Chapter 4. Banach and Hilbert Spaces

4.1 Definition: Let W be an inner product space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. For a subset $\mathcal{A}\subseteq W$, we say that $\mathcal{A}$ is **orthogonal** when $\langle u,v\rangle=0$ for all $u,v\in\mathcal{A}$ with $u\neq v$, and we say that $\mathcal{A}$ is **orthonormal** when $\mathcal{A}$ is orthogonal with $\|u\|=1$ for every $u\in\mathcal{A}$.

4.2 Theorem: *Let W be an inner product space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}\subseteq W$.*

(1) If $\mathcal{A}$ is an orthogonal set of nonzero vectors then for $x\in\operatorname{Span}\mathcal{A}$ with say $x=\sum\limits_{k=1}^{n}c_ku_k$ where $c_k\in\mathbb{F}$ and $u_k\in\mathcal{A}$, we have $c_k=\langle x,u_k\rangle/\|u_k\|^2$ for all indices k, and in particular, $\mathcal{A}$ is linearly independent.

(2) If $\mathcal{A}$ is orthonormal then for $x\in\operatorname{Span}\mathcal{A}$ with say $x=\sum\limits_{k=1}^{n}c_ku_k$ where $c_k\in\mathbb{F}$ and $u_k\in\mathcal{A}$, we have $c_k=\langle x,u_k\rangle$ for all k, and in particular, $\mathcal{A}$ is linearly independent.

Proof: To prove Part (1), suppose that $\mathcal{A}$ is an orthogonal set of nonzero vectors and let $x=\sum\limits_{j=1}^{n}c_ju_j$ with each $c_j\in\mathbb{F}$ and each $u_j\in\mathcal{A}$. Then for all indices k, since $\langle u_j,u_k\rangle=0$ whenever $j\neq k$ we have $\langle x,u_k\rangle=\left\langle\sum\limits_{j=1}^{n}c_ju_j\,,\,u_k\right\rangle=\sum\limits_{j=1}^{n}c_j\langle u_j,u_k\rangle=c_k\langle u_k,u_k\rangle=c_k\|u_k\|^2$ and so $c_k=\frac{\langle x,u_k\rangle}{\|u_k\|^2}$, as required. In particular, when $x=0$ we find that $c_k=0$ for all k, and this shows that $\mathcal{A}$ is linearly independent. This proves Part (1), and Part (2) follows immediately from Part (1).

4.3 Theorem: *(The Gram-Schmidt Procedure) Let W be a finite or countable dimensional inner product space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}=\{u_1,u_2,\cdots\}$ be an ordered basis for W. Let $v_1=u_1$ and for $n\geq 2$ let $v_n=u_n-\sum\limits_{k=1}^{n-1}\dfrac{\langle u_n,v_k\rangle}{\|v_k\|^2}\,v_k$. Then the set $\mathcal{B}=\{v_1,v_2,\cdots\}$ is an orthogonal basis for W 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 prove, by induction on n, that $\{v_1,v_2,\cdots,v_n\}$ is an orthogonal basis for $\operatorname{Span}\{u_1,u_2,\cdots,u_n\}$. When $n=1$ this is clear since $v_1=u_1$. Let $n\geq 2$ and suppose, inductively, that $\{v_1,\cdots,v_{n-1}\}$ is an orthogonal basis for $\operatorname{Span}\{u_1,\cdots,u_{n-1}\}$. Since $v_n=u_n-\sum_{k=1}^{n-1}\frac{\langle u_n,v_k\rangle}{\|v_k\|^2}\,v_k$, we see that u_n is equal to v_n plus a linear combination of the vectors $v_1,\cdots,v_{n-1}$, and so we have $\operatorname{Span}\{v_1,\cdots,v_{n-1},v_n\}=\operatorname{Span}\{v_1,\cdots,v_{n-1},u_n\}$. By the induction hypothesis, we have $\operatorname{Span}\{v_1,\cdots,v_{n-1}\}=\operatorname{Span}\{u_1,\cdots,u_{n-1}\}$ so we have

$$\operatorname{Span}\{v_1,\cdots,v_{n-1},v_n\}=\operatorname{Span}\{v_1,\cdots,v_{n-1},u_n\}=\operatorname{Span}\{u_1,\cdots,u_{n-1},u_n\}.$$

It remains to show that the set $\{v_1,v_2,\cdots,v_n\}$ is an orthogonal set. By the induction hypothesis, we have $\langle v_j,v_k\rangle=0$ for all $1\leq j,k<n$, so it suffices to show that $\langle v_n,v_k\rangle=0$ for all indices $1\leq k<n$ and indeed, for $1\leq k<n$ we have

$$\begin{aligned}\langle v_n,v_k\rangle&=\left\langle u_n-\sum_{j=1}^{n-1}\tfrac{\langle u_n,v_j\rangle}{\|v_j\|^2}\,v_j\;,\;v_k\right\rangle=\langle u_n,v_k\rangle-\sum_{j=1}^{n-1}\tfrac{\langle u_n,v_j\rangle}{\|v_j\|^2}\,\langle v_j,v_k\rangle\\&=\langle u_n,v_k\rangle-\frac{\langle u_n,v_k\rangle}{\|v_k\|^2}\,\langle v_k,v_k\rangle=0.\end{aligned}$$

4.4 Corollary: *Every finite or countable dimensional inner product space W over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ has an orthonormal basis.*

Proof: The proof is left as an exercise.

4.5 Remark: It is not the case that every uncountable dimensional inner product space has an orthonormal basis. For example, we shall see below that an infinite dimensional separable Hilbert space does not have an orthonormal basis.

4.6 Corollary: *Let W be a finite or countable dimensional inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $U \subseteq W$ be a finite dimensional subspace. Then every orthogonal (or orthonormal) basis $\mathcal{A}$ for U extends to an orthogonal (or orthonormal) basis for W.*

Proof: The proof is left as an exercise.

4.7 Remark: The above corollary does not hold in general in the case that the subspace U is countable dimensional, as we shall soon see in Example 4.12.

4.8 Corollary: *Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let U and V be finite or countable dimensional inner product spaces over $\mathbb{F}$. 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 $\mathbb{F}^\infty$.*

Proof: The proof is left as an exercise.

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

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

4.10 Theorem: *Let W be an inner product space over $F = \mathbb{R}$ or $\mathbb{C}$. Let $U \subseteq W$ be a subspace. Then*
(1) $U^\perp$ is a subspace of W,
(2) if $\mathcal{A}$ is a basis for U then $U^\perp = \big\{ x \in W \big| \langle x, u \rangle = 0$ for all $u \in U \big\}$,
(3) $U \cap U^\perp = \{0\}$, and
(4) $U \subseteq (U^\perp)^\perp$.
(5) if U is finite dimensional then $U \oplus U^\perp = W$, and
(6) if $U \oplus U^\perp = W$ then $U = (U^\perp)^\perp$.

Proof: We leave the proofs of Parts (1) to (4) as an exercise. To prove Part (5), suppose that U is finite-dimensional. Let $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ be an orthonormal basis for $\mathcal{A}$. We need to show that for every $x \in W$ there exist unique vectors $u, v \in W$ with $u \in U$, $v \in U^\perp$ and $u + v = x$. First we prove uniqueness. Let $x \in W$, and suppose that $u \in U$, $v \in U^\perp$ and $u + v = x$. Note that for all indices k we have

$$\langle x, u_k \rangle = \langle u + v, u_k \rangle = \langle u, u_k \rangle + \langle v, u_k \rangle = \langle u, u_k \rangle.$$

and so, by Theorem 4.2, we have

$$u = \sum_{k=1}^{n} \langle u, u_k \rangle u_k = \sum_{k=1}^{n} \langle x, u_k \rangle u_k.$$

This proves uniqueness, since given $x \in W$, the vector u must be given by $u = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$ and then the vector v must be given by $v = x - u$.

To prove existence, let $x \in W$ and choose u and v to be the vectors $u = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$ and $v = x - u$. Then we have $u \in U$ and $u + v = x$, so it suffices to show that $v \in U^\perp$. For all indices k we have

$$\begin{aligned} \langle v, u_k \rangle &= \langle x - u, u_k \rangle = \langle x, u_k \rangle - \langle u, u_k \rangle = \langle x, u_k \rangle - \Big\langle \sum_{j=1}^{n} \langle x, u_j \rangle u_j \, , \, u_k \Big\rangle \\ &= \langle x, u_k \rangle - \sum_{j=1}^{n} \langle x, u_j \rangle \langle u_j, u_k \rangle = \langle x, u_k \rangle - \sum_{j=1}^{n} \langle x, u_j \rangle \delta_{j,k} = \langle x, u_k \rangle - \langle x, u_k \rangle = 0. \end{aligned}$$

Since $\langle v, u_k \rangle = 0$ for all $1 \le k \le n$, from Part (2) we have $v \in U^\perp$. This proves Part (5).

To prove Part (6). Suppose that $U \subseteq (U^\perp)^\perp$. We know, from Part 4, that $U \subseteq (U^\perp)^\perp$. Let $x \in (U^\perp)^\perp$. Choose $u, v \in W$ with $u \in U$, $v \in U^\perp$ such that $u + v = x$. Since $u \in U \subseteq (U^\perp)^\perp$ and $x \in (U^\perp)^\perp$, we have $v = x - u \in (U^\perp)^\perp$. Since $v \in U^\perp \cap (U^\perp)^\perp = \{0\}$, we have $v = 0$ so that $x = u + v = u \in U$.

4.11 Remark: Parts (5) and (6) of the above theorem do not always hold when U is infinite dimensional, as the following example shows.

4.12 Example: Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $W = \mathbb{F}^\infty$ and let $U = \big\{ a \in \mathbb{F}^\infty \big| \sum_{k=1}^{\infty} a_k = 0 \big\}$. Note that W is a countable-dimensional inner product space with standard basis $\{e_1, e_2, e_3, \cdots\}$ and U is a countable-dimensional proper subspace of W with basis $\mathcal{A} = \{u_1, u_2, u_3, \cdots\}$ where $u_k = e_1 - e_{k+1} = (1, 0, \cdots, 0, -1, 0, \cdots)$. We have

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

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 $U \subsetneqq (U^T)^T = W$ and $U \oplus U^\perp = U \oplus \{0\} = U \subsetneqq W$. Also notice that, although we could apply the Gram-Schmidt Procedure to the basis $\mathcal{A}$ to obtain an orthogonal basis $\mathcal{B} = \{v_1, v_2, \cdots\}$ for U, the basis $\mathcal{B}$ cannot be extended to an orthogonal basis for W because there is no nonzero vector $0 \neq x \in W$ with $\langle x, v_k \rangle$ for all k.

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

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

Since $U = (U^\perp)^\perp$, for u and v as above we have $\text{Proj}_{U^\perp}(x) = v$. When $y \in W$ and $U = \text{Span}\{y\}$, we also write $\text{Proj}_y(x) = \text{Proj}_U(x)$.

4.14 Note: Let W be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let U be a finite dimensional subspace of W. Let $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ be an orthogonal basis for U. Then for $x \in W$, as in the proof of Part (5) of Theorem 4.15, we see that

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

4.15 Example: As an exercise, show 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 that for any such y we have $\mathrm{Proj}_U(x) = Ay$. In particular, when $\mathrm{rank}(A) = m$ show that A^*A is invertible so that $\mathrm{Proj}_U(x) = A(A^*A)^{-1}A^*x$.

4.16 Theorem: *Let W be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $U \subseteq W$ be a subspace of W such that $W = U \oplus U^\perp$. Let $x \in W$. Then $\mathrm{Proj}_U(x)$ is the unique point in U which is nearest to x.*

Proof: Let $u, v \in W$ be the vectors with $u \in U$, $v \in V$ and $u + v = x$, so that we have $\mathrm{Proj}_U(x) = u$. Let $w \in U$ with $w \neq u$. Since $\langle w-u, x-u\rangle = \langle w-u, v\rangle = \langle w, v\rangle - \langle u, v\rangle = 0$, Pythagoras' Theorem gives

$$\|x - w\|^2 = \|(x - u) - (w - u)\|^2 = \|x - u\|^2 + \|w - u\|^2 > \|x - u\|^2$$

and so $\|x - w\| > \|x - u\|$.

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

4.18 Theorem: *Let H be a Hilbert space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. 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 $d = \mathrm{dist}(a, S) = \inf\big\{\|x - a\| \,\big|\, x \in S\big\}$. 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 $m \in \mathbb{Z}^+$ 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

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

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 $n \in \mathbb{Z}^+$ 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

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

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

4.19 Theorem: *Let H be a Hilbert space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Let $U\subseteq H$ be a closed subspace. Then we have $H=U\oplus U^\perp$. This means that for all $x\in H$ there exist unique points $u\in U$ and $v\in U^\perp$ such that $u+v=x$. In this case, the point u is the unique point in U nearest to x.*

Proof: 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

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

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

It remains to show 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

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

so that $\|x-u'\|>\|x-u\|$. Thus u is the point in U which is nearest to x, as required.

4.20 Theorem: *Every inner product space contains a maximal orthonormal set.*

Proof: Let W be an inner product space. Let S be the set of all orthonormal sets in W, 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.

4.21 Theorem: *Let H be a Hilbert space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}$ be an orthonormal set in H and let $U=\mathrm{Span}_F\,\mathcal{A}$. Then $\mathcal{A}$ is maximal if and only if U is dense in H.*

Proof: If $\mathcal{A}$ is not maximal then we can choose $v\in U^\perp$ with $\|v\|=1$ (so that $\mathcal{A}\cup\{v\}$ is orthonormal) and then for all $u\in U$, since $\langle v,u\rangle=0$, we have $\|u-v\|^2=\|u\|^2+\|v\|^2\geq\|v\|^2=1$. Thus U is not dense in H, indeed there is no $u\in U$ with $\|u-v\|\leq\frac{1}{2}$.

Suppose, conversely, that U is not dense in H, that is $\overline{U}\neq H$. Note that $\overline{U}$ is a vector space, indeed given $a,b\in\overline{U}$ we can choose $\{x_n\}$ and $\{y_n\}$ with $x_n\to a$ and $y_n\to b$ in H and then $(x_n+y_n)\to(a+b)$ so that $a+b\in\overline{U}$, and for $c\in\mathbb{F}$ we have $cx_n\to ca$ so that $ca\in\overline{U}$. By the above theorem, we have $H=\overline{U}\oplus\overline{U}^\perp$. Since $H\neq\overline{U}$ we must have $\overline{U}^\perp\neq\{0\}$. Choose $v\in\overline{U}^\perp$ with $\|v\|=1$. Since $\langle v,u\rangle=0$ for all $u\in\overline{U}$ we certainly have $\langle v,u\rangle=0$ for all $u\in U$, so the set $\mathcal{A}\cup\{v\}$ is orthonormal. And we cannot have $v\in U$ since $U\cap U^\perp=\{0\}$, and so $\mathcal{A}\subsetneqq\mathcal{A}\cup\{v\}$ so that $\mathcal{A}$ is not maximal.

4.22 Theorem: *Let H be a Hilbert space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{A}$ be a maximal orthonormal set in H. Then H is separable if and only if $\mathcal{A}$ is at most countable.*

Proof: Suppose that $\mathcal{A}$ is uncountable. Let S be any dense subset of H. For each $u\in\mathcal{A}$ choose $s_u\in S$ with $\|s_u-u\|\leq\frac{\sqrt{2}}{4}$. For $u,v\in\mathcal{A}$ 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 S is uncountable so H is not separable.

Suppose, conversely, that $\mathcal{A}=\{u_1,u_2,\cdots\}$ is finite or countable. By the above theorem, $U=\text{Span}_F\,\mathcal{A}$ is dense in H. Note that $\text{Span}_{\mathbb{Q}}\mathcal{A}$ is dense in $\text{Span}_{\mathbb{R}}\mathcal{A}$ and $\text{Span}_{\mathbb{Q}[i]}\mathcal{A}$ is dense in $\text{Span}_{\mathbb{C}}\mathcal{A}$. Indeed given $c_1,c_2,\cdots,c_n\in\mathbb{F}$ (where $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$) we can choose $r_1,r_2,\cdots,r_n\in\mathbb{P}$ (where $\mathbb{P}=\mathbb{Q}$ or $\mathbb{Q}[i]$) such that $|r_k-c_k|<\frac{\epsilon}{n}$ and then

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

4.23 Theorem: *Let H be a separable Hilbert space over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, let $\mathcal{A}=\{u_1,u_2,\cdots\}$ be a countable orthonormal set in H, and let $U=\text{Span}_F\,\mathcal{A}$. Then the following are equivalent.*

(1) $\mathcal{A}$ is maximal.
(2) U is dense in H.
(3) For every $x\in H$ we have $x=\sum_{k=1}^\infty\langle x,u_k\rangle u_k$ in H.
(4) For every $x\in H$ we have $\|x\|^2=\sum_{k=1}^\infty\big|\langle x,u_k\rangle\big|^2$ in $\mathbb{R}$.
(5) For all $x,y\in H$ we have $\langle x,y\rangle=\sum_{k=1}^\infty\langle x,u_k\rangle\overline{\langle y,u_k\rangle}$.

Proof: The equivalence of Parts (1) and (2) follows from Theorem 4.21. Let us prove that (2) implies (3). Suppose that U is dense in H. Let $x\in H$. For each $n\in\mathbb{Z}^+$, let $U_n=\text{Span}\{u_1,u_2,\cdots,u_n\}$ and let $w_n=\text{Proj}_{U_n}(x)=\sum_{k=1}^n\langle x,u_k\rangle u_k$. Let $\epsilon>0$. Since U is dense in H we can choose $u\in U$ with $\|u-x\|<\epsilon$. Say $u=\sum_{k=1}^m c_ku_k$. For all $n\geq m$, since $u\in U_n$ and w_n is the point in U_n nearest to x we have $\|w_n-x\|\leq\|u-x\|<\epsilon$. Thus $\lim_{n\to\infty}w_n=x$ in H. This means that $x=\sum_{k=1}^\infty\langle x,u_k\rangle u_k$ in H.

Let us prove that (3) implies (4). Suppose that for every $x\in H$ we have $x=\lim_{n\to\infty}w_n$ where $w_n=\sum_{k=1}^n\langle x,u_k\rangle u_k$. Note that

$$\|w_n\|^2=\Big\langle\sum_{k=1}^n\langle x,u_k\rangle u_k\,,\,\sum_{\ell=1}^n\langle x,u_\ell\rangle u_\ell\Big\rangle=\sum_{k=1}^n\sum_{\ell=1}^n\langle x,u_k\rangle\overline{\langle x,u_\ell\rangle}\delta_{k,\ell}=\sum_{k=1}^n\big|\langle x,u_k\rangle\big|^2.$$

Let $\epsilon > 0$. Choose $m \in \mathbb{Z}^+$ such that for all $n \geq m$ we have $\|w_n - x\| < \epsilon$. By the Triangle Inequality, for all $n \geq m$ we have $\Big|\|w_n\| - \|x\|\Big| \leq \|w_n - x\| < \epsilon$. This shows that $\lim_{n\to\infty} \|w_n\| = \|x\|$ in $\mathbb{R}$, hence $\|x\|^2 = \lim_{n\to\infty} \|w_n\|^2 = \sum_{k=1}^{\infty} \big|\langle x, u_k\rangle\big|^2$ in $\mathbb{R}$.

Next we prove that (4) implies (5). Suppose that (4) holds. Let $x, y \in H$. Let $x_n = \sum_{k=1}^{n} \langle x, u_k\rangle u_k$ and $y_n = \sum_{k=1}^{n} \langle y, u_k\rangle u_k$. Note that

$$\langle x_n, y_n\rangle = \Big\langle \sum_{k=1}^{n} \langle x, u_k\rangle u_k \, , \, \sum_{\ell=1}^{n} \langle y, u_k\rangle u_k \Big\rangle = \sum_{k=1}^{n}\sum_{\ell=1}^{n} \langle x, u_k\rangle \overline{\langle y, u_\ell\rangle} \delta_{k,\ell} = \sum_{k=1}^{n} \langle x, u_k\rangle \overline{\langle y, u_k\rangle}$$

and note that for $c \in \mathbb{C}$ we have $x_n + cy_n = \sum_{k=1}^{n} \langle x + cy, u_k\rangle u_k$. Since (4) holds, we have $\lim_{n\to\infty} \|x_n\|^2 = \|x\|^2$, $\lim_{n\to\infty} \|y_n\|^2 = \|y\|^2$, and $\lim_{n\to\infty} \|x_n + cy_n\|^2 = \|x + cy\|^2$. By the Polarization Identity,

$$\begin{aligned}
\langle x, y\rangle &= \tfrac{1}{4}\Big(\|x+y\|^2 + i\,\|x+iy\|^2 - \|x-y\|^2 - i\,\|x-iy\|^2\Big) \\
&= \lim_{n\to\infty} \tfrac{1}{4}\Big(\|x_n+y_n\|^2 + i\,\|x_n+iy_n\|^2 - \|x_n-y_n\|^2 - i\,\|x_n-iy_n\|^2\Big) \\
&= \lim_{n\to\infty} \langle x_n, y_n\rangle = \sum_{k=1}^{\infty} \langle x, u_k\rangle \overline{\langle y, u_k\rangle}.
\end{aligned}$$

Note that (4) follows immediately from (5) by taking $y = x$. We finish the proof by proving that (4) implies (2). Suppose that for all $x \in H$ we have $\|x\|^2 = \sum_{k=1}^{\infty} \big|\langle x, u_k\rangle\big|^2$. As above, let $w_n = \text{Proj}_{U_n} = \sum_{k=1}^{n} \langle x, u_k\rangle u_k$ so that $\|w_n\|^2 = \sum_{k=1}^{n} \big|\langle x, u_k\rangle\big|^2$. Then we have $\lim_{n\to\infty} \|w_n\|^2 = \|x\|^2$. Given $\epsilon > 0$, choose $n \in \mathbb{Z}^+$ so that $\|x\|^2 - \|w_n\|^2 < \epsilon^2$. Since $w_n = \text{Proj}_{U_n}(x)$ we have $w_n \in U_n$ and $x - w_n \in {U_n}^\perp$ so that $\langle x - w_n, w_n\rangle = 0$. By Pythagoras' Theorem, $\|x - w_n\|^2 = \|x\|^2 - \|w_n\|^2 < \epsilon^2$, hence $\|x = w_n\| < \epsilon$. Since $\epsilon > 0$ was arbitrary and $w_n \in U$, this shows that U is dense in H.

4.24 Definition: A maximal orthonormal set in a Hilbert space H (over $\mathbb{R}$ or $\mathbb{C}$) is called a **Hilbert basis** for H (over $\mathbb{R}$ or $\mathbb{C}$).

4.25 Theorem: *Let H be an infinite dimensional separable Hilbert space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $\mathcal{A} = \{u_1, u_2, u_3, \cdots\}$ be a countable Hilbert basis for H.*

(1) For all $x \in H$, if $x = \sum_{k=1}^{\infty} a_k u_k$ and $x = \sum_{k=1}^{\infty} b_k u_k$ then $a_k = b_k = \langle x, u_k\rangle$.

(2) For all $c_k \in \mathbb{F}$, $\sum_{k=1}^{\infty} c_k u_k$ converges in H if and only if $\sum_{k=1}^{\infty} |c_k|^2$ converges in $\mathbb{R}$.

(3) The map $\phi : H \to \ell_2(F)$ given by $\phi(x) = \big(\langle x, u_1\rangle, \langle x, u_2\rangle, \cdots\big)$ is an inner product space isomorphism.

Proof: The proof is left as an exercise.

4.26 Definition: When U and V are normed linear spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, a linear map $L : U \to V$ is also called a **linear operator**, and a linear map $f : U \to \mathbb{F}$ is also called a **linear functional** on U.

4.27 Definition: Let U and V be normed linear spaces over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and let $L : U \to V$ be a linear operator. The **operator norm** of L is given by

$$\|L\|_{op} = \sup\left\{ \|Lx\| \,\Big|\, x \in U \text{ with } \|x\| \leq 1 \right\}$$

and we say that L is **bounded** when $\|L\|_{op} < \infty$. Since $Lx = \|x\|\, L\big(\frac{x}{\|x\|}\big)$ for all $0 \neq x \in U$, it follows that

$$\|Lx\| \leq \|L\|_{op}\, \|x\| \text{ for all } x \in U.$$

4.28 Theorem: *Let U and V be normed linear spaces and let $L : U \to V$ be a linear operator. Then the following are equivalent.*
(1) L is continuous at 0.
(2) L is bounded.
(3) L is uniformly continuous in X.

Proof: Suppose that L is continuous at 0. Choose $\delta > 0$ so that $\|x\| \leq \delta \Longrightarrow \|Lx\| \leq 1$. Then $\|x\| \leq 1 \Longrightarrow \|\delta x\| \leq \delta \Longrightarrow \|L(\delta x)\| \leq 1 \Longrightarrow \|L(x)\| = \frac{1}{\delta}\, \|L(\delta x)\| \leq \frac{1}{\delta}$ so $\|L\|_{op} \leq \frac{1}{\delta}$.

Now suppose that L is bounded. For $x, y \in U$ we have

$$\big\|Lx - Ly\big\| = \big\|L(x-y)\big\| = \left\|L\left(\tfrac{x-y}{\|x-y\|}\right)\right\|\, \|x-y\| \leq \|L\|_{op}\|y-x\|.$$

Thus given $\epsilon > 0$ we can choose $\delta = \frac{1}{\|L\|_{op}+1}$ and then $\|x-y\| < \delta \Longrightarrow \|Lx - Ly\| < \epsilon$.

Finally, we note that if L is uniformly continuous in U then L is continuous at 0.

4.29 Theorem: *(The Uniform Boundedness Principle) Let U be a Banach space and let V be a normed linear space. Let S be a set of bounded linear operators $L : U \to V$. Suppose that for every $x \in U$ there exists $m_x \geq 0$ such that $\|Lx\| \leq m_x$ for all $L \in S$. Then there exists $m \geq 0$ such that $\|L\|_{op} \leq m$ for all $L \in S$.*

Proof: For each $n \in \mathbb{Z}^+$, let $A_n = \left\{x \in U \,\big|\, \|Lx\| \leq n \text{ for all } L \in S\right\}$. Note that A_n is closed because the sets $\left\{x \in U \,\big|\, \|Lx\| \leq n\right\}$ are closed for each $L \in S$, and A_n is the intersection of these sets. By the hypothesis of the theorem, we have $U = \bigcup_{n=1}^{\infty} A_n$. By the Baire Category Theorem (since U is complete), the sets A_n cannot all be nowhere dense. Choose $n \in \mathbb{Z}^+$ so that A_n is not nowhere dense. Chose $a \in A_n$ and $r > 0$ so that $\overline{B}(a,r) \subseteq A_n$. For all $x \in U$, if $x \in B(a,r)$ then $x \in A_n$ so we have $\|L(x)\| \leq n$ for all $L \in S$. If $\|x\| < r$ then $x + a \in B(a,r)$ and $a \in B(a,r)$ and so

$$\|L(x)\| = \|L(x+a) - L(a)\| \leq \|L(x+a)\| + \|L(a)\| \leq 2n \text{ for all } L \in S.$$

For all $L \in S$ and $x \in U$, if $\|x\| \leq 1$ then $\|rx\| \leq r$ and so $\|L(x)\| = \frac{1}{r}\, \|L(rx)\| \leq \frac{2n}{r}$. Thus we have $\|L\|_{op} \leq \frac{2n}{r}$ for all $L \in S$.

4.30 Theorem: *(Condensation of Singularities) Let U be a Banach space and let V be a normed linear space. For each $m, n \in \mathbb{Z}^+$, let $L_{m,n} : U \to V$ be a bounded linear operator. Suppose that for each $m \in \mathbb{Z}^+$ there exists $x_m \in U$ such that $\limsup_{n\to\infty} \|L_{mn}(x_m)\| = \infty$. Then the set $E = \left\{ x \in U \,\middle|\, \limsup_{n\to\infty} \|L_{mn}(x)\| = \infty \text{ for all } m \in \mathbb{Z}^+ \right\}$ is a dense $\mathcal{G}_\delta$ set.*

Proof: Fix $m \in \mathbb{Z}^+$. For each $\ell \in \mathbb{Z}^+$, let $A_\ell = \left\{ x \in U \,\middle|\, \|L_{n,m}(x)\| \leq \ell \text{ for all } n \in \mathbb{Z}^+ \right\}$ and note that each set A_ℓ is closed. As in the proof of the Uniform Boundedness Principle, if one of the sets A_ℓ was not nowhere dense then we could choose $m \geq 0$ such that $\|L_{m,n}\| \leq m$ for all $n \in \mathbb{Z}^+$. But then for all $x \in U$ we would have $\|L_{m,n}(x)\| \leq m\|x\|$ for all n so that $\limsup_{n\to\infty} \|L_{m,n}(x)\| \leq m\|x\|$, contradicting the hypothesis of the theorem. Thus all of the sets A_ℓ must be nowhere dense. Let $B_m = \bigcup_{\ell=1}^{\infty} A_\ell = \left\{ x \in U \,\middle|\, \limsup_{n\to\infty} \|L_{m,n}(x)\| < \infty \right\}$ and let $C = \bigcup_{m=1}^{\infty} B_m = \left\{ x \in U \,\middle|\, \limsup_{n\to\infty} \|L_{m,n}(x)\| < \infty \text{ for some } m \in \mathbb{Z}^+ \right\}$, and note that $E = U \setminus C$. Then C is a countable union of closed nowhere dense sets, so E is a countable intersection of open dense sets. By the Baire Category Theorem, E is dense.