsilon > 0,$$

in which case there exists $x_{n_k} \in U$ such that $\|(K - K_{n_k})x_{n_k}\| \ge \varepsilon$ for all n_k. Since K is compact, by passing to a subsequence if necessary, we may assume $\{Kx_{n_k}\}_{n_k=1}^\infty$ is convergent in B. Letting $y := \lim_{k\to\infty} Kx_{n_k}$,

$$\begin{aligned}\|(K - K_{n_k})x_{n_k}\| &= \|(1 - P_{n_k})Kx_{n_k}\| \\ &\le \|(1 - P_{n_k})(Kx_{n_k} - y)\| + \|(1 - P_{n_k})y\| \\ &\le \|Kx_{n_k} - y\| + \|(1 - P_{n_k})y\| \to 0 \text{ as } k \to \infty.\end{aligned}$$

But this contradicts the assumption that ε is positive and hence we must have $\lim_{n\to\infty} \|K - K_n\| = 0$, i.e. K is an operator norm limit of finite rank operators. The converse direction follows from Corollary 14.40 and Lemma 14.41. ∎

Corollary 14.43. *If K is compact then so is K^*.*

Proof. First Proof. Let $K_n = P_n K$ be as in the proof of Proposition 14.42, then $K_n^* = K^* P_n$ is still finite rank. Furthermore, using Proposition 14.16,

$$\|K^* - K_n^*\| = \|K - K_n\| \to 0 \text{ as } n \to \infty$$

showing K^* is a limit of finite rank operators and hence compact.

Second Proof. Let $\{x_n\}_{n=1}^\infty$ be a bounded sequence in B, then

$$\|K^* x_n - K^* x_m\|^2 = (x_n - x_m, KK^*(x_n - x_m)) \le 2C\,\|KK^*(x_n - x_m)\| \tag{14.24}$$

where C is a bound on the norms of the x_n. Since $\{K^* x_n\}_{n=1}^\infty$ is also a bounded sequence, by the compactness of K there is a subsequence $\{x_n'\}$ of the $\{x_n\}$ such that $KK^* x_n'$ is convergent and hence by Eq. (14.24), so is the sequence $\{K^* x_n'\}$. ∎

14.3.1 The Spectral Theorem for Self Adjoint Compact Operators

For the rest of this section, $K \in \mathcal{K}(H) := \mathcal{K}(H,H)$ will be a self-adjoint compact operator or **S.A.C.O.** for short. Because of Proposition 14.42, we might expect compact operators to behave very much like finite dimensional matrices. This is typically the case as we will see below.

Example 14.44 (Model S.A.C.O.). Let $H = \ell_2$ and K be the diagonal matrix

$$K = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix},$$

where $\lim_{n\to\infty} |\lambda_n| = 0$ and $\lambda_n \in \mathbb{R}$. Then K is a self-adjoint compact operator. This assertion was proved in Example 11.17 above.

The main theorem (Theorem 14.46) of this subsection states that up to unitary equivalence, Example 14.44 is essentially the most general example of an S.A.C.O.

Theorem 14.45. *Let K be a S.A.C.O., then either $\lambda = \|K\|$ or $\lambda = -\|K\|$ is an eigenvalue of K.*

Proof. Without loss of generality we may assume that K is non-zero since otherwise the result is trivial. By Theorem 14.37, there exists $f_n \in H$ such that $\|f_n\| = 1$ and

$$\frac{|\langle f_n | K f_n \rangle|}{\|f_n\|^2} = |\langle f_n | K f_n \rangle| \longrightarrow \|K\| \text{ as } n \to \infty. \tag{14.25}$$

By passing to a subsequence if necessary, we may assume that $\lambda := \lim_{n\to\infty} \langle f_n | K f_n \rangle$ exists and $\lambda \in \{\pm \|K\|\}$. By passing to a further subsequence if necessary, we may assume, using the compactness of K, that Kf_n is convergent as well. We now compute:

$$\begin{aligned} 0 \le \|Kf_n - \lambda f_n\|^2 &= \|Kf_n\|^2 - 2\lambda \langle K f_n | f_n \rangle + \lambda^2 \\ &\le \lambda^2 - 2\lambda \langle K f_n | f_n \rangle + \lambda^2 \\ &\to \lambda^2 - 2\lambda^2 + \lambda^2 = 0 \text{ as } n \to \infty. \end{aligned}$$

Hence

$$K f_n - \lambda f_n \to 0 \text{ as } n \to \infty \tag{14.26}$$

and therefore

$$f := \lim_{n\to\infty} f_n = \frac{1}{\lambda} \lim_{n\to\infty} K f_n$$

exists. By the continuity of the inner product, $\|f\| = 1 \neq 0$. By passing to the limit in Eq. (14.26) we find that $Kf = \lambda f$. ■

Theorem 14.46 (Compact Operator Spectral Theorem). *Suppose that $K : H \to H$ is a non-zero S.A.C.O., then*

1. *there exists at least one eigenvalue $\lambda \in \{\pm \|K\|\}$.*
2. *There are at most countable many* ***non-zero*** *eigenvalues, $\{\lambda_n\}_{n=1}^N$, where $N = \infty$ is allowed. (Unless K is finite rank (i.e. $\dim \operatorname{Ran}(K) < \infty$), N will be infinite.)*
3. *The λ_n's (including multiplicities) may be arranged so that $|\lambda_n| \ge |\lambda_{n+1}|$ for all n. If $N = \infty$ then $\lim_{n\to\infty} |\lambda_n| = 0$. (In particular any eigenspace for K with* ***non-zero*** *eigenvalue is finite dimensional.)*
4. *The eigenvectors $\{\phi_n\}_{n=1}^N$ can be chosen to be an O.N. set such that $H = \overline{\operatorname{span}\{\phi_n\}} \oplus \operatorname{Nul}(K)$.*
5. *Using the $\{\phi_n\}_{n=1}^N$ above,*
$$K\psi = \sum_{n=1}^N \lambda_n \langle \psi | \phi_n \rangle \phi_n \text{ for all } \psi \in H.$$
6. *The spectrum of K is, $\sigma(K) = \{0\} \cup \{\lambda_n : n < N+1\}$.*

Proof. We will find λ_n's and ϕ_n's recursively. Let $\lambda_1 \in \{\pm\|K\|\}$ and $\phi_1 \in H$ such that $K\phi_1 = \lambda_1 \phi_1$ as in Theorem 14.45. Take $M_1 = \operatorname{span}(\phi_1)$ so $K(M_1) \subset M_1$. By Lemma 14.38, $KM_1^\perp \subset M_1^\perp$. Define $K_1 : M_1^\perp \to M_1^\perp$ via $K_1 = K|_{M_1^\perp}$. Then K_1 is again a compact operator. If $K_1 = 0$, we are done. If $K_1 \neq 0$, by Theorem 14.45 there exists $\lambda_2 \in \{\pm \|K\|_1\}$ and $\phi_2 \in M_1^\perp$ such that $\|\phi_2\| = 1$ and $K_1\phi_2 = K\phi_2 = \lambda_2 \phi_2$. Let $M_2 := \operatorname{span}(\phi_1, \phi_2)$. Again $K(M_2) \subset M_2$ and hence $K_2 := K|_{M_2^\perp} : M_2^\perp \to M_2^\perp$ is compact. Again if $K_2 = 0$ we are done. If $K_2 \neq 0$. Then by Theorem 14.45 there exists $\lambda_3 \in \{\pm\|K\|_2\}$ and $\phi_3 \in M_2^\perp$ such that $\|\phi_3\| = 1$ and $K_2\phi_3 = K\phi_3 = \lambda_3\phi_3$. Continuing this way indefinitely or until we reach a point where $K_n = 0$, we construct a sequence $\{\lambda_n\}_{n=1}^N$ of eigenvalues and orthonormal eigenvectors $\{\phi_n\}_{n=1}^N$ such that $|\lambda_i| \ge |\lambda_{i+1}|$ with the further property that

$$|\lambda_i| = \sup_{\phi \perp \{\phi_1, \phi_2, \ldots \phi_{i-1}\}} \frac{\|K\phi\|}{\|\phi\|} \tag{14.27}$$

If $N = \infty$ then $\lim_{i\to\infty} |\lambda_i| = 0$ for if not there would exist $\varepsilon > 0$ such that $|\lambda_i| \ge \varepsilon > 0$ for all i. In this case $\{\phi_i/\lambda_i\}_{i=1}^\infty$ is sequence in H bounded by ε^{-1}. By compactness of K, there exists a subsequence i_k such that $\phi_{i_k} = K\phi_{i_k}/\lambda_{i_k}$ is convergent. But this is impossible since $\{\phi_{i_k}\}$ is an orthonormal set. Hence we must have that $\varepsilon = 0$. Let $M := \operatorname{span}\{\phi_i\}_{i=1}^N$ with $N = \infty$ **possible**. Then $K(M) \subset M$ and hence $K(M^\perp) \subset M^\perp$. Using Eq. (14.27),

$$\|K|_{M^\perp}\| \le \|K|_{M_n^\perp}\| = |\lambda_n| \longrightarrow 0 \text{ as } n \to \infty$$

showing $K|M^\perp \equiv 0$. Define P_0 to be orthogonal projection onto $M^\perp$. Then for $\psi \in H$,

$$\psi = P_0\psi + (1 - P_0)\psi = P_0\psi + \sum_{i=1}^{N} \langle\psi|\phi_i\rangle\phi_i$$

and

$$K\psi = KP_0\psi + K\sum_{i=1}^{N}\langle\psi|\phi_i\rangle\phi_i = \sum_{i=1}^{N}\lambda_i\langle\psi|\phi_i\rangle\phi_i.$$

Since $\{\lambda_n\} \subset \sigma(K)$ and $\sigma(K)$ is closed, it follows that $0 \in \sigma(K)$ and hence $\{\lambda_n\}_{n=1}^{\infty} \cup \{0\} \subset \sigma(K)$. Suppose that $z \notin \{\lambda_n\}_{n=1}^{\infty} \cup \{0\}$ and let d be the distance between z and $\{\lambda_n\}_{n=1}^{\infty}\cup\{0\}$. Notice that $d > 0$ because $\lim_{n\to\infty} \lambda_n = 0$. A few simple computations show that:

$$(K - zI)\psi = \sum_{i=1}^{N}\langle\psi|\phi_i\rangle(\lambda_i - z)\phi_i - zP_0\psi,$$

$(K - z)^{-1}$ exists,

$$(K - zI)^{-1}\psi = \sum_{i=1}^{N}\langle\psi|\phi_i\rangle(\lambda_i - z)^{-1}\phi_i - z^{-1}P_0\psi,$$

and

$$\begin{aligned}\|(K - zI)^{-1}\psi\|^2 &= \sum_{i=1}^{N}|\langle\psi|\phi_i\rangle|^2\frac{1}{|\lambda_i - z|^2} + \frac{1}{|z|^2}\|P_0\psi\|^2 \\ &\le \left(\frac{1}{d}\right)^2\left(\sum_{i=1}^{N}|\langle\psi|\phi_i\rangle|^2 + \|P_0\psi\|^2\right) = \frac{1}{d^2}\|\psi\|^2.\end{aligned}$$

We have thus shown that $(K - zI)^{-1}$ exists, $\|(K - zI)^{-1}\| \le d^{-1} < \infty$ and hence $z \notin \sigma(K)$. ■

Theorem 14.47 (Structure of Compact Operators). *Let $K : H \to B$ be a compact operator. Then there exists $N \in \mathbb{N}\cup\{\infty\}$, orthonormal subsets $\{\phi_n\}_{n=1}^{N} \subset H$ and $\{\psi_n\}_{n=1}^{N} \subset B$ and a sequences $\{\alpha_n\}_{n=1}^{N} \subset \mathbb{R}_+$ such that $\lambda_1 \ge \lambda_2 \ge \dots$, $\lim_{n\to\infty} \alpha_n = 0$ if $N = \infty$, $\|\psi_n\| \le 1$ for all n and*

$$Kf = \sum_{n=1}^{N}\alpha_n\langle f|\phi_n\rangle\psi_n \text{ for all } f \in H. \tag{14.28}$$

Proof. Since K^*K is a selfadjoint compact operator, Theorem 14.46 implies there exists an orthonormal set $\{\phi_n\}_{n=1}^{N} \subset H$ and positive numbers $\{\lambda_n\}_{n=1}^{N}$ such that

$$K^*K\psi = \sum_{n=1}^{N}\lambda_n\langle\psi|\phi_n\rangle\phi_n \text{ for all } \psi \in H.$$

Let A be the positive square root of K^*K defined by

$$A\psi := \sum_{n=1}^{N}\sqrt{\lambda_n}\langle\psi|\phi_n\rangle\phi_n \text{ for all } \psi \in H.$$

A simple computation shows, $A^2 = K^*K$, and therefore,

$$\begin{aligned}\|A\psi\|^2 &= \langle A\psi|A\psi\rangle = \langle\psi|A^2\psi\rangle \\ &= \langle\psi|K^*K\psi\rangle = \langle K\psi|K\psi\rangle = \|K\psi\|^2\end{aligned}$$

for all $\psi \in H$. Hence we may define a unitary operator, $u : \overline{\mathrm{Ran}(A)} \to \overline{\mathrm{Ran}(K)}$ by the formula

$$uA\psi = K\psi \text{ for all } \psi \in H.$$

We then have

$$K\psi = uA\psi = \sum_{n=1}^{N}\sqrt{\lambda_n}\langle\psi|\phi_n\rangle u\phi_n \tag{14.29}$$

which proves the result with $\psi_n := u\phi_n$ and $\alpha_n = \sqrt{\lambda_n}$.

It is instructive to find ψ_n explicitly and to verify Eq. (14.29) by bruit force. Since $\phi_n = \lambda_n^{-1/2}A\phi_n$,

$$\psi_n = \lambda_n^{-1/2}uA\phi_n = \lambda_n^{-1/2}uA\phi_n = \lambda_n^{-1/2}K\phi_n$$

and

$$\langle K\phi_n|K\phi_m\rangle = \langle\phi_n|K^*K\phi_m\rangle = \lambda_n\delta_{mn}.$$

This verifies that $\{\psi_n\}_{n=1}^{N}$ is an orthonormal set. Moreover,

$$\begin{aligned}\sum_{n=1}^{N}\sqrt{\lambda_n}\langle\psi|\phi_n\rangle\psi_n &= \sum_{n=1}^{N}\sqrt{\lambda_n}\langle\psi|\phi_n\rangle\lambda_n^{-1/2}K\phi_n \\ &= K\sum_{n=1}^{N}\langle\psi|\phi_n\rangle\phi_n = K\psi\end{aligned}$$

since $\sum_{n=1}^{N}\langle\psi|\phi_n\rangle\phi_n = P\psi$ where P is orthogonal projection onto $\mathrm{Nul}(K)^{\perp}$.

Second Proof. Let $K = u\,|K|$ be the polar decomposition of K. Then $|K|$ is self-adjoint and compact, by Corollary 37.12, and hence by Theorem 14.46 there exists an orthonormal basis $\{\phi_n\}_{n=1}^{N}$ for $\mathrm{Nul}(|K|)^{\perp} = \mathrm{Nul}(K)^{\perp}$ such that $|K|\,\phi_n = \lambda_n\phi_n$, $\lambda_1 \ge \lambda_2 \ge \dots$ and $\lim_{n\to\infty}\lambda_n = 0$ if $N = \infty$. For $f \in H$,

$$Kf = u\,|K|\sum_{n=1}^{N}\langle f|\phi_n\rangle\phi_n = \sum_{n=1}^{N}\langle f|\phi_n\rangle u\,|K|\,\phi_n = \sum_{n=1}^{N}\lambda_n\langle f|\phi_n\rangle u\phi_n$$

which is Eq. (14.28) with $\psi_n := u\phi_n$. ■

14.4 Weak Convergence

Suppose H is an infinite dimensional Hilbert space and $\{x_n\}_{n=1}^\infty$ is an orthonormal subset of H. Then, by Eq. (14.1), $\|x_n - x_m\|^2 = 2$ for all $m \neq n$ and in particular, $\{x_n\}_{n=1}^\infty$ has no convergent subsequences. From this we conclude that $C := \{x \in H : \|x\| \leq 1\}$, the closed unit ball in H, is not compact. To overcome this problems it is sometimes useful to introduce a weaker topology on X having the property that C is compact.

Definition 14.48. *Let $(X, \|\cdot\|)$ be a Banach space and X^* be its continuous dual. The weak topology, τ_w, on X is the topology generated by X^*. If $\{x_n\}_{n=1}^\infty \subset X$ is a sequence we will write $x_n \overset{w}{\to} x$ as $n \to \infty$ to mean that $x_n \to x$ in the weak topology.*

Because $\tau_w = \tau(X^*) \subset \tau_{\|\cdot\|} := \tau(\{\|x - \cdot\| : x \in X\})$, it is harder for a function $f : X \to \mathbb{F}$ to be continuous in the τ_w – topology than in the norm topology, $\tau_{\|\cdot\|}$. In particular if $\phi : X \to \mathbb{F}$ is a linear functional which is τ_w – continuous, then ϕ is $\tau_{\|\cdot\|}$ – continuous and hence $\phi \in X^*$.

Exercise 14.10. Show the vector space operations of X are continuous in the weak topology, i.e. show:

1. $(x, y) \in X \times X \to x + y \in X$ is $(\tau_w \otimes \tau_w, \tau_w)$ – continuous and
2. $(\lambda, x) \in \mathbb{F} \times X \to \lambda x \in X$ is $(\tau_{\mathbb{F}} \otimes \tau_w, \tau_w)$ – continuous.

Proposition 14.49. *Let $\{x_n\}_{n=1}^\infty \subset X$ be a sequence, then $x_n \overset{w}{\to} x \in X$ as $n \to \infty$ iff $\phi(x) = \lim_{n\to\infty} \phi(x_n)$ for all $\phi \in X^*$.*

Proof. By definition of τ_w, we have $x_n \overset{w}{\to} x \in X$ iff for all $\Gamma \subset\subset X^*$ and $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|\phi(x) - \phi(x_n)| < \varepsilon$ for all $n \geq N$ and $\phi \in \Gamma$. This later condition is easily seen to be equivalent to $\phi(x) = \lim_{n\to\infty} \phi(x_n)$ for all $\phi \in X^*$. ■

The topological space (X, τ_w) is still Hausdorff as follows from the Hahn Banach Theorem, see Theorem 7.26. For the moment we will concentrate on the special case where $X = H$ is a Hilbert space in which case $H^* = \{\phi_z := \langle \cdot | z \rangle : z \in H\}$, see Theorem 14.15. If $x, y \in H$ and $z := y - x \neq 0$, then

$$0 < \varepsilon := \|z\|^2 = \phi_z(z) = \phi_z(y) - \phi_z(x).$$

Thus

$$V_x := \{w \in H : |\phi_z(x) - \phi_z(w)| < \varepsilon/2\} \text{ and}$$
$$V_y := \{w \in H : |\phi_z(y) - \phi_z(w)| < \varepsilon/2\}$$

are disjoint sets from τ_w which contain x and y respectively. This shows that (H, τ_w) is a Hausdorff space. In particular, this shows that weak limits are unique if they exist.

Remark 14.50. Suppose that H is an infinite dimensional Hilbert space $\{x_n\}_{n=1}^\infty$ is an orthonormal subset of H. Then Bessel's inequality (Proposition 14.18) implies $x_n \overset{w}{\to} 0 \in H$ as $n \to \infty$. This points out the fact that if $x_n \overset{w}{\to} x \in H$ as $n \to \infty$, it is no longer necessarily true that $\|x\| = \lim_{n\to\infty} \|x_n\|$. However we do always have $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$ because,

$$\|x\|^2 = \lim_{n\to\infty} \langle x_n | x \rangle \leq \liminf_{n\to\infty} [\|x_n\| \|x\|] = \|x\| \liminf_{n\to\infty} \|x_n\|.$$

Proposition 14.51. *Let H be a Hilbert space, $\beta \subset H$ be an orthonormal basis for H and $\{x_n\}_{n=1}^\infty \subset H$ be a bounded sequence, then the following are equivalent:*

1. *$x_n \overset{w}{\to} x \in H$ as $n \to \infty$.*
2. *$\langle x|y \rangle = \lim_{n\to\infty} \langle x_n|y \rangle$ for all $y \in H$.*
3. *$\langle x|y \rangle = \lim_{n\to\infty} \langle x_n|y \rangle$ for all $y \in \beta$.*

Moreover, if $c_y := \lim_{n\to\infty} \langle x_n|y \rangle$ exists for all $y \in \beta$, then $\sum_{y\in\beta} |c_y|^2 < \infty$ and $x_n \overset{w}{\to} x := \sum_{y\in\beta} c_y y \in H$ as $n \to \infty$.

Proof. 1. $\Longrightarrow$ 2. This is a consequence of Theorem 14.15 and Proposition 14.49. 2. $\Longrightarrow$ 3. is trivial. 3. $\Longrightarrow$ 1. Let $M := \sup_n \|x_n\|$ and H_0 denote the algebraic span of β. Then for $y \in H$ and $z \in H_0$,

$$|\langle x - x_n|y \rangle| \leq |\langle x - x_n|z \rangle| + |\langle x - x_n|y - z \rangle| \leq |\langle x - x_n|z \rangle| + 2M \|y - z\|.$$

Passing to the limit in this equation implies $\limsup_{n\to\infty} |\langle x - x_n|y \rangle| \leq 2M \|y - z\|$ which shows $\limsup_{n\to\infty} |\langle x - x_n|y \rangle| = 0$ since H_0 is dense in H. To prove the last assertion, let $\Gamma \subset\subset \beta$. Then by Bessel's inequality (Proposition 14.18),

$$\sum_{y\in\Gamma} |c_y|^2 = \lim_{n\to\infty} \sum_{y\in\Gamma} |\langle x_n|y \rangle|^2 \leq \liminf_{n\to\infty} \|x_n\|^2 \leq M^2.$$

Since $\Gamma \subset\subset \beta$ was arbitrary, we conclude that $\sum_{y\in\beta} |c_y|^2 \leq M < \infty$ and hence we may define $x := \sum_{y\in\beta} c_y y$. By construction we have

$$\langle x|y \rangle = c_y = \lim_{n\to\infty} \langle x_n|y \rangle \text{ for all } y \in \beta$$

and hence $x_n \overset{w}{\to} x \in H$ as $n \to \infty$ by what we have just proved. ■

Theorem 14.52. *Suppose $\{x_n\}_{n=1}^\infty$ is a bounded sequence in a Hilbert space, H. Then there exists a subsequence $y_k := x_{n_k}$ of $\{x_n\}_{n=1}^\infty$ and $x \in X$ such that $y_k \overset{w}{\to} x$ as $k \to \infty$.*

Proof. This is a consequence of Proposition 14.51 and a Cantor's diagonalization argument which is left to the reader, see Exercise 14.11. ■

Theorem 14.53 (Alaoglu's Theorem for Hilbert Spaces). *Suppose that H is a separable Hilbert space, $C := \{x \in H : \|x\| \leq 1\}$ is the closed unit ball in H and $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis for H. Then*

$$\rho(x, y) := \sum_{n=1}^{\infty} \frac{1}{2^n} |\langle x - y | e_n \rangle| \tag{14.30}$$

defines a metric on C which is compatible with the weak topology on C, $\tau_C := (\tau_w)_C = \{V \cap C : V \in \tau_w\}$. Moreover (C, ρ) is a compact metric space. (This theorem will be extended to Banach spaces, see Theorems 35.14 and 35.15 below.)

Proof. The routine check that ρ is a metric is left to the reader. Let τ_ρ be the topology on C induced by ρ. For any $y \in H$ and $n \in \mathbb{N}$, the map $x \in H \to \langle x - y | e_n \rangle = \langle x | e_n \rangle - \langle y | e_n \rangle$ is τ_w continuous and since the sum in Eq. (14.30) is uniformly convergent for $x, y \in C$, it follows that $x \to \rho(x, y)$ is τ_C – continuous. This implies the open balls relative to ρ are contained in τ_C and therefore $\tau_\rho \subset \tau_C$. For the converse inclusion, let $z \in H$, $x \to \phi_z(x) = \langle x | z \rangle$ be an element of H^*, and for $N \in \mathbb{N}$ let $z_N := \sum_{n=1}^N \langle z | e_n \rangle e_n$. Then $\phi_{z_N} = \sum_{n=1}^N \overline{\langle z | e_n \rangle} \phi_{e_n}$ is ρ continuous, being a finite linear combination of the ϕ_{e_n} which are easily seen to be ρ – continuous. Because $z_N \to z$ as $N \to \infty$ it follows that

$$\sup_{x \in C} |\phi_z(x) - \phi_{z_N}(x)| = \|z - z_N\| \to 0 \text{ as } N \to \infty.$$

Therefore $\phi_z|_C$ is ρ – continuous as well and hence $\tau_C = \tau(\phi_z|_C : z \in H) \subset \tau_\rho$. The last assertion follows directly from Theorem 14.52 and the fact that sequential compactness is equivalent to compactness for metric spaces. ■

14.5 Supplement 1: Converse of the Parallelogram Law

Proposition 14.54 (Parallelogram Law Converse). *If $(X, \|\cdot\|)$ is a normed space such that Eq. (14.2) holds for all $x, y \in X$, then there exists a unique inner product on $\langle\cdot|\cdot\rangle$ such that $\|x\| := \sqrt{\langle x|x\rangle}$ for all $x \in X$. In this case we say that $\|\cdot\|$ is a Hilbertian norm.*

Proof. If $\|\cdot\|$ is going to come from an inner product $\langle\cdot|\cdot\rangle$, it follows from Eq. (14.1) that

$$2\mathrm{Re}\langle x|y\rangle = \|x + y\|^2 - \|x\|^2 - \|y\|^2$$

and

$$-2\mathrm{Re}\langle x|y\rangle = \|x - y\|^2 - \|x\|^2 - \|y\|^2.$$

Subtracting these two equations gives the "polarization identity,"

$$4\mathrm{Re}\langle x|y\rangle = \|x + y\|^2 - \|x - y\|^2.$$

Replacing y by iy in this equation then implies that

$$4\mathrm{Im}\langle x|y\rangle = \|x + iy\|^2 - \|x - iy\|^2$$

from which we find

$$\langle x|y\rangle = \frac{1}{4} \sum_{\varepsilon \in G} \varepsilon \|x + \varepsilon y\|^2 \tag{14.31}$$

where $G = \{\pm 1, \pm i\}$ – a cyclic subgroup of $S^1 \subset \mathbb{C}$. Hence if $\langle\cdot|\cdot\rangle$ is going to exists we must define it by Eq. (14.31). Notice that

$$\begin{aligned}\langle x|x\rangle &= \frac{1}{4} \sum_{\varepsilon \in G} \varepsilon \|x + \varepsilon x\|^2 = \|x\|^2 + i\|x + ix\|^2 - i\|x - ix\|^2 \\ &= \|x\|^2 + i\left|1 + i\right|^2 \|x\|^2 - i\left|1 - i\right|^2 \|x\|^2 = \|x\|^2.\end{aligned}$$

So to finish the proof of (4) we must show that $\langle x|y\rangle$ in Eq. (14.31) is an inner product. Since

$$\begin{aligned}4\langle y|x\rangle &= \sum_{\varepsilon \in G} \varepsilon \|y + \varepsilon x\|^2 = \sum_{\varepsilon \in G} \varepsilon \|\varepsilon(y + \varepsilon x)\|^2 \\ &= \sum_{\varepsilon \in G} \varepsilon \|\varepsilon y + \varepsilon^2 x\|^2 \\ &= \|y + x\|^2 + \|-y + x\|^2 + i\|iy - x\|^2 - i\|-iy - x\|^2 \\ &= \|x + y\|^2 + \|x - y\|^2 + i\|x - iy\|^2 - i\|x + iy\|^2 \\ &= 4\overline{\langle x|y\rangle}\end{aligned}$$

it suffices to show $x \to \langle x|y\rangle$ is linear for all $y \in H$. (The rest of this proof may safely be skipped by the reader.) For this we will need to derive an identity from Eq. (14.2). To do this we make use of Eq. (14.2) three times to find

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

Solving this equation for $\|x + y + z\|^2$ gives

$$\|x + y + z\|^2 = \|y + z\|^2 + \|x + y\|^2 - \|x - z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2. \tag{14.32}$$

Using Eq. (14.32), for $x, y, z \in H$,

$$\begin{aligned}4\,\mathrm{Re}\langle x + z|y\rangle &= \|x + z + y\|^2 - \|x + z - y\|^2 \\ &= \|y + z\|^2 + \|x + y\|^2 - \|x - z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2 \\ &\quad - \left(\|z - y\|^2 + \|x - y\|^2 - \|x - z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2\right) \\ &= \|z + y\|^2 - \|z - y\|^2 + \|x + y\|^2 - \|x - y\|^2 \\ &= 4\,\mathrm{Re}\langle x|y\rangle + 4\,\mathrm{Re}\langle z|y\rangle.\end{aligned} \tag{14.33}$$

Now suppose that $\delta \in G$, then since $|\delta| = 1$,

$$4\langle \delta x|y\rangle = \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\|\delta x+\varepsilon y\|^2 = \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\|x+\delta^{-1}\varepsilon y\|^2$$
$$= \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\delta\|x+\delta\varepsilon y\|^2 = 4\delta\langle x|y\rangle \tag{14.34}$$

where in the third inequality, the substitution $\varepsilon \to \varepsilon\delta$ was made in the sum. So Eq. (14.34) says $\langle \pm ix|y\rangle = \pm i\langle ix|y\rangle$ and $\langle -x|y\rangle = -\langle x|y\rangle$. Therefore

$$\operatorname{Im}\langle x|y\rangle = \operatorname{Re}\left(-i\langle x|y\rangle\right) = \operatorname{Re}\langle -ix|y\rangle$$

which combined with Eq. (14.33) shows

$$\operatorname{Im}\langle x+z|y\rangle = \operatorname{Re}\langle -ix-iz|y\rangle = \operatorname{Re}\langle -ix|y\rangle + \operatorname{Re}\langle -iz|y\rangle$$
$$= \operatorname{Im}\langle x|y\rangle + \operatorname{Im}\langle z|y\rangle$$

and therefore (again in combination with Eq. (14.33)),

$$\langle x+z|y\rangle = \langle x|y\rangle + \langle z|y\rangle \text{ for all } x,y\in H.$$

Because of this equation and Eq. (14.34) to finish the proof that $x \to \langle x|y\rangle$ is linear, it suffices to show $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ for all $\lambda > 0$. Now if $\lambda = m \in \mathbb{N}$, then

$$\langle mx|y\rangle = \langle x+(m-1)x|y\rangle = \langle x|y\rangle + \langle (m-1)x|y\rangle$$

so that by induction $\langle mx|y\rangle = m\langle x|y\rangle$. Replacing x by x/m then shows that $\langle x|y\rangle = m\langle m^{-1}x|y\rangle$ so that $\langle m^{-1}x|y\rangle = m^{-1}\langle x|y\rangle$ and so if $m,n\in\mathbb{N}$, we find

$$\langle \frac{n}{m}x|y\rangle = n\langle \frac{1}{m}x|y\rangle = \frac{n}{m}\langle x|y\rangle$$

so that $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ for all $\lambda > 0$ and $\lambda\in\mathbb{Q}$. By continuity, it now follows that $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ for all $\lambda > 0$. ■

14.6 Supplement 2. Non-complete inner product spaces

Part of Theorem 14.24 goes through when H is a not necessarily complete inner product space. We have the following proposition.

Proposition 14.55. *Let $(H, \langle\cdot|\cdot\rangle)$ be a not necessarily complete inner product space and $\beta\subset H$ be an orthonormal set. Then the following two conditions are equivalent:*

1. *$x = \sum_{u\in\beta}\langle x|u\rangle u$ for all $x\in H$.*

2. *$\|x\|^2 = \sum_{u\in\beta}|\langle x|u\rangle|^2$ for all $x\in H$.*

Moreover, either of these two conditions implies that $\beta\subset H$ is a maximal orthonormal set. However $\beta\subset H$ being a maximal orthonormal set is not sufficient to conditions for 1) and 2) hold!

Proof. As in the proof of Theorem 14.24, 1) implies 2). For 2) implies 1) let $\Lambda \subset\subset \beta$ and consider

$$\left\|x-\sum_{u\in\Lambda}\langle x|u\rangle u\right\|^2 = \|x\|^2 - 2\sum_{u\in\Lambda}|\langle x|u\rangle|^2 + \sum_{u\in\Lambda}|\langle x|u\rangle|^2$$
$$= \|x\|^2 - \sum_{u\in\Lambda}|\langle x|u\rangle|^2.$$

Since $\|x\|^2 = \sum_{u\in\beta}|\langle x|u\rangle|^2$, it follows that for every $\varepsilon > 0$ there exists $\Lambda_\varepsilon \subset\subset\beta$ such that for all $\Lambda\subset\subset\beta$ such that $\Lambda_\varepsilon\subset\Lambda$,

$$\left\|x-\sum_{u\in\Lambda}\langle x|u\rangle u\right\|^2 = \|x\|^2 - \sum_{u\in\Lambda}|\langle x|u\rangle|^2 < \varepsilon$$

showing that $x = \sum_{u\in\beta}\langle x|u\rangle u$. Suppose $x = (x_1, x_2, \dots, x_n, \dots)\in\beta^\perp$. If 2) is valid then $\|x\|^2 = 0$, i.e. $x = 0$. So β is maximal. Let us now construct a counter example to prove the last assertion. Take $H = \operatorname{Span}\{e_i\}_{i=1}^\infty \subset \ell^2$ and let $\tilde{u}_n = e_1 - (n+1)e_{n+1}$ for $n = 1, 2\dots$. Applying Gramn-Schmidt to $\{\tilde{u}_n\}_{n=1}^\infty$ we construct an orthonormal set $\beta = \{u_n\}_{n=1}^\infty \subset H$. I now claim that $\beta\subset H$ is maximal. Indeed if $x = (x_1, x_2, \dots, x_n, \dots)\in\beta^\perp$ then $x\perp u_n$ for all n, i.e.

$$0 = \langle x|\tilde{u}_n\rangle = x_1 - (n+1)x_{n+1}.$$

Therefore $x_{n+1} = (n+1)^{-1}x_1$ for all n. Since $x\in\operatorname{Span}\{e_i\}_{i=1}^\infty$, $x_N = 0$ for some N sufficiently large and therefore $x_1 = 0$ which in turn implies that $x_n = 0$ for all n. So $x = 0$ and hence β is maximal in H. On the other hand, β is not maximal in ℓ^2. In fact the above argument shows that $\beta^\perp$ in ℓ^2 is given by the span of $v = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots)$. Let P be the orthogonal projection of ℓ^2 onto the $\operatorname{Span}(\beta) = v^\perp$. Then

$$\sum_{i=1}^\infty\langle x|u_n\rangle u_n = Px = x - \frac{\langle x|v\rangle}{\|v\|^2}v,$$

so that $\sum_{i=1}^\infty\langle x|u_n\rangle u_n = x$ iff $x\in\operatorname{Span}(\beta) = v^\perp\subset\ell^2$. For example if $x = (1,0,0,\dots)\in H$ (or more generally for $x = e_i$ for any i), $x\notin v^\perp$ and hence $\sum_{i=1}^\infty\langle x|u_n\rangle u_n \neq x$. ■

14.7 Exercises

Exercise 14.11. Prove Theorem 14.52. **Hint:** Let $H_0 := \overline{\text{span}\{x_n : n \in \mathbb{N}\}}$ – a separable Hilbert subspace of H. Let $\{\lambda_m\}_{m=1}^\infty \subset H_0$ be an orthonormal basis and use Cantor's diagonalization argument to find a subsequence $y_k := x_{n_k}$ such that $c_m := \lim_{k\to\infty} \langle y_k|\lambda_m\rangle$ exists for all $m \in \mathbb{N}$. Finish the proof by appealing to Proposition 14.51.

Exercise 14.12. Suppose that $\{x_n\}_{n=1}^\infty \subset H$ and $x_n \xrightarrow{w} x \in H$ as $n \to \infty$. Show $x_n \to x$ as $n \to \infty$ (i.e. $\lim_{n\to\infty} \|x - x_n\| = 0$) iff $\lim_{n\to\infty} \|x_n\| = \|x\|$.

Exercise 14.13 (Banach-Saks). Suppose that $\{x_n\}_{n=1}^\infty \subset H$, $x_n \xrightarrow{w} x \in H$ as $n \to \infty$, and $c := \sup_n \|x_n\| < \infty$.[4] Show there exists a subsequence, $y_k = x_{n_k}$ such that

$$\lim_{N\to\infty} \left\| x - \frac{1}{N}\sum_{k=1}^N y_k \right\| = 0,$$

i.e. $\frac{1}{N}\sum_{k=1}^N y_k \to x$ as $N \to \infty$. **Hints:** 1. show it suffices to assume $x = 0$ and then choose $\{y_k\}_{k=1}^\infty$ so that $|\langle y_k|y_l\rangle| \le l^{-1}$ (or even smaller if you like) for all $k \le l$.

Exercise 14.14 (The Mean Ergodic Theorem). Let $U : H \to H$ be a unitary operator on a Hilbert space H, $M = \text{Nul}(U - I)$, $P = P_M$ be orthogonal projection onto M, and $S_n = \frac{1}{n}\sum_{k=0}^{n-1} U^k$. Show $S_n \to P_M$ **strongly**, i.e. $\lim_{n\to\infty} S_n x = P_M x$ for all $x \in H$. **Hints:** 1. verify the result for $x \in \text{Nul}(U - I)$ and $x \in \text{Ran}(U - I)$, 2. show $\text{Nul}(U^* - I) = \text{Nul}(U - I)$, 3. finish the result with a limiting argument making use of items 1. and 2. and Lemma 14.17.

[4] The assumption that $c < \infty$ is superfluous because of the "uniform boundedness principle," see Theorem 35.8 below.