Chapter 2. Hilbert Spaces

Review of Inner Product Spaces from Linear Algebra

2.1 Definition: Let V be a vector space (over any field $\mathbb{F}$). Recall that a (Hamel) **basis** for V is a maximal linearly independent set in V or, equivalently, a linearly independent set which spans V. Also recall that any two Hamel bases for V have the same cardinality, and we define the (Hamel) **dimension** of V, denoted by $\dim(V)$, to be the cardinality of any Hamel basis.

2.2 Definition: Let V be an inner product space (over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). For a subset $\mathcal{B} \subseteq V$, we say $\mathcal{B}$ is **orthogonal** when $\langle u, v\rangle = 0$ for all $u, v \in \mathcal{B}$ with $u \neq v$, and we say $\mathcal{B}$ is **orthonormal** when $\mathcal{B}$ is orthogonal with $\|u\| = 1$ for every $u \in \mathcal{B}$. For a finite or countable ordered set $\mathcal{B} = (u_1, u_2, u_3, \cdots)$ in V we say $\mathcal{B}$ is **orthogonal** when $\langle u_k, u_\ell\rangle = 0$ for all $k \neq \ell$, and we say $\mathcal{B}$ is **orthonormal** when it is orthogonal and $\|u_k\| = 1$ for all k.

2.3 Theorem: *Let V be an inner product space. Let $\mathcal{B} \subseteq V$ be orthonormal. Let $x, y \in \operatorname{Span}\mathcal{B}$ with say $x = \sum_{k=1}^{n} a_k u_k$ and $y = \sum_{k=1}^{n} b_k u_k$ where $a_k, b_k \in \mathbb{F}$ and $u_k \in \mathcal{B}$. Then*

$$\langle x, u_k\rangle = a_k \ , \ \langle x, y\rangle = \sum_{k=1}^{n} a_k \overline{b_k} \ , \text{ and } \|x\|^2 = \sum_{k=1}^{n} |a_k|^2.$$

In particular, $\mathcal{B}$ is linearly independent.

Proof: We omit the proof.

2.4 Theorem: *(The Gram-Schmidt Procedure) Let V be an inner product space, which is of finite or countable Hamel dimension. Let $\mathcal{A} = (u_1, u_2, u_3, \cdots)$ be a finite or countable ordered Hamel basis for V. Let $v_1 = u_1$ and for $n \geq 2$ let $v_n = u_n - \sum_{k=1}^{n-1} \frac{\langle u_n, v_k\rangle}{\|v_k\|^2} v_k$. Then $\mathcal{B} = (v_1, v_2, v_3, \cdots)$ is an orthogonal Hamel basis for V with the property that for every index $n \geq 1$ we have $\operatorname{Span}\{v_1, \cdots, v_n\} = \operatorname{Span}\{u_1, \cdots, u_n\}$.*

Proof: We omit the proof

2.5 Corollary: *Every inner product space which is of finite or countable Hamel dimension has an orthonormal Hamel basis.*

2.6 Corollary: *Let V be an inner product space which is of finite or countable Hamel dimension. Let $U \subseteq V$ be a finite dimensional subspace. Then every orthogonal (or orthonormal) Hamel basis $\mathcal{B}$ for U extends to an orthogonal (or orthonormal) Hamel basis for V.*

2.7 Corollary: *Let U and V be inner product spaces of finite or countable Hamel dimension. Then U and V are isomorphic (as inner product spaces) if and only if $\dim(U) = \dim(V)$. In particular, if $\dim(U) = n$ then U is isomorphic to $\mathbb{F}^n$ and if $\dim(U) = \aleph_0$ then U is isomorphic to $\mathbb{F}^\infty$ (which is the space of sequences in $\mathbb{F}$ with only finitely many nonzero terms, using the inner product $\langle x, y\rangle = \sum_{k=1}^{\infty} x_k \overline{y_k}$).*

2.8 Corollary: *Every finite-dimensional inner product space is complete, and every inner product space which is of countable Hamel dimension is not complete.*

2.9 Definition: *When W is a vector space (over any field $\mathbb{F}$) and $U, V \subseteq W$ are subspaces, we write $U + V = \big\{u + v \,\big|\, u \in U, v \in V\big\}$ and we write $W = U \oplus V$ when $W = U + V$ and $U \cap V = \{0\}$, that is when for every $x \in W$, $x = u + v$ for some unique $u \in U$, $v \in V$.*

2.10 Definition: *Let V be an inner product space (over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). For a subspace $U \subseteq V$, we define the* **orthogonal complement** *of U in V to be the set*

$$U^\perp = \big\{x \in V \big| \langle x, u\rangle = 0 \text{ for all } u \in U\big\}.$$

2.11 Theorem: *Let V be an inner product space and let $U \subseteq V$ be a subspace. Then*
(1) $U^\perp$ is a subspace of V,
(2) if $\mathcal{B}$ is a basis for U then $U^\perp = \big\{x \in V \big| \langle x, u\rangle = 0 \text{ for all } u \in \mathcal{B}\big\}$,
(3) $U \cap U^\perp = \{0\}$, and
(4) $U \subseteq (U^\perp)^\perp$.
(5) if U is finite-dimensional then $U \oplus U^\perp = V$, and
(6) if $U \oplus U^\perp = W$ then $U = (U^\perp)^\perp$.

Proof: We omit the proof.

2.12 Definition: Let V be an inner product space. Let $U \subseteq V$ be a subspace such that $V = U \oplus U^\perp$. For $x \in V$, we define the **orthogonal projection** of x onto U, denoted by $\mathrm{Proj}_U(x)$, as follows. Since $V = U \oplus U^\perp$, we can choose unique vectors $u, v \in V$ with $u \in U$, $v \in V$ and $u + v = x$. We then define

$$\mathrm{Proj}_U(x) = u.$$

When U is finite-dimensional so $U = (U^\perp)^\perp$, for u and v as above we have $\mathrm{Proj}_{U^\perp}(x) = v$. When $y \in V$ and $U = \mathrm{Span}\,\{y\}$, we also write $\mathrm{Proj}_y(x) = \mathrm{Proj}_U(x)$.

2.13 Theorem: *Let V be an inner product space. Let $U \subseteq V$ be a subspace of V such that $V = U \oplus U^\perp$. Let $x \in V$. Then $\mathrm{Proj}_U(x)$ is the unique point in U nearest to x.*

Proof: We omit the proof.

2.14 Example: Let V be an inner product space. Let U be a finite dimensional subspace of V. Let $\mathcal{B} = \{u_1, u_2, \cdots, u_n\}$ be an orthogonal basis for U. Recall (or verify) that

$$\mathrm{Proj}_U(x) = \sum_{k=1}^{n} \frac{\langle x, u_k\rangle}{\|u_k\|^2}\, u_k.$$

2.15 Example: Recall (or verify) that for $A \in M_{n\times m}(\mathbb{C})$ and $U = \mathrm{Col}(A)$, given $x \in \mathbb{C}^n$ there exists $y \in \mathbb{C}^m$ such that $A^*Ay = A^*x$ and for any such y, we have $\mathrm{Proj}_U(x) = Ay$. In particular, if $\mathrm{rank}(A) = m$ then A^*A is invertible so that $\mathrm{Proj}_U(x) = A(A^*A)^{-1}A^*x$.

2.16 Note: Let W be an inner product space and let $U \subseteq W$ be a subspace. Note that $\overline{U}$ is also a subspace because given $u, v \in \overline{U}$ and $t \in \mathbb{F}$ we can choose sequences (x_n) and (y_n) in U with $x_n \to u$ and $y_n \to v$ and then we have $(x_n + ty_n \to u_t v$ so that $u + tv \in \overline{U})$ Also note that $\overline{U}^\perp = U^\perp$. Indeed, since $U \subseteq \overline{U}$ we have $\overline{U}^\perp \subseteq U^\perp$ so it suffices to prove that $U^\perp \subseteq \overline{U}^\perp$ Let $v \in U^\perp$ and let $u \in \overline{U}$. Choose a sequence (x_n) in U with $x_n \to u$. Then we have $\langle v, u\rangle = \langle v, \lim\limits_{n\to\infty} x_n\rangle = \lim\limits_{n\to\infty} \langle v, x_n\rangle = 0$, indeed

$$\big|\langle v, u\rangle\big| = \big|\langle v, u\rangle - \langle v, x_n\rangle\big| = \big|\langle v, u - x_n\rangle\big| \le \|v\|\,\|u - x_n\| \to 0.$$

Closed Subspaces of Hilbert Spaces and Orthogonal Projections

2.17 Example: Properties of finite-dimensional subspaces of inner product spaces do not always carry over to infinite dimensional subspaces. For example, let $V=\mathbb{F}^\infty$ (the space of sequences in $\mathbb{F}$ with finitely many nonzero terms) with inner product $\langle x,y\rangle=\sum\limits_{k=1}^\infty x_k\overline{y_k}$, and let $U=\left\{a\in\mathbb{F}^\infty\middle|\sum\limits_{k=1}^\infty a_k=0\right\}$. Note that V has countable Hamel dimension with standard orthonormal Hamel basis $\mathcal{S}=\{e_1,e_2,e_3,\cdots\}$, and $U\subsetneqq V$ has countable Hamel dimension, with Hamel basis $\mathcal{B}=\{u_1,u_2,u_3,\cdots\}$ where $u_k=e_1-e_{k+1}$. We have

$$\begin{aligned}U^\perp&=\left\{x\in V\middle|\langle x,u_k\rangle=0\text{ for all }k\right\}=\left\{x\in V\middle|\langle x,e_1-e_{k+1}\rangle=0\text{ for all }k\right\}\\&=\left\{x\in V\middle|x_1=x_{k+1}\text{ for all }k\right\}=\left\{x\in V\middle|x_1=x_2=x_3=\cdots\right\}=\{0\}\end{aligned}$$

because for $x\in\mathbb{F}^\infty$ we have $x_n=0$ for all but finitely many indices n. Notice that in this example we have $U\subsetneqq(U^\perp)^\perp=V$ and $V\neq U\oplus U^\perp$. Although we could apply the Gram-Schmidt Procedure to $\mathcal{B}$ to obtain an orthogonal Hamel basis $\mathcal{C}=\{v_1,v_2,\cdots\}$ for U, we cannot extend $\mathcal{C}$ to an orthogonal Hamel basis for V because there is no nonzero vector $0\neq x\in V$ with $\langle x,v_k\rangle=0$ for all k.

2.18 Definition: Let V be a vector space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. For a subset $S\subseteq V$, we say that S is **convex** when for all $a,b\in S$ we have $a+t(b-a)\in S$ for all $0\leq t\leq 1$.

2.19 Theorem: *Let H be a Hilbert space. Let $S\subseteq H$ be nonempty, closed and convex. Then for every $a\in H$ there exists a unique point $b\in S$ which is nearest to a, that is such that $\|a-b\|\leq\|a-x\|$ for all $x\in S$.*

Proof: Let $a\in H$. Let $d=\text{dist}(a,S)=\inf\left\{\|x-a\|\,\middle|\,x\in S\right\}$. Choose a sequence $\{x_n\}$ in S so that $\|x_n-a\|\to d$, hence $\|x_n-a\|^2\to d^2$. Let $\epsilon>0$ and choose $m\in\mathbf{Z}^+$ so that for all $n\geq m$ we have $\|x_n-a\|^2\leq d^2+\frac{\epsilon^2}{4}$. Let $k,l\geq m$. By the Parallelogram Law we have

$$\big\|(x_k-a)+(x_l-a)\big\|^2+\big\|(x_k-a)-(x_l-a)\big\|^2=2\big\|x_k-a\big\|^2+2\big\|x_l-a\big\|^2$$

Since S is convex, we have $\frac{x_k+x_l}{2}\in S$, hence $\big\|\frac{x_k+x_l}{2}-a\big\|\geq d$, and so

$$\begin{aligned}\big\|x_k-x_l\big\|^2&=\big\|(x_k-a)-(x_l-a)\big\|^2\\&=2\big\|x_k-a\big\|^2+2\big\|x_l-a\big\|^2-\big\|(x_k-a)+(x_l-a)\big\|^2\\&=2\big\|x_k-a\big\|^2+2\big\|x_l-a\big\|^2-4\big\|\tfrac{x_k+x_l}{2}-a\big\|^2\\&\leq2\big(d^2+\tfrac{\epsilon^2}{4}\big)+2\big(d^2+\tfrac{\epsilon^2}{4}\big)-4d^2=\epsilon^2.\end{aligned}$$

so that $\|x_k-x_l\|\leq\epsilon$. This shows that the sequence $\{x_n\}$ is Cauchy. Since H is complete, $\{x_n\}$ converges in H, and since S is closed in H, the limit lies in S. Let $b=\lim\limits_{n\to\infty}x_n\in S$. Since $b\in S$ we have $\|b-a\|\geq d$, and we have $\|b-a\|\leq\|b-x_n\|+\|x_n-a\|$ for all $n\in\mathbf{Z}^+$ so that $\|b-a\|\leq\lim\limits_{n\to\infty}\big(\|b-x_n\|+\|x_n-a\|\big)=d$, and so $\|b-a\|=d$. This shows that $\|b-a\|\geq\|x-a\|$ for all $x\in S$. Finally, we note that the point b is unique because given $c\in S$ with $\|c-a\|=d$, since S is convex we have $\frac{b+c}{2}\in S$ so that $\big\|\frac{b+c}{2}-a\big\|\geq d$, and so the Parallelogram Law gives

$$\begin{aligned}\|b-c\|^2&=\big\|(b-a)-(c-a)\big\|^2=2\|b-a\|^2+2\|c-a\|^2-\big\|(b-a)+(c-a)\big\|\\&=4d^2-4\big\|\tfrac{b-c}{2}-a\big\|^2\leq4d^2-4d^2=0\end{aligned}$$

so that $\|b-c\|=0$ hence $b=c$.

2.20 Theorem: *Let H be a Hilbert space. Let $U \subseteq H$ be a subspace. Then U is closed if and only if $H = U \oplus U^\perp$. In this case, $U^\perp$ is closed and $(U^\perp)^\perp = U$ and for $x \in H$, if $x = u + v$ with $u \in U$ and $v \in U^\perp$ then u is the unique point in U nearest to x and v is the unique point in $U^\perp$ nearest to x.*

Proof: Suppose that $H = U \oplus U^\perp$. Let $(x_n)_{n \geq 1}$ be a sequence in U which converges in H, say $x_n \to a \in H$. Note that since $x_n \to a$ in H, we have $\langle x_n, y\rangle \to \langle a, y\rangle$ for all $y \in H$ because $\big|\langle x_n, y\rangle - \langle a, y\rangle\big| = \big|\langle x_n - a, y\rangle\big| \leq \|x_n - a\|\,\|y\|$. Since $H = U \oplus U^\perp$, we can write $a = u + v$ with $u \in U$ and $v \in U^\perp$. Then, since $\langle u, v\rangle = 0$, we have

$$\|v\|^2 = \langle v, v\rangle = \langle u + v, v\rangle = \langle a, v\rangle = \lim_{n\to\infty} \langle x_n, v\rangle = \lim_{n\to\infty} 0 = 0$$

and so $v = 0$ so that $a = u + v = u \in U$. Thus U is closed.

Suppose that U is closed. Let $x \in H$. Since U is a vector space it is convex, so by the previous theorem there is a unique point $u \in U$ which is nearest to x. Let u be this nearest point and let $v = x - u$ so that $u + v = x$. We claim that $v \in U^\perp$. Suppose, for a contradiction, that $v \notin U^\perp$. Choose $u_1 \in U$ with $\langle v, u_1\rangle \neq 0$. Write $\langle v, u_1\rangle = re^{i\theta}$ with $r > 0$ and $\theta \in \mathbb{R}$ (when $\mathbb{F} = \mathbb{R}$ we have $e^{i\theta} = \pm 1$) and let $u_2 = e^{i\theta} u_1$. Note that $u_2 \in U$ and $\langle v, u_2\rangle = \langle v, e^{i\theta} u_1\rangle = e^{-i\theta}\langle v, u_1\rangle = e^{-i\theta} r\, e^{i\theta} = r > 0$. For all $t \in \mathbb{R}$ we have

$$\big\|x - (u + tu_2)\big\|^2 = \|v - tu_2\|^2 = \|v\|^2 - 2t\,\mathrm{Re}\,\langle v, u_2\rangle + t^2\|u_2\|^2 = \|v\|^2 - 2r\,t + \|u_2\|^2 t^2.$$

It follows that for small $t > 0$ we have $\big\|x - (u + tu_2)\big\|^2 \leq \|v\|^2 = \|x - u\|^2$ which is not possible, since u is the point in U which is nearest to x.

We claim that the points $u \in U$ and $v \in U^\perp$ with $u + v = x$, which we found in the previous paragraph, are the only such points. Let $x \in H$. Suppose that $u \in U$, $v \in U^\perp$ and $u + v = x$. We claim that u must be equal to the (unique) point in U which is nearest to x. Let $u' \in U$ with $u' \neq u$. Since $v \in U^\perp$ and $u' - u \in U$ we have $\langle x - u, u' - u\rangle = \langle v, u' - u\rangle = 0$ and so

$$\begin{aligned}\big\|x - u'\big\|^2 &= \big\|(x - u) - (u' - u)\big\|^2 = \|x - u\|^2 - 2\mathrm{Re}\,\langle x - u, u' - u\rangle + \|u' - u\|^2 \\ &= \|x - u\| + \|u' - u\| > \|x - u\|^2\end{aligned}$$

so that $\|x - u'\| > \|x - u\|$. Thus u is the point in U which is nearest to x, so u (hence also v) is uniquely determined. Thus we have $H = U \oplus U^\perp$.

We claim that since $H = U \oplus U^\perp$, it follows that $(U^\perp)^\perp = U$. We always have $U \subseteq (U^\perp)^\perp$, so we only need to show that $(U^\perp)^\perp \subseteq U$. Let $x \in (U^\perp)^\perp$. Choose $u \in U$ and $v \in U^\perp$ such that $x = u + v$. Since $u \in U$ and $v \in U^\perp$ we have $\langle u, v\rangle = 0$, and since $v \in U^\perp$ and $x \in (U^\perp)^\perp$ we have $\langle x, v\rangle = 0$. It follows that

$$\|v\|^2 = \langle v, v\rangle = \langle u, v\rangle + \langle v, v\rangle = \langle u + v, v\rangle = \langle x, v\rangle = 0$$

and so $v = 0$ and hence $x = u + v = u \in U$. Thus $(U^\perp)^\perp \subseteq U$, as required.

Finally note that since $H = U \oplus U^\perp$ and $(U^\perp)^\perp = U$ we have $H = U^\perp \oplus (U^\perp)^\perp$, and so $U^\perp$ is closed, as proven in the first paragraph (with U replaced by $U^\perp$).

2.21 Definition: When H is a Hilbert space and $U \subseteq H$ is a closed subspace, we define the **orthogonal projection** onto U to be the map $P : H \to U$ given by $Px = u$ where u is the unique point in U nearest to x. Equivalently, $Px = u$ where $x = u + v$ with $u \in U$ and $v \in U^\perp$.

Unordered Series

2.22 Definition: Let $(a_n)_{n\geq 1}$ be a sequence in a normed linear space V. We say that $\sum\limits_{k=1}^{\infty} a_k$ **converges absolutely** in V when $\sum\limits_{k=1}^{\infty} \|a_k\|$ converges in $\mathbb{R}$, and we say that $\sum\limits_{k=1}^{\infty} a_k$ **converges unconditionally** in V when $\sum\limits_{n=1}^{\infty} a_{\sigma(n)}$ converges in V for every bijective map $\sigma : \mathbf{Z}^+ \to \mathbf{Z}^+$ (that is, when every rearrangement of the series converges).

2.23 Example: Recall that for a sequence $(a_n)_{n\geq 1}$ in $\mathbb{R}$, the series $\sum\limits_{n=1}^{\infty} |a_n|$ converges if and only if $\sum\limits_{n=1}^{\infty} a_{\sigma(n)}$ converges for every bijective map $\sigma : \mathbf{Z}^+ \to \mathbf{Z}^+$ so, in $\mathbb{R}$, unconditional convergence is the same thing as absolute convergence. But verify that in ℓ_2, the series $\sum\limits_{n=1}^{\infty} \frac{1}{n}\, e_n$ converges unconditionally to $\left(\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \cdots\right)$, but it does not converge absolutely.

2.24 Definition: Let K be a nonempty set (possibly uncountable). When A is any set, an **indexed set** in A with index set K is a function $a : K \to A$, and we write $a_k = a(k)$ and $a = (a_k)_{k\in K}$. When X is a normed linear space and $(a_k)_{k\in K}$ is an indexed set in X, the **unordered series** $\sum\limits_{k\in K} a_k$ is defined to be the indexed set $(s_F)_{F\in \mathrm{Fin}(K)}$ where $\mathrm{Fin}(K)$ is the set of finite subsets of K and $s_F = \sum\limits_{k\in F} a_k$ for each $F \in \mathrm{Fin}(K)$. We say that the unordered series $\sum\limits_{k\in K} a_k$ **converges** (unconditionally) in X when there exists $s \in X$ such that

$$\forall\, \epsilon > 0 \ \ \exists\, F \in \mathrm{Fin}(K) \ \ \forall\, I \in \mathrm{Fin}(K) \ \left(I \supseteq F \Longrightarrow \|s_I - s\| < \epsilon\right).$$

In this case, the element $s \in X$ is unique, it is called the (unordered) **sum** of the unordered series $\sum\limits_{k\in K} a_k$, and we write $\sum\limits_{k\in K} a_k = s$. As usual, we write $\sum\limits_{k\in K} a_k$ both to denote the unordered series (which may or may not converge) and its sum (when it does converge).

We say that the unordered series $\sum\limits_{k\in K} a_k$ **converges absolutely** in X when $\sum\limits_{k\in K} \|a_k\|$ converges in $\mathbb{R}$.

When $(a_k)_{k\in K}$ is an indexed set in $\mathbb{R}$ with each $a_k \geq 0$, whether or not the unordered series $\sum\limits_{k\in K} a_k$ converges, we define its (unordered) **sum** to be

$$s = \sup\left\{ \sum_{k\in F} a_k \,\middle|\, F \in \mathrm{Fin}(K)\right\}$$

and we write $\sum\limits_{k\in K} a_k = s$. Verify, as an exercise, that the unordered series converges if and only if its sum is finite and that, in this case, our two definitions of the sum agree.

2.25 Exercise: Let X be a normed linear space. Show that X is a Banach space if and only if it has the property that every absolutely convergent unordered series in X converges.

2.26 Theorem: *Let $(a_k)_{k\in K}$ be an indexed set in $\mathbb{R}$ with each $a_k \geq 0$. If $\sum\limits_{k\in K} a_k$ converges then there are at most countably many indices $k \in K$ for which $a_k \neq 0$.*

Proof: For each $n \in \mathbf{Z}^+$, let $K_n = \{k \in K \,|\, a_k \geq \frac{1}{n}\}$. If one of the set K_n was infinite we would have $\sum\limits_{k\in K} a_k = \infty$. Thus if $\sum\limits_{k\in K} a_k < \infty$, then every set K_n is finite, and so the set $\{k \in K \,|\, a_k > 0\} = \bigcup\limits_{n=1}^{\infty} K_n$ is at most countable.

2.27 Definition: Let $(a_k)_{k\in K}$ be an indexed set in a normed linear space X. We say that the unordered series $\sum\limits_{k\in K} a_k$ is **Cauchy** when

$$\forall \epsilon > 0 \;\; \exists F \in \text{Fin}(K) \;\; \forall I, J \in \text{Fin}(K) \; \big(I, J \supseteq F \Longrightarrow \|s_I - s_J\| < \epsilon\big).$$

As an exercise, verify that $\sum\limits_{k\in K} a_k$ is Cauchy if and only if

$$\forall \epsilon > 0 \;\; \exists F \in \text{Fin}(K) \;\; \forall L \in \text{Fin}(K) \; \big(L \cap F = \emptyset \Longrightarrow \|s_L\| < \epsilon\big).$$

2.28 Theorem: *(Cauchy Criterion for Unordered Series) Let $(a_k)_{k\in K}$ be an indexed set in a normed linear space X.*

(1) If $\sum\limits_{k\in K} a_k$ converges in X then it is Cauchy.

(2) If X is complete and $\sum\limits_{k\in K} a_k$ is Cauchy, then $\sum\limits_{k\in K} a_k$ converges in X.

Proof: To prove Part 1, suppose that $\sum\limits_{k\in K} a_k$ converges in X, say $s = \sum\limits_{k\in K} a_k$. Let $\epsilon > 0$. Choose $F \in \text{Fin}(K)$ such that for all $I \in \text{Fin}(K)$ with $I \supseteq F$ we have $\|s_I - s\| < \frac{\epsilon}{2}$. Let $I, J \in \text{Fin}(K)$ with $I, J \supseteq F$. Then we have $\|s_I - s_J\| \leq \|s_I - s\| + \|s - s_J\| < \epsilon$, and so $\sum\limits_{k\in K} a_k$ is Cauchy.

To prove Part 2, suppose X is complete and $\sum\limits_{k\in K} a_k$ is Cauchy in X. Since $\sum\limits_{k\in K} a_k$ is Cauchy, we can choose sets $F_n \in \text{Fin}(K)$ with $F_1 \subseteq F_2 \subseteq F_3 \subseteq \cdots$ such that for all $I, J \in \text{Fin}(K)$ with $I, J \supseteq F_n$ we have $\|s_I - s_J\| < \frac{1}{2^n}$ (indeed, having chosen F_n we can choose $G_n \in \text{Fin}(K)$ so that $G_n \subseteq I, J \in \text{Fin}(K) \Longrightarrow \|s_I - s_J\| < \frac{1}{2^{n+1}}$ and then set $F_{n+1} = F_n \cup G_n$). Then the sequence $(s_{F_n})_{n\geq 1}$ is Cauchy in X $\big($because when $\ell > m$ we have $\|s_{F_\ell} - s_{F_m}\| \leq \|s_{F_\ell} - s_{F_{\ell-1}}\| + \cdots + \|s_{F_{m+1}} - s_{F_m}\| < \frac{1}{2^{\ell-1}} + \cdots + \frac{1}{2^m} < \frac{1}{2^{m-1}}\big)$. Since X is complete, the sequence $(s_{F_n})_{n\geq 1}$ converges, say $s_{F_n} \to s$ in X. We claim that $\sum\limits_{k\in K} a_k = s$. Let $\epsilon > 0$. Choose $m \in \mathbf{Z}^+$ with $\frac{1}{2^m} < \frac{\epsilon}{2}$ such that $n \geq m \Longrightarrow \|s_{F_n} - s\| < \frac{\epsilon}{2}$. Then for all $I \in \text{Fin}(K)$ with $I \supseteq F_m$ we have $\|s_I - s\| \leq \|s_I - s_{F_m}\| + \|s_{F_m} - s\| < \frac{1}{2^m} + \frac{\epsilon}{2} < \epsilon$. Thus $\sum\limits_{k\in K} a_k = s$, as claimed.

Formulas Involving Orthonormal Indexed Sets

2.29 Definition: An indexed set $(u_k)_{k\in K}$ in an inner product space V is called **orthonormal** when $\|u_k\|=1$ for all $k\in K$ and $\langle u_k,u_\ell\rangle=0$ for all $k,\ell\in K$ with $k\neq\ell$.

2.30 Theorem: *Let H be a Hilbert space. Let $(u_k)_{k\in K}$ be an orthonormal indexed set in H, and let $\mathcal{B}=\{u_k\,|\,k\in K\}$. Let $x,y\in\overline{\operatorname{Span}\mathcal{B}}\in H$ and for each $k\in K$, let $a_k=\langle x,u_k\rangle$ and $b_k=\langle y,u_k\rangle$. Then*

(1) $\sum\limits_{k\in K}a_ku_k=x$,

(2) $\sum\limits_{k\in K}|a_k|^2=\|x\|^2$ *and*

(3) $\sum\limits_{k\in K}a_k\overline{b_k}=\langle x,y\rangle$.

Proof: Let us prove Part 1. For each $F\in\operatorname{Fin}(K)$, let $U_F=\operatorname{Span}\{u_k\,|\,k\in F\}$ and let $s_F=\sum\limits_{k\in F}a_ku_k=\operatorname{Proj}_{U_F}(x)$. Let $\epsilon>0$. Since $x\in\overline{\operatorname{Span}\mathcal{B}}$ we can choose $u\in\operatorname{Span}\mathcal{B}$ with $\|u-x\|<\epsilon$. Since elements in $\operatorname{Span}\mathcal{B}$ are finite linear combinations of elements in $\mathcal{B}$, we have $u\in U_F$ for some finite index set $F\in\operatorname{Fin}(K)$. For $I\in\operatorname{Fin}(K)$ with $I\supseteq F$, since s_I is the point in U_I nearest to x, and since we have $u\in U_F\subseteq U_I$, it follows that

$$\Big\|\sum_{k\in I}a_ku_k-x\Big\|=\|s_I-x\|\leq\|u-x\|<\epsilon.$$

Thus $\sum\limits_{k\in K}a_ku_k=x$ in H, as required.

Let us prove Part 2. For each $F\in\operatorname{Fin}(K)$, let $s_F=\sum\limits_{k\in F}a_ku_k$. Let $\epsilon>0$. Since $\sum\limits_{k\in K}a_ku_k=x$, so we can choose $F\in\operatorname{Fin}(K)$ such that for all $I\in\operatorname{Fin}(K)$ with $I\supseteq F$ we have $\|s_I-x\|<\min\left\{1,\frac{\epsilon}{2\|x\|+1}\right\}$. Since $\|s_I-x\|<1$ we have $\|s_I\|<\|x\|+1$ hence $\|s_I\|+\|x\|<2\|x\|+1$, and since $\|s_I-x\|<\frac{\epsilon}{2\|x\|+1}$ we have $\big|\|s_I\|-\|x\|\big|\leq\|s_I-x\|<\frac{\epsilon}{2\|x\|+1}$, and so

$$\Big|\sum_{k\in I}|a_I|^2-\|x\|^2\Big|=\big|\|s_I\|^2-\|x\|^2\big|=\big|\|s_I\|-\|x\|\big|\big(\|s_I\|+\|x\|\big)<\epsilon.$$

Thus $\sum\limits_{k\in K}|a_k|^2=\|x\|^2$, as required.

Let us prove Part 3. For each $F\in\operatorname{Fin}(K)$, let $r_F=\sum\limits_{k\in F}a_ku_k$ and $s_I=\sum\limits_{k\in F}b_ku_k$. Note that by Part 2, we have $\|r_I\|^2=\sum\limits_{k\in F}|a_k|^2\leq\sum\limits_{k\in K}|a_k|^2=\|x\|^2$ so that $\|r_F\|\leq\|x\|$, and similarly $\|s_F\|\leq\|y\|$. Let $\epsilon>0$. By Part 1, we can choose $F\in\operatorname{Fin}(K)$ such that for all $I\in\operatorname{Fin}(K)$ with $I\supseteq F$ we have $\|r_I-x\|<\frac{\epsilon}{\|y\|+1}$ and $\|s_I-y\|<\frac{\epsilon}{\|x\|+1}$. Then for $F\subseteq I\in\operatorname{Fin}(K)$ we have

$$\begin{aligned}\Big|\sum_{k\in I}a_k\overline{b_k}-\langle x,y\rangle\Big|&=\big|\langle r_I,s_I\rangle-\langle x,y\rangle\big|=\Big|\langle r_I,s_I\rangle-\langle x,s_I\rangle+\langle x,s_I\rangle-\langle x,y\rangle\Big|\\&\leq\big|\langle r_I,s_I\rangle-\langle x,s_I\rangle\big|+\big|\langle x,s_I\rangle-\langle x,y\rangle\big|=\big|\langle r_i-x,s_I\rangle\big|+\big|\langle x,s_I-y\rangle\big|\\&\leq\|r_I-x\|\,\|s_I\|+\|x\|\,\|s_I-y\|\leq\|r_I-x\|\,\|y\|+\|x\|\,\|s_I-y\|\\&<\epsilon.\end{aligned}$$

Thus $\sum\limits_{k\in K}a_k\overline{b_k}=\langle x,y\rangle$, as required.

2.31 Theorem: *Let H be a Hilbert space. Let $(u_k)_{k\in K}$ be an orthonormal indexed set in H, let $\mathcal{B}=\{u_k\,|\,k\in K\}$, let $U=\overline{\operatorname{Span}\mathcal{B}}$, and let $(c_k)_{k\in K}$ be an indexed set in $\mathbb{F}$.*

(1) If $\sum\limits_{k\in K}c_ku_k$ converges in H and $x=\sum\limits_{k\in K}c_ku_k$, then $x\in U$ and $c_k=\langle x,u_k\rangle$ for all $k\in K$.

(2) $\sum\limits_{k\in K}c_ku_k$ converges in H if and only if $\sum\limits_{k\in K}|c_k|^2$ converges in $\mathbb{R}$.

Proof: To prove Part 1, suppose the series $\sum\limits_{k\in K}c_ku_k$ converges in H and let $x=\sum\limits_{k\in K}c_ku_k$. Since the series converges in H, it is Cauchy. Note that for each $F\in\operatorname{Fin}(K)$, we have $s_F=\sum\limits_{k\in F}c_ku_k\in\operatorname{Span}\mathcal{B}\subseteq U$, and so the series $\sum\limits_{k\in K}c_ku_k$ is a Cauchy series in U. Since U is closed in H it is complete, so the Cauchy series $\sum\limits_{k\in K}c_ku_k$ converges in U, hence $x\in U$. Since $x\in U$, we know from Part 1 of the previous theorem that $x=\sum\limits_{k\in K}a_ku_k$ where $a_k=\langle x,u_k\rangle$. Since $\sum\limits_{k\in K}c_ku_k=x=\sum\limits_{k\in K}a_ku_k$, it follows from linearity that $\sum\limits_{k\in K}(c_k-x_k)u_k=0$. From Part 2 of the previous theorem, we have $\sum\limits_{k\in K}|c_k-a_k|^2=0$, and so we must have $c_k-a_k=0$ for all $k\in K$. This completes the proof of Part 1.

To prove Part 2, suppose first that $\sum\limits_{k\in K}c_ku_k$ converges, and let $x=\sum\limits_{k\in K}c_ku_k$. By Part 1, we have $x\in U$ and $c_k=\langle x,u_k\rangle$. By Part 2 of the previous theorem, we have $\sum\limits_{k\in K}|c_k|^2\le\|x\|^2$, so the series $\sum\limits_{k\in K}|c_k|^2$ converges in $\mathbb{R}$.

Suppose, conversely, that $\sum\limits_{k\in K}|c_k|^2$ converges in $\mathbb{R}$ and let $m=\sum\limits_{k\in K}|c_k|^2<\infty$. For $I\in\operatorname{Fin}(K)$, let $s_I=\sum\limits_{k\in I}c_ku_k$. Let $\epsilon>0$. Choose $F\in\operatorname{Fin}(K)$ such that $m-\epsilon^2<\sum\limits_{k\in F}|c_k|^2\le m$. For $I,J\in\operatorname{Fin}(K)$ with $I,J\supseteq F$, writing $I\triangle J=(I\cup J)\setminus(I\cap J)=(I\setminus J)\cup(J\setminus I)$, we have

$$\begin{aligned}\|s_I-s_J\|^2=\|s_{I\setminus J}-s_{J\setminus I}\|^2&=\sum_{k\in I\triangle J}|c_k|^2=\sum_{k\in I\cup J}|c_k|^2-\sum_{k\in I\cap J}|c_k|^2\\&\le\sum_{k\in I\cup J}|c_k|^2-\sum_{k\in F}|c_k|^2<m-(m-\epsilon^2)=\epsilon^2\end{aligned}$$

so that $\|s_I-s_J\|<\epsilon$. Thus the unordered series $\sum\limits_{k\in K}c_ku_k$ is Cauchy, so it converges in H.

2.32 Theorem: *(Bessel's Inequality) Let V be an inner product space. Let $(u_k)_{k\in K}$ be an orthonormal indexed set in V. For all $x\in V$, we have $\sum\limits_{k\in K}\big|\langle x,u_k\rangle\big|^2\le\|x\|^2$.*

Proof: Let $x\in V$. For $k\in K$, let $a_k=\langle x,u_k\rangle$. Let $F\in\operatorname{Fin}(K)$ and let $w_F=\sum\limits_{k\in F}a_ku_k$. Then

$$\begin{aligned}0\le\|x-w_F\|^2&=\|x\|^2-2\operatorname{Re}\langle x,w_F\rangle+\|w_F\|^2\\&=\|x\|^2-2\operatorname{Re}\Big(\sum_{k\in F}\overline{a_k}\langle x,u_k\rangle\Big)+\sum_{k,\ell\in F}a_k\overline{a_\ell}\langle u_k,u_\ell\rangle\\&=\|x\|^2-\sum_{k\in F}|a_k|^2.\end{aligned}$$

Since $\sum\limits_{k\in F}|a_k|^2\le\|x\|^2$ for every $F\in\operatorname{Fin}(K)$, it follows that $\sum\limits_{k\in K}|a_k|^2\le\|x\|^2$, as required.

2.33 Theorem: *(Orthogonal Projection) Let H be a Hilbert space, let $(u_k)_{k\in K}$ be an orthonormal indexed set in H, let $\mathcal{B}=\{u_k\,|\,k\in K\}$ and let $U=\overline{\operatorname{Span}\mathcal{B}}$. The orthogonal projection $P:H\to U$ is given by $Px=\sum\limits_{k\in K}a_ku_k$ where $a_k=\langle x,u_k\rangle$ and we have $\|P\|=1$.*

Proof: First note that when $x\in H$, by Bessel's Inequality we have $\sum\limits_{k\in K}|a_k|^2\le\|x\|^2$ so that $\sum\limits_{k\in K}|a_k|^2$ converges, hence by Part 2 of Theorem 2.31, the unordered series $\sum\limits_{k\in K}a_ku_k$ converges, and hence by Part 1 of Theorem 2.31, the sum $Px=\sum\limits_{k\in K}a_ku_k$ lies in U. Thus the map $P:H\to U$ given by $\sum\limits_{k\in K}a_ku_k$ is well-defined with $\operatorname{Range}(P)\subseteq U$.

To show that P is the orthogonal projection onto U, it suffices to show that when $x\in H$ and $u=Px$ and $v=x-u$, we have $v\in U^\perp$. Let $x\in H$, $u=Px=\sum\limits_{k\in K}a_ku_k\in U$ and $v=x-u$. Since $u=\sum\limits_{k\in K}a_ku_k$, by Part 1 of Theorem 2.31, for all $k\in K$ we have $\langle u,u_k\rangle=a_k=\langle x,u_k\rangle$, and hence $\langle v,u_k\rangle=\langle x-u,u_k\rangle=\langle x,u_k\rangle-\langle u,u_k\rangle=0$. Thus we have $v\in\operatorname{Span}\mathcal{B}^\perp=U^\perp$, as required. Thus P is the orthogonal projection onto U.

It remains to show that $\|P\|=1$. When $u\in U$ we have $Pu=u$ so that $\|Pu\|=\|u\|$ and it follows that $\|P\|\ge1$. For $x\in H$, if we let $u=Px=\sum\limits_{k\in K}a_ku_k$ then, as mentioned above, we have $\langle u,u_k\rangle=a_k=\langle x,u_k\rangle$, so by Part 2 of Theorem 2.30 and Bessel's Inequality, we have $\|Px\|^2=\|u\|^2=\sum\limits_{k\in K}|a_k|^2\le\|x\|^2$ so that $\|Px\|\le\|x\|$. Since $\|Px\|\le\|x\|$ for all $x\in H$, we have $\|P\|\le1$.

2.34 Theorem: *Let H be a Hilbert space and let $\mathcal{B}$ be an orthonormal set in H. Then $\mathcal{B}$ is a maximal orthonormal set if and only* $\operatorname{Span}\mathcal{B}$ *is dense H.*

Proof: Suppose $\mathcal{B}$ is not maximal. Choose an orthonormal set $\mathcal{C}$ in H with $\mathcal{B} \subsetneq \mathcal{C}$. Let $v \in \mathcal{C} \setminus \mathcal{B}$. Then $\|v\| = 1$ and $\langle v, u\rangle = 0$ for all $u \in \mathcal{B}$, hence $\langle v, u\rangle = 0$ for all $u \in \operatorname{Span}\mathcal{B}$. For all $u \in \operatorname{Span}\mathcal{B}$ we have $\|u - v\| = \|u\|^2 + \|v\|^2 = \|u\|^2 + 1 \geq 1$ so $v \notin \overline{\operatorname{Span}\mathcal{B}}$.

Suppose, conversely, that $\operatorname{Span}\mathcal{B}$ is not dense in H. Let $U = \overline{\operatorname{Span}\mathcal{B}} \neq H$, and recall (from Note 2.16) that $U^\perp = (\operatorname{Span}\mathcal{B})^\perp$. Since U is closed, by Theorem 2.20 we have $H = U \oplus U^\perp$. Since $U \neq H$ we have $U^\perp \neq \{0\}$. Choose $v \in U^\perp$ with $\|v\| = 1$. Then $\mathcal{B} \cup \{v\}$ is an orthonormal set which properly contains $\mathcal{B}$, so $\mathcal{B}$ is not maximal.

2.35 Theorem:

(1) Every inner product space contains a maximal orthonormal set.
(2) In a Hilbert space, any two maximal orthonormal sets have the same cardinality

Proof: To prove Part 1, let V be an inner product space. Let S be the set of all orthonormal sets in V, ordered by inclusion. If C is a chain in S (that is a totally ordered subset of S) then $\bigcup C$ is an upper bound for C in S. Since every chain in S has an upper bound, it follows from Zorn's Lemma that S has a maximal element.

To prove Part 2, let H be a Hilbert space and let $(u_k)_{k\in K}$ and $(v_\ell)_{\ell\in L}$ be two indexed orthonormal sets in H, and suppose that $\mathcal{B} = \{u_k | k \in K\}$ and $\mathcal{C} = \{v_\ell | \ell \in L\}$ are both maximal. If K or L is finite, then $\mathcal{B}$ and $\mathcal{C}$ are both Hamel bases for H and they have the same cardinality. Suppose K and L are infinite. For $k \in K$, let $L_k = \big\{\ell \in L \big| \langle u_k, v_\ell\rangle \neq 0\big\}$. Since for each $\ell \in L$ we have $\sum\limits_{k\in K} \big|\langle u_k, v_\ell\rangle\big| = \|v_\ell\|^2 = 1 > 0$, it follows that for each $\ell \in L$ there exists $k \in K$ such that $\langle u_k, v_\ell\rangle \neq 0$, so we have $L = \bigcup\limits_{k\in K} L_k$. Since for each $k \in K$ we have $\sum\limits_{\ell\in L} \big|\langle u_k, v_\ell\rangle\big|^2 = \|u_k\|^2 = 1 < \infty$, it follows from Theorem 2.26 that each set L_k is at most countable, that is $|L_k| \leq \aleph_0$. Thus, using some cardinal arithmetic, we have

$$|L| = \Big|\bigcup_{k\in K} L_k\Big| \leq \sum_{k\in K} |L_k| \leq \sum_{k\in K} \aleph_0 = |K| \cdot \aleph_0 = |K|.$$

A similar argument shows that $|K| \leq |L|$.

2.36 Definition: A **Hilbert basis** for a Hilbert space H is a maximal orthonormal set in H. The (Hilbert) **dimension** of a Hilbert space H, denoted by $\dim H$, is the cardinality of any Hilbert basis for H. We do not distinguish notationally between the Hamel dimension of H (that is the dimension of H as a vector space) and the Hilbert dimension of H (that is the dimension of H as a Hilbert space). Unless otherwise stated, when H is a Hilbert space, $\dim H$ will denote the Hilbert dimension.

2.37 Theorem: *Let H be a Hilbert space, let $(u_k)_{k\in K}$ be an orthonormal indexed set in H, and let $\mathcal{B} = \{u_k \,|\, k \in K\}$. Then the following are equivalent.*

(1) $\mathcal{B}$ is a Hilbert basis for H.
(2) For every $x \in H$ we have $x = \sum\limits_{k\in K} a_k u_k$, where $a_k = \langle x, u_k\rangle$.
(3) For every $x \in H$ we have $\|x\|^2 = \sum\limits_{k\in K} |a_k|^2$ where $a_k = \langle x, u_k\rangle$.
(4) For every $x, y \in H$ we have $\langle x, y\rangle = \sum\limits_{k\in K} a_k \overline{b_k}$ where $a_k = \langle x, u_k\rangle$ and $b_k = \langle y, u_k\rangle$.

Proof: The proof is left as an exercise.

2.38 Theorem: *Let H be a Hilbert space with Hilbert basis $\mathcal{B}$. Then H is separable if and only if $\mathcal{B}$ is at most countable.*

Proof: Suppose that $\mathcal{B}$ is uncountable. Let S be any dense subset of H. For each $u \in \mathcal{B}$ choose $s_u \in S$ with $\|s_u - u\| \leq \frac{\sqrt{2}}{4}$. For $u, v \in \mathcal{B}$ with $u \neq v$ we have $\|u\| = 1$ and $\|v\| = 1$ and $\langle u, v \rangle = 0$ so that $\|u - v\|^2 = \|u\|^2 + \|v\|^2 = 2$ and so

$$\|s_u - s_v\| = \big\|(s_u - u) + (u - v) + (v - s_v)\big\| \geq \|u - v\| - \big(\|s_u - u\| + \|s_v - v\|\big) = \sqrt{2} - \tfrac{\sqrt{2}}{4} - \tfrac{\sqrt{2}}{4} > 0$$

so that $s_u \neq s_v$. Thus $\mathcal{B}$ is at most countable.

Suppose, conversely, that $\mathcal{B} = \{u_1, u_2, \cdots\}$ is finite or countable. By Theorem 2.34, $\text{Span}_{\mathbb{F}}\mathcal{B}$ is dense in H. Note that $\text{Span}_{\mathbb{Q}}\mathcal{B}$ is dense in $\text{Span}_{\mathbb{R}}\mathcal{B}$ and $\text{Span}_{\mathbb{Q}[i]}\mathcal{B}$ is dense in $\text{Span}_{\mathbb{C}}\mathcal{B}$. Indeed given $c_1, \cdots, c_n \in \mathbb{F}$ (where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) we can choose $r_1, \cdots, r_n \in \mathbb{K}$ (where $\mathbb{K} = \mathbb{Q}$ or $\mathbb{Q}[i]$) such that $|r_k - c_k| < \frac{\epsilon}{n}$ and then

$$\begin{aligned}\Big\| \sum_{k=1}^{n} r_k u_k - \sum_{k=1}^{n} c_k u_k \Big\| &= \Big\| \sum_{k=1}^{n} (r_k - c_k) u_k \Big\| \leq \sum_{k=1}^{n} \big\|(r_k - c_k) u_k\big\| \\ &= \sum_{k=1}^{n} |r_k - c_k| \, \|u_k\| = \sum_{k=1}^{n} |r_k - c_k| < \epsilon.\end{aligned}$$

2.39 Exercise: For any nonempty set K and for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, let

$$\begin{aligned}\mathbb{F}^K &= \big\{(c_k)_{k \in K} \big| \text{each } c_k \in \mathbb{F}\big\}, \\ \ell_2(K, \mathbb{F}) &= \Big\{(c_k)_{k \in K} \in \mathbb{F}^K \,\Big|\, \sum_{k \in K} |c_k|^2 < \infty\Big\}.\end{aligned}$$

(a) For $a, b \in \ell_2(K, \mathbb{F})$, show that $\sum_{k \in K} a_k \overline{b_k}$ converges and let $\langle a, b \rangle = \sum_{k \in K} a_k \overline{b_k}$.
(b) Prove that this defines an inner product on $\ell_2(K, \mathbb{F})$.
(c) Prove that $\ell_2(K, \mathbb{F})$ is complete under this inner product.
(d) For each $\ell \in K$, let $e_\ell \in \ell_2(K, \mathbb{F})$ be given by $e_\ell = (e_{\ell,k})_{k \in K}$ with $e_{\ell,\ell} = 1$ and $e_{\ell,k} = 0$ when $k \neq \ell$. Prove that $(e_\ell)_{\ell \in K}$ is a Hilbert basis for $\ell_2(K, \mathbb{F})$.
(e) Prove that if H is a Hilbert space over $\mathbb{F}$ with $\dim H = |K|$ then $H \cong \ell_2(K, \mathbb{F})$.

2.40 Example: When $|K| = n \in \mathbf{Z}^+$ we have $\ell_2(K, \mathbb{F}) \cong \mathbb{F}^n$ (using the standard inner product). When $|K| = \aleph_0$ we have $\ell_2(K, \mathbb{F}) \cong \ell_2(\mathbb{F})$. For every separable Hilbert space H (over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) we have $H \cong \ell_2 = \ell_2(\mathbb{F})$. For example, we have $L_2[a, b] \cong \ell_2$.

The Dual Space and the Adjoint Map

2.41 Theorem: *(The Riesz Representation Theorem for Hilbert Spaces) Let H be a Hilbert space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. The map $\phi =: H \to H^*$ given by $\phi(u)(x)=\langle x,u\rangle$ is a bijective norm-preserving map which is linear when $\mathbb{F}=\mathbb{R}$ and conjugate-linear when $\mathbb{F}=\mathbb{C}$.*

Proof: For $u\in H$, write $\phi_u=\phi(u)$ so that $\phi_u(x)=\langle x,u\rangle$. Since $\phi_u(u)=\langle u,u\rangle=\|u\|^2$ it follows that $\|\phi_u\|\geq\|u\|$. Since for all $x\in H$ we have $\big|\phi_u(x)\big|=\big|\langle x,u\rangle\big|\leq\|x\|\,\|u\|$ it follows that $\|\phi_u\|\leq\|u\|$. Thus ϕ_u is a bounded linear map $\phi_u:H\to\mathbb{F}$ (that is $\phi_u\in H^*$) with $\|\phi_u\|=\|u\|$. Hence ϕ is a norm-preserving map $\phi:H\to H^*$. Note that ϕ is linear when $\mathbb{F}=\mathbb{R}$ and conjugate-linear when $\mathbb{F}=\mathbb{C}$. Since norm-preserving maps are injective, it remains to show that ϕ is surjective,

Let $f\in H^*$, that is let $f:H\to\mathbb{F}$ be a bounded linear map. If $f=0$ then we can take $u=0$ to get $\phi_u=f$. Suppose that $f\neq 0$. Let $U=\ker(f)$. Since f is linear, U is a subspace of H, and since f is bounded (hence continuous), U is closed, and it follows from Theorem 2.20 that $H=U\oplus U^\perp$. Since $f\neq 0$ it follows that $U\neq H$ so we have $U^\perp\neq\{0\}$. Choose $v\in U^\perp$ with $\|v\|=1$. Let $x\in H$. For $y=f(x)v-f(v)x$ we have $f(y)=f(x)f(v)-f(v)f(x)=0$ so that $y\in\ker(f)=U$. Since $y\in U$ and $v\in U^\perp$ we have $\langle y,v\rangle=0$, and so

$$\begin{aligned} f(x)=f(x)\|v\|^2=f(x)\langle v,v\rangle=\big\langle f(x)v,v\big\rangle=\big\langle y+f(v)x\,,\,v\big\rangle \\ =\big\langle f(v)x,v\big\rangle=f(v)\langle x,v\rangle=\big\langle x\,,\,\overline{f(v)}\,v\big\rangle. \end{aligned}$$

Thus we can choose $u=\overline{f(v)}\,v$ to get $\phi(u)=\phi_u=f$.

2.42 Definition: When H is a Hilbert space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, we use the bijection ϕ of the above theorem to define an inner product on H^*, as follows. Given $f,g\in H^*$ we let $u=\phi^{-1}(f)$ and $v=\phi^{-1}(g)$ (that is we let u and v be the elements in H such that $f(x)=\langle x,u\rangle$ and $g(x)=\langle x,v\rangle$) and then we define $\langle f,g\rangle=\langle v,u\rangle$ (note that the order of u and v is reversed so that the inner product is sesquilinear when $\mathbb{F}=\mathbb{C}$).

2.43 Definition: Recall that when U and V are vector spaces (over any field $\mathbb{F}$) and $F:U\to V$ is a linear map, we write $U^\#$ and $V^\#$ to denote the algebraic dual spaces, and we define the **dual** (or the **transpose** or the **algebraic adjoint**) of F to be the linear map $F^T:V^\#\to U^\#$ given by $F^T(g)=g\circ F$, that is by $F^T(g)(u)=g\big(F(u)\big)$ when $g\in V^\#$ and $u\in U$. In the case that U and V are normed linear spaces over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, and $F\in\mathcal{B}(U,V)$ (that is if $F:U\to V$ is a continuous linear map), F^T restricts to give a well-defined map $F^T:V^*\to U^*$ (because if $g\in V^\#$ is continuous then so is $g\circ F$).

When H and K are Hilbert spaces over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$ and $F:H\to K$ is a continuous linear map, we define the (Hilbert space) **adjoint** of F to be the liner map $F^*:K\to H$ given by $F^*=\phi^{-1}\circ F^T\circ\psi$ where $\phi:H\to H^*$ is the bijective map given by $\phi(u)(x)=\langle x,u\rangle$ and $\psi:K\to K^*$ is the bijective map given by $\psi(v)(y)=\langle y,v\rangle$. Equivalently, the adjoint of F is the map $F^*:K\to H$ such that $\phi\circ F^*=F^T\circ\psi$, that is the map such that

$$\phi\big(F^*(y)\big)=F^T\big(\psi(y)\big)=\psi(y)\circ F \text{ for all } y\in K,\text{ that is}$$
$$\phi\big(F^*y\big)(x)=\psi(y)(Fx)\text{ for all } x\in H, y\in K,\text{ that is}$$
$$\langle x,F^*y\rangle=\langle Fx,y\rangle\text{ for all } x\in H,y\in K.$$

2.44 Exercise: Show that when H and K are Hilbert spaces and $F:H\to K$ is a bounded linear map, we have $\|F^*\|=\|F^T\|=\|F\|$.

Weak Convergence

2.45 Definition: Let V be an inner product space over F, where $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, let (u_n) be a sequence in V and let $w\in V$. We say that (u_n) **converges weakly** to w in V, and we write $u_n\to w$ weakly in V, when $\langle u_n,x\rangle\to\langle w,x\rangle$ in F for all $x\in V$.

2.46 Note: When V is an inner product space and (u_n) is a sequence in V, it is easy to see that if $u_n\to w$ in V then we also have $u_n\to w$ weakly in V, but the converse is not always true. For example, when (u_n) is an orthonormal sequence in a Hilbert space H, verify that $u_n\to 0$ weakly in H (by Part 3 of Theorem 2.37), but $u_n\not\to 0$ in H.

2.47 Theorem: *Every bounded sequence in a Hilbert space has a weakly convergent subsequence.*

Proof: Let H be a Hilbert space. We claim that for every $a\in H$ and every bounded sequence $u=(u_n)$ in H, there is a subsequence (u_{n_ℓ}) of (u_n) such that the sequence $\big(\langle u_{n_\ell},a\rangle\big)$ converges in $\mathbb{F}$. Let $a\in H$, let $(u_n)_{n\geq 1}$ be a bounded sequence in H, and let $M=\sup\big\{\|u_n\|\,\big|\,n\in\mathbf{Z}^+\big\}$. Then for all $n\in\mathbf{Z}^+$ we have $\big|\langle u_n,a\rangle\big|\leq\|u_n\|\,\|a\|\leq M\|a\|$, and so the sequence $\big(\langle u_n,a\rangle\big)$ is bounded in $\mathbb{F}$. By the Bolzano-Weierstrass Theorem, we can choose a subsequence (u_{n_ℓ}) of (u_n) such that $\big(\langle u_n,a\rangle\big)$ converges in $\mathbb{F}$, as claimed.

Suppose that H is separable and let $S=\{a_1,a_2,\cdots\}\subseteq H$ be a countable dense subset. Let $u=(u_n)$ be a bounded sequence in H and let $M=\sup\big\{\|u_n\|\,\big|\,n\in\mathbf{Z}^+\big\}$. By the above claim, we can choose a subsequence (u_{n_ℓ}) of (u_n) such that $\lim\limits_{\ell\to\infty}\langle u_{n_\ell},a_1\rangle$ exists in $\mathbb{F}$, then we can choose a subsequence $(u_{n_{\ell_k}})$ of (u_{n_ℓ}) such that $\lim\limits_{k\to\infty}\langle u_{n_{\ell_k}},a_2\rangle$ exists in $\mathbb{F}$, then we can choose a subsequence $(u_{n_{\ell_{k_j}}})$ of $(u_{n_{\ell_k}})$ so that $\lim\limits_{j\to\infty}\langle u_{n_{\ell_{k_j}}},a_3\rangle$ exists in $\mathbb{F}$, and so on. Then the diagonal sequence $v=(v_1,v_2,v_3,\cdots)=\big(u_{n_1},u_{n_{\ell_2}},u_{n_{\ell_{k_3}}},\cdots\big)$ is then a subsequence of the original sequence (u_n) with the property that $\big(\langle v_k,a_m\rangle\big)$ converges in $\mathbb{F}$ for every $m\in\mathbf{Z}^+$, that is $\big(\langle v_k,a\rangle\big)$ converges for every $a\in S$.

Define $f:S\to\mathbb{F}$ by $f(a)=\lim\limits_{k\to\infty}\langle v_k,a\rangle$ for $a\in S$. Note that f is uniformly continuous on S because for $a,b\in S$ we have $\big|\langle v_k,a-b\rangle\big|\leq\|v_k\|\,\|a-b\|\leq M\|a-b\|$ for all k so that

$$\big|f(a)-f(b)\big|=\Big|\lim_{k\to\infty}\langle v_k,a\rangle-\lim_{k\to\infty}\langle v_k,b\rangle\Big|=\lim_{k\to\infty}\big|\langle v_k,a-b\rangle\big|\leq M\|a-b\|.$$

Since $f:S\to\mathbb{F}$ is uniformly continuous on S and S is dense in H, it follows that f extends (uniquely) to a continuous map $f:H\to\mathbb{F}$ defined by $f(x)=\lim\limits_{n\to\infty}f(a_n)$ where $x\in H$ and (a_n) is any sequence in S with $a_n\to x$ in H. Verify that this map f is linear and bounded (so we have $f\in H^*$) with $\|f\|\leq M$.

By The Riesz Representation Theorem, we can choose $w\in H$ such that $f(x)=\langle x,w\rangle$ for all $x\in H$. Verify that we have $\lim\limits_{k\to\infty}\langle v_k,x\rangle=\langle w,x\rangle$ for all $x\in H$, so (v_k) converges weakly to w in H. This completes the proof of the theorem in the case that H is separable.

Suppose that H is not separable and let (u_n) be a bounded sequence in H. Let $\mathcal{B}=\big\{e_k\big|k\in K\big\}$ be a Hilbert basis for H. For each $n\in\mathbf{Z}^+$, by Theorem 2.30 we have $u_n=\sum\limits_{k\in K}c_{n,k}e_k$ where $c_{n,k}=\langle u_n,e_k\rangle$ and we have $\sum\limits_{k\in K}|c_{n,k}|^2=\|u_k\|^2$. By theorem 2.26, for each $n\in\mathbf{Z}^+$ there are at most countably many indices $k\in K$ for which $c_{n,k}\neq 0$. Thus the set $L=\big\{k\in K\big|\exists n\in\mathbf{Z}^+\,c_{n,k}\neq 0\big\}$ is at most countable, and all of the elements u_n lie in the separable Hilbert space $U=\overline{\text{Span}\,\big\{e_\ell\big|\ell\in L\big\}}$. Since (u_n) is bounded, as proven above we can find a subsequence of (u_n) which converges weakly in U to an element $w\in U$. Verify that since $H=U\oplus U^\perp$, the subsequence also converges weakly in H to w.

2.48 Definition: Let H be a Hilbert space. A **compact operator** on H is a linear map $F : H \to H$ which sends weakly convergent sequences to convergent sequences, that is a linear map such that if $u_n \to w$ weakly in H then $Fu_n \to Fw$ in H.

2.49 Note: When H is a Hilbert space, every compact operator on H is continuous (because if $u_n \to w$ in H then $u_n \to w$ weakly in H) but the converse is not always true. For example, when H is an infinite-dimensional Hilbert space, the identity map $I : H \to H$ is continuous but not compact (since if (u_n) is an orthonormal sequence in H then $u_n \to 0$ weakly in H but $u_n \not\to 0$ in H).

2.50 Definition: Let H be a Hilbert space. A **self-adjoint** operator on H is a continuous linear map $F : H \to H$ such that $F^* = F$, that is such that $\langle Fx, y\rangle = \langle x, Fy\rangle$ for all $x, y \in H$.

2.51 Theorem: *Let H be a Hilbert space and let $F : H \to H$ be a continuous self-adjoint operator. Then*

(1) For every $u \in H$, we have $\langle Fu, u\rangle \in \mathbb{R}$. In particular, every eigenvalue of F is real.

(2) We have $\|F\| = \sup\left\{ \big|\langle Fu, u\rangle\big| \,\Big|\, u \in H, \|u\| = 1\right\}$. In particular, for every eigenvalue λ of F we have $|\lambda| \le \|F\|$.

Proof: To prove Part 1, note that since F is self-adjoint we have $\langle Fu, u\rangle = \langle u, F^*u\rangle = \langle u, Fu\rangle = \overline{\langle Fu, u\rangle}$, and so $\langle Fu, u\rangle \in \mathbb{R}$. In particular, when λ is an eigenvalue of F and $u \in H$ is a corresponding eigenvector with $\|u\| = 1$, we have $\lambda = \lambda\langle u, u\rangle = \langle \lambda u, u\rangle = \langle Fu, u\rangle \in \mathbb{R}$.

To prove Part 2, let $M = \sup\left\{ \big|\langle Fu, u\rangle\big| \,\Big|\, u \in H, \|u\| = 1\right\}$. Note that for all $u \in H$ with $\|u\| = 1$, we have $\big|\langle Fu, u\rangle\big| \le \|Fu\|\,\|u\| \le \|F\|\|u\| \cdot \|u\| = \|F\|$, and so $M \le \|F\|$.

To show that $\|F\| \le M$ we shall use a formula similar to the Polarization Identity. Verify (by expanding and cancelling) that for all $u, v \in H$ we have

$$\Big(\big\langle F(u{+}v), u{+}v\big\rangle - \big\langle F(u{-}v), u{-}v\big\rangle\Big) + i\Big(\big\langle F(u{+}iv), u{+}iv\big\rangle - \big\langle F(u{-}iv), u{-}iv\big\rangle\Big) = 4\langle Fu, v\rangle.$$

By Part 1, all of the inner products on the left are real, so if $\langle Fu, v\rangle \in \mathbb{R}$ then we have

$$\langle Fu, v\rangle = \tfrac{1}{4}\Big(\big\langle F(u{+}v), u{+}v\big\rangle - \big\langle F(u{-}v), u{-}v\big\rangle\Big).$$

Since $\big|\langle Fu, u\rangle\big| \le M$ for all $u \in H$ with $\|u\| = 1$, it follows that $\big|\langle Fw, w\rangle\big| \le M\|w\|^2$ for all $w \in H$ $\big($indeed when $w \ne 0$ we have $\big|\langle Fw, w\rangle\big| = \|w\|^2\big|\big\langle F\frac{w}{\|w\|}, \frac{w}{\|w\|}\big\rangle\big| \le \|w\|^2 M\big)$. Applying this fact with $w = u \pm v$ to the above displayed formula for $\langle Fu, v\rangle$, then using the Parallelogram Law, when $\langle Fu, v\rangle \in \mathbb{R}$ we have

$$\begin{aligned}\big|\langle Fu, v\rangle\big| &\le \tfrac{1}{4}\Big(\big|\big\langle F(u{+}v), u{+}v\big\rangle\big| + \big|\big\langle F(u{-}v), u{-}v\big\rangle\big|\Big)\\ &\le \tfrac{M}{4}\big(\|u+v\|^2 + \|u-v\|^2\big) = \tfrac{M}{2}\big(\|u\|^2 + \|v\|^2\big).\end{aligned}$$

In particular, if $\|u\| = \|v\| = 1$ and $\langle Fu, v\rangle \in \mathbb{R}$ then $\big|\langle Fu, v\rangle\big| \le M$. So for all $u \in H$ with $\|u\| = 1$, if $Fu = 0$ then $\|Fu\| \le M$ and if $Fu \ne 0$ then $\|Fu\| = \big|\big\langle Fu, \frac{Fu}{\|Fu\|}\big\rangle\big| \le M$. Thus we have $\|F\| \le M$, as required. Finally, note that when λ is an eigenvalue of F and u is a corresponding eigenvector with $\|u\| = 1$, we have $|\lambda| = \big|\lambda\langle u, u\rangle\big| = \big|\langle Fu, u\rangle\big| \le \|F\|$.

2.52 Example: The map $F : L_2[0, 1] \to L_2[0, 1]$ given by $F(f)(x) = xf(x)$ is a continuous self-adjoint map with no eigenvalues.

2.53 Theorem: *Let H be a Hilbert space and let $F : H \to H$ be a compact self-adjoint operator. Then F has an eigenvalue λ with $|\lambda| = \|F\|$.*

Proof: When $F = 0$, $\lambda = \|F\| = 0$ is an eigenvalue of F. Suppose $F \neq 0$. Since F is self-adjoint, we know that $\langle Fu, u\rangle \in \mathbb{R}$ for all $u \in H$ with $\|F\| = \sup\left\{ \left|\langle Fu, u\rangle\right| \,\middle|\, \|u\| = 1\right\}$. It follows that either $\|F\| = \lambda$ where $\lambda = \sup\left\{\langle Fu, u\rangle \middle| \|u\| = 1\right\} > 0$ or $\|F\| = -\lambda$ where $\lambda = \inf\left\{\langle Fu, u\rangle \middle| \|u\| = 1\right\} < 0$. Suppose the former (the proof in the latter case is similar). Since $\lambda = \sup\left\{\langle Fu, u\rangle \middle| \|u\|^2 = 1\right\}$, we can choose a sequence (u_n) in H with each $\|u_n\| = 1$ such that $\langle Fu_n, u_n\rangle \to \lambda$ in $\mathbb{R}$. Since (u_n) is bounded, we can choose a weakly convergent subsequence $(v_k) = (u_{n_k})$, say $v_k \to w$ weakly in H. Note that each $\|v_k\| = 1$, we have $\langle Fv_k, v_k\rangle \in \mathbb{R}$ for all k with $\langle Fv_k, v_k\rangle \to \lambda$ in $\mathbb{R}$, and $\lambda = \|F\|$, and so

$$\begin{aligned}\left\|Fv_k - \lambda v_k\right\|^2 &= \|Fv_k\|^2 - 2\mathrm{Re}\,\langle Fv_k, \lambda v_k\rangle + \|\lambda v_k\|^2 \\ &= \|Fv_k\|^2 - 2\lambda\langle Fv_k, v_k\rangle + \lambda^2 \\ &\leq \|F\|^2 - 2\lambda\langle Fv_k, v_k\rangle + \lambda^2 \longrightarrow \|F\|^2 - \lambda^2 = 0.\end{aligned}$$

Since $v_k \to w$ weakly in H and F is compact, we have $Fv_k \to Fw$ in H, and hence

$$\lambda v_k = (\lambda v_k - Fv_k) + Fv_k \longrightarrow 0 + Fw = Fw.$$

Since F is continuous we have

$$F(Fw) = F\big(\lim_{k\to\infty} \lambda v_k\big) = \lambda \lim_{k\to\infty} Fv_k = \lambda\, Fw$$

and so λ is an eigenvalue of F with eigenvector Fw.

2.54 Note: Let H be a Hilbert space. We use the following remarks in the next theorem.

(1) When $F : H \to H$ is a continuous linear operator and λ is an eigenvalue of F, the eigenspace E_λ is closed because $\{0\}$ is closed in H and $E_\lambda = G^{-1}(\{0\})$ where $G = F - \lambda I$, which is continuous.

(2) When $F : H \to H$ is a continuous self-adjoint operator and λ and μ are distinct eigenvalues of F, the eigenspaces E_λ and E_μ are orthogonal. Indeed, if $\lambda, \mu \in \mathbb{R}$ with $\lambda \neq \mu$ and $u \in E_\lambda$ and $v \in E_\mu$, then $\lambda\langle u, v\rangle = \langle \lambda u, v\rangle = \langle Fu, v\rangle = \langle u, Fv\rangle = \langle u, \mu v\rangle = \mu\langle u, v\rangle$ hence $\langle u, v\rangle = 0$.

(3) When $U \subseteq H$ is a closed subspace, the orthogonal projection P onto U is self-adjoint. Indeed, given $x, y \in H$, write $x = u + v$ and $y = r + s$ with $u, r \in U$ and $v, s \in U^\perp$ and then $\langle Px, y\rangle = \langle u, r + s\rangle = \langle u, r\rangle - \langle u + v, r\rangle = \langle x, Py\rangle$.

(4) When $F, G : H \to H$ are self-adjoint, so is $F + c\,G$ where $c \in \mathbb{R}$, because for all $x, y \in H$ we have $\langle (F + c\,G)x, y\rangle = \langle Fx, y\rangle + c\langle Gx, y\rangle = \langle x, Fy\rangle + c\langle x, Gy\rangle = \langle x, (F + c\,G)y\rangle$.

(5) When $U \subseteq H$ is a finite-dimensional subspace, the orthogonal projection P onto U is compact. Indeed, suppose $w_n \to w$ weakly in H. Write $w_n = u_n + v_n$ and $w = u + v$ with $u_n, u \in U$ and $v_n, v \in U^\perp$. For all $x \in U$, we have $\langle u_n, x\rangle = \langle w_n, x\rangle \to \langle w, x\rangle = \langle u, x\rangle$ so $u_n \to u$ weakly in U. Since U is finite-dimensional, we can choose an orthonormal basis $\{e_1, \cdots, e_n\}$ for U and then $u_n = \sum_{k=1}^n \langle u_n, e_k\rangle e_k \to \sum_{k=1}^n \langle u, e_k\rangle e_k = u$ in U. Thus $Pw_n = u_n \to u = Pw$, so P is compact.

(6) When $F, G : H \to H$ are compact, so is $F + c\,G$ where $c \in \mathbb{F}$, because if (w_n) converges weakly in H then (Fw_n) and (Gw_n), hence also $\big((F + c\,G)w_n\big)$, converge in H.

(7) When $F : H \to H$ is a continuous self-adjoint operator, λ is a nonzero eigenvalue of F, and P is the orthogonal projection onto the eigenspace E_λ, we have $\lambda P = FP = PF$ because for all $x \in H$ we have $Px \in E_\lambda$ so that $FPx = \lambda Px$, and for all $x, y \in H$ we have $\langle PFx, y\rangle = \langle Fx, Py\rangle = \langle x, FPy\rangle = \langle x, \lambda Py\rangle = \langle \lambda Px, y\rangle$.

2.55 Theorem: *(The Spectral Theorem for Compact Self-Adjoint Operators) Let H be a Hilbert space and let $F : H \to H$ be a nonzero compact self-adjoint operator on H. Then the set of nonzero eigenvalues of H is at most countable, and the eigenspace of each nonzero eigenvalue is finite-dimensional. When H has finitely many nonzero eigenvalues, say $\lambda_1, \cdots, \lambda_n$, we have $F = \lambda_1 P_{\lambda_1} + \cdots + \lambda_n P_{\lambda_n}$ where P_{λ_k} is the orthogonal projection onto the eigenspace E_{λ_k}. When H has countably many nonzero eigenvalues, they can be arranged into a sequence $\lambda_1, \lambda_2, \lambda_3, \cdots$ in nonincreasing order of absolute value with $\lambda_n \to 0$, and in the space of bounded linear operators on H, we have*

$$F = \sum_{k=1}^{\infty} \lambda_k P_{\lambda_k}$$

where P_{λ_k} is the orthogonal projection onto the eigenspace E_{λ_k}.

Proof: First we note that because F is compact, it follows that the eigenspace E_λ of any nonzero eigenvalue $\lambda \neq 0$ must be finite dimensional, because if E_λ was infinite dimensional we could choose an orthonormal sequence $(e_n)_{n\geq 1}$ in E_λ, but this is not possible because we would have $e_n \to 0$ weakly in H but $Fe_n = \lambda e_n \not\to 0$ in H.

Using Theorem 2.52, choose an eigenvalue λ of F with $|\lambda| = \|F\|$ and note that $\lambda \neq 0$. Since F is continuous, the eigenspace $E_\lambda = E_\lambda(F)$ is closed. Let P be the orthogonal projection onto E_λ. and note that P is compact and self-adjoint and we have $FP = PF = \lambda P$. Let

$$G = F - \lambda P$$

and note that G is also compact and self-adjoint.

We claim that λ is not an eigenvalue of G. Let $u \in H$ with $Gu = \lambda u$, that is $\lambda u = Gu = Fu - \lambda Pu$. Apply P on both sides, using $PF = \lambda P$ and $P^2 = P$ to get $\lambda Pu = P(Fu - \lambda Pu) = \lambda Pu - \lambda Pu = 0$, and hence $Pu = 0$. Since $Pu = 0$ and P is the orthogonal projection onto E_λ, we have $u \in E_\lambda^\perp$. Since $u \in E_\lambda^\perp$ and $u \in E_\lambda$, we have $u = 0$. Thus λ is not an eigenvalue of G, as claimed.

We claim that every non-zero eigenvalue μ of G is also an eigenvalue of F, and that $E_\mu(G) = E_\mu(F)$ (that is, the eigenspace of μ for G is equal to the eigenspace of μ for F). Let $0 \neq \mu$ be an eigenvalue of G and let w be an eigenvector of μ for G, so we have $Gw = \mu w$. Note that since $\lambda P = FP = PF$ we have $G = F - \lambda P = F(I - P) = (I - P)F$, and since $P^2 = P$ we have $(I-P)^2 = (I-2P+P^2) = (I-P)$. Thus we have $\mu w = Gw = (I-P)Fw$ and hence $(I-P)\mu w = (I-P)^2 Fw = (I-P)Fw = Gw = \mu w$. Since $\mu \neq 0$ we can divide both sides by μ to get $(I-P)w = w$, and so $Fw = F(I-P)w = Gw = \mu w$. Thus μ is also an eigenvalue of F with w as an eigenvector, so we have $E_\mu(G) \subseteq E_\mu(F)$.

It remains to show that $E_\mu(F) \subseteq E_\mu(G)$. Let $v \in E_\mu(F)$, so we have $Fv = \mu v$. Since μ is an eigenvalue of G but λ is not, we have $\mu \neq \lambda$ so that the eigenspaces $E_\mu(F)$ and $E_\lambda(F)$ are orthogonal, and hence $Pv = 0$. Thus $Gv = (F - \lambda P)v = Fv = \mu v$ and hence $E_\mu(F) \subseteq E_\mu(G)$, as required.

Let $F_1 = F$, $\lambda_1 = \lambda$ and $F_2 = G = F_1 - \lambda_1 P_{\lambda_1}$, then repeat the above procedure by choosing an eigenvalue λ_2 of F_2 with $|\lambda_2| = \|F_2\|$, and letting $F_3 = F_2 - \lambda_2 P_{\lambda_2}$, and so on, to obtain a sequence of eigenvalues $\lambda_1, \lambda_2, \cdots$ and maps $F_{n+1} = F - \sum_{k=1}^{n} \lambda_k P_{\lambda_k}$ where at each stage, λ_n is an eigenvalue for F_n with $|\lambda_n| = \|F_n\|$, and $E_{\lambda_n}(F_n) = E_{\lambda_n}(F)$. Note that the eigenvalues are distinct (because λ_{n-1} is an eigenvalue for F_{n-1} but not for F_n) and they are in nonincreasing order of absolute value $\big($because λ_n is an eigenvalue of F_{n-1} so that $|\lambda_n| \leq \|F_{n-1}\| = |\lambda_{n-1}|\big)$.

Either the procedure comes to an end after finitely many steps with $F_{n+1} = 0$, in which case we have $F = \sum_{k=1}^{n} \lambda_k P_{\lambda_k}$, or it continues indefinitely to give an infinite (countable) sequence of distinct eigenvalues in nonincreasing order of absolute value. Suppose that the procedure continues indefinitely, so we obtain an infinite sequence $\lambda_1, \lambda_2, \lambda_3, \cdots$ of distinct eigenvalues of F with $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots$.

We claim that $|\lambda_n| \to 0$. Suppose, for a contradiction, that $|\lambda_n| \to r > 0$. For each $n \in \mathbf{Z}^+$, choose an eigenvector $u_n \in E_{\lambda_n}(F)$ with $\|u_n\| = 1$. Since (u_n) is bounded, we can choose a weakly convergent subsequence (u_{n_k}). Since F is compact, the sequence (Fu_{n_k}) converges in H. But this is not possible because (since the eigenspaces are orthogonal) we have

$$\|Fu_{n_k} - Fu_{n_\ell}\|^2 = \left\|\lambda_{n_k} u_{n_k} - \lambda_{n_\ell} u_{n_\ell}\right\|^2 = \lambda_{n_k}^2 + \lambda_{n_\ell}^2 \geq 2r^2$$

so the sequence (Fu_{n_k}) is not Cauchy. Thus $|\lambda_n| \to 0$, as claimed.

Note that since $|\lambda_n| \to 0$, it follows that $F = \sum_{k=1}^{\infty} \lambda_k P_{\lambda_k}$ (in the space of bounded linear operators on H, using the operator norm) because

$$\left\|F - \sum_{k=1}^{n} \lambda_k P_{\lambda_k}\right\| = \|F_{n+1}\| = |\lambda_{n+1}| \to 0.$$

It remains to show that the eigenvalues $\lambda_1, \lambda_2, \cdots$ constitute all of the nonzero eigenvalues of F. Let us consider the case that our procedure yields infinitely many eigenvalues $\lambda_1, \lambda_2, \cdots$ and that $F = \sum_{k=1}^{\infty} \lambda_k P_{\lambda_k}$ (the case of finitely many eigenvalues is simpler). Each eigenspace E_{λ_k} is finite-dimensional and can be given an orthonormal basis. These bases can be combined to give a countable orthonormal set (or an orthonormal sequence). This orthonormal set is a Hamel basis for the space of sums $\sum_{k=1}^{\infty} u_k$ where each $u_k \in E_{\lambda_k}$ with only finitely many of the terms u_k non-zero. Let U be the closure of this space in H. By Theorems 2.30 and 2.31, U is the space of sums $\sum_{k=1}^{\infty} u_k$ in H with $\sum_{k=1}^{\infty} \|u_k\|^2 < \infty$, where each $u_k \in E_{\lambda_k}$, and the elements u_k are uniquely determined. Since U is closed in H, we have $H = U \oplus U^\perp$, and so every element $w \in H$ can be written uniquely in the form $w = v + \sum_{k=1}^{\infty} u_k$ with $v \in U^\perp$ and $u_k \in E_{\lambda_k}$, and then we have $u_k = P_{\lambda_k} w$.

Let $0 \neq \mu \in \mathbb{R}$ with $\mu \neq \lambda_k$ for any k, let $w \in H$, and suppose that $Fw = \mu w$. Write $w = v + \sum_{k=1}^{\infty} u_k$ with $v \in U^\perp$ and $u_k \in E_{\lambda_k}$. Then $Fw = \sum_{k=1}^{\infty} \lambda_k u_k$ and $\mu w = \mu v + \sum_{k=1}^{\infty} \mu u_k$, so that $0 = \mu w - Fw = \mu v + \sum_{k=1}^{\infty} (\mu - \lambda_k) u_k$, hence $\mu v = 0$ and $(\mu - \lambda_k) u_k = 0$ for all k. Since since $\mu \neq 0$ we have $v = 0$ and since $\mu \neq \lambda_k$ we have $u_k = 0$. Thus $w = 0$ so that μ is not an eigenvalue of F.