EL3370 Mathematical Methods in Signals, Systems and Control

Topic 5: Orthogonal Expansions

Cristian R. Rojas

Division of Decision and Control Systems
KTH Royal Institute of Technology

Outline

Outline

Orthonormal Sets

The notion of basis is very important, since it allows to define "coordinates" in a space, thus allowing explicit computations in Hilbert spaces.

Definition

In an inner product space V, a family $(e_\alpha)_{\alpha\in I}$ in $V \setminus \{0\}$ is an *orthogonal set* if $e_\alpha \perp e_\beta$ for $\alpha \neq \beta$. If also $\|e_\alpha\| = 1$ for all $\alpha \in I$, $(e_\alpha)_{\alpha\in I}$ is an *orthonormal set*. In case I is finite, $\mathbb{N}$ or $\mathbb{Z}$, (e_α) is an *orthogonal*/*orthonormal sequence*.

Examples of orthonormal sets

1. In $\mathbb{C}^n$, take the standard basis vectors.
2. In ℓ_2, take $(e_n)_{n\in\mathbb{N}}$ with $e_n = (0,\dots,0,1,0,\dots)$. (The 1 is in the n-th position.)
3. In $L_2[-\pi,\pi]$, take $(e_n)_{n\in\mathbb{Z}}$, with $e_n(t) = (2\pi)^{1/2}e^{int}$ for $n \in \mathbb{Z}$. (*Fourier basis*)

Orthonormal Sets (cont.)

Definition

If (e_n) is an orthonormal sequence in a Hilbert space H, then, for every $x \in H$, (x, e_n) is the *n-th Fourier coefficient of x w.r.t.* (e_n), and $\sum_{n=1}^{\infty} (x, e_n) e_n$ is the *Fourier series w.r.t.* (e_n).

Lemma

Let $\{e_1, \ldots, e_n\}$ be an orthonormal set in an inner product space V; $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and $x \in V$. Then, $\left\| x - \sum_{k=1}^{n} \lambda_k e_k \right\|^2 = \|x\|^2 + \sum_{k=1}^{n} |\lambda_k - c_k|^2 - \sum_{k=1}^{n} |c_k|^2$, where $c_k := (x, e_k)$.
(*Exercise!*)

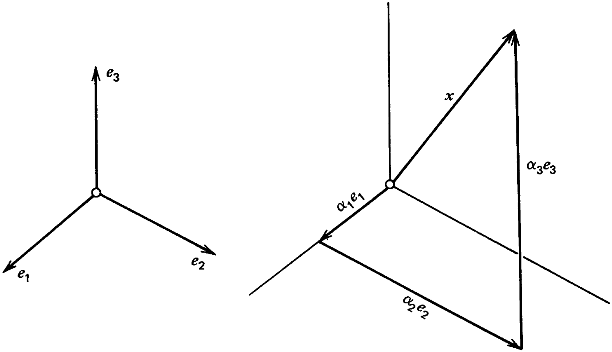
Since $\{e_1, \ldots, e_n\}$ span $\text{lin}\{e_1, \ldots, e_n\}$, we have

Theorem

Let $\{e_1, \ldots, e_n\}$ be an orthonormal set in an inner product space V. The closest point y of $\text{lin}\{e_1, \ldots, e_n\}$ to a point $x \in V$ is $y = \sum_{k=1}^{n} (x, e_k) e_k$, and $\|x - y\|^2 = \|x\|^2 - \sum_{k=1}^{n} |(x, e_k)|^2$.

Orthonormal Sets (cont.)

Corollary
If $x \in \text{lin}\{e_1, \ldots, e_n\}$, then $x = \sum_{k=1}^{n} (x, e_k) e_k$, and $\|x\|^2 = \sum_{k=1}^{n} |(x, e_k)|^2$.



$\alpha_k = (x, e_k)$

Outline

Bessel Inequality

Theorem (Bessel Inequality)
If (e_n) is an orthonormal sequence in an inner product space V, and $x \in V$, then

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leqslant \|x\|^2.$$

Proof. For $N \in \mathbb{N}$, $\sum_{k=1}^{N} |(x, e_k)|^2 = \|x\|^2 - \left\| x - \sum_{k=1}^{N} (x, e_k) e_k \right\|^2 \leqslant \|x\|^2$. Take $N \to \infty$. □

We want to study the meaning of $\sum_{k=1}^{\infty} (x, e_k) e_k$.

Definition (Infinite sum in a normed space)
Let (x_n) be a sequence in a normed space V. We say that $\sum_{n=1}^{\infty} x_n$ *converges* and has *sum* x (*i.e.*, $\sum_{n=1}^{\infty} x_n = x$) if $\sum_{n=1}^{N} x_n \to x$ as $N \to \infty$, *i.e.*, $\left\| x - \sum_{n=1}^{N} x_n \right\| \to 0$ as $N \to \infty$.

Bessel Inequality (cont.)

Theorem

Let (e_n) is an orthonormal sequence in a Hilbert space H, and let (λ_n) be a sequence in $\mathbb{C}$. Then $\sum_{n=1}^{\infty} \lambda_n e_n$ converges in H iff $\sum_{n=1}^{\infty} |\lambda_n|^2 < \infty$.

Proof

($\Rightarrow$) Let $x = \sum_{n=1}^{\infty} \lambda_n e_n$ and $x_N = \sum_{n=1}^{N} \lambda_n e_n$. Then, $(x_N, e_n) = \lambda_n$ for $n \leqslant N$, and taking $N \to \infty$ gives $(x, e_n) = \lambda_n$. Then, by Bessel inequality: $\sum_{n=1}^{\infty} |\lambda_n|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 \leqslant \|x\|^2 < \infty$.

($\Leftarrow$) Assume that $\sum_{n=1}^{\infty} |\lambda_n|^2 < \infty$, and let $x_N = \sum_{n=1}^{N} \lambda_n e_n$. Then,

$$\|x_{N+P} - x_N\|^2 = \left\| \sum_{n=N+1}^{N+P} \lambda_n e_n \right\|^2 = \sum_{n=N+1}^{N+P} \|\lambda_n e_n\|^2 = \sum_{n=N+1}^{N+P} |\lambda_n|^2 \to 0 \quad \text{as } N \to \infty.$$

Therefore, (x_n) is Cauchy, and it converges in H. □

Observation

If $H = L_2[a,b]$, then the above convergence is *in norm* (or L_2 convergence).
A *different* type is *point-wise convergence*: $\sum_{n=1}^{\infty} x_n(t) = x(t)$ for all $t \in [a,b]$.

Outline

Goal: When does $\sum_{n=1}^{\infty}(x, e_n)e_n = x$? $\Rightarrow$ We need conditions on (e_n).

Approximation error: $y = x - \sum_{n=1}^{\infty}(x, e_n)e_n$.
Notice that $(y, e_k) = 0$ for all k. We want to ensure $y = 0$.

Definitions

- An orthonormal set A in an inner product space V is *maximal* if the only point in V which is orthogonal to every $x \in A$ is 0, *i.e.*, A cannot be extended to a larger orthonormal set.
- A set A in a normal space V is *total* (or *fundamental*) if its span is dense in V.
- If A is a total orthonormal set in an inner product space V, every $x \in V$ can be written as $x = \sum_{e \in A}(x, e)e$, and A is called an *orthonormal basis* of V.

Note. By Bessel's inequality, given $x \in V$ and an orthonormal set A, since $\sum_{e \in A} |(x, e)|^2 \leqslant \|x\|$, at most a countable number of terms (x, e), as e runs over A, can be non-zero: for every $n \in \mathbb{N}$, the number of terms s.t. $|(x, e)|^2 > 1/n$ can be at most $n\|x\|^2$, and $\{e \in A : (x, e) \neq 0\} = \bigcup_{n \in \mathbb{N}}\{e \in A : |(x, e)|^2 > 1/n\}$, which is at most countable. Thus, sums like $\sum_{e \in A}(x, e)e$ and $\sum_{e \in A} |(x, e)|^2$ can be reduced to sums over sequences.

Total Orthonormal Sequences (cont.)

Theorem. If A is an orthonormal set in a Hilbert space H, the following are equivalent:

(1) A is total.

(2) $\|x\|^2 = \sum_{e \in A} |(x,e)|^2$ for all $x \in H$.

(3) A is maximal.

If H is an incomplete inner product space, then (1) and (2) are still equivalent, and they imply (3), but not conversely (*see bonus slides for an example*).

Proof

(1) ⇔ (2): For a given $x \in H$, sort the elements of $\{e \in A : (x,e) \neq 0\}$ into a sequence (e_n). Then, take $N \to \infty$ in $\sum_{n=1}^{N} |(x,e_n)|^2 = \|x\|^2 - \left\| x - \sum_{n=1}^{N} (x,e_n)e_n \right\|^2$.

(2) ⇒ (3): If A is not maximal, take a nonzero $x \perp A$. Then $\|x\|^2 > 0 = \sum_{e \in A} |(x,e)|^2$.

(3) ⇒ (1): Given an $x \in H$, $\sum_{e \in A} (x,e)e$ is convergent (due to the completeness of H), and $x - \sum_{e \in A} (x,e)e$ is orthogonal to every $e \in A$, so by maximality of A, $x = \sum_{e \in A} (x,e)e$, which implies that A is an orthonormal basis.

Only the implication (3) ⇒ (1) requires H to be complete. □

Total Orthonormal Sequences (cont.)

Theorem. Let H be an inner product space. Then,

(1) If H is separable, then every orthonormal set in H is countable.

(2) If H contains a total orthonormal sequence, then H is separable.

Proof

(1) If $A \subseteq H$ is an orthonormal set, distinct points $x, y \in A$ are at a distance $\sqrt{(x-y, x-y)} = \sqrt{2}$, so if A were uncountable, a set dense in H would be uncountable too.

(2) If (e_n) is a total orthonormal set, consider the set D, consisting of all linear combinations $\lambda_1 e_1 + \cdots + \lambda_n e_n$ where $n \in \mathbb{N}$ and $\lambda_k = a_k + i b_k$ with $a_k, b_k \in \mathbb{Q}$ for $k = 1, \dots, n$. D is a countable set dense in H (*why?*). □

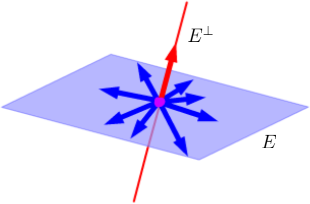
Observation: A separable Hilbert space is isomorphic to $\mathbb{C}^n$ (for some n) or to ℓ_2 (*see bonus slides for proof*).

See bonus slides for examples of non-separable Hilbert spaces.

Outline

We can use orthogonality to decompose a Hilbert space.

Definition
Let H be a Hilbert space. The *orthogonal complement* of $E \subseteq H$ is $E^\perp := \{x \in H : x \perp E\}$.



Theorem
For every subset E of a Hilbert space, $E^\perp$ is a closed linear space. (*Exercise!*)

The projection theorem gives the following characterization of $E^\perp$:

Lemma
Let M be a linear subspace of an inner product space V, and let $x \in V$. Then $x \in M^\perp$ iff $\|x - y\| \geqslant \|x\|$ for all $y \in M$.

Orthogonal Complements (cont.)

Definition

Let $M, N \subseteq V$, where V is a vector space. V is the *direct sum* of M and N, denoted $V = M \oplus N$, if every $x \in V$ has a *unique* decomposition $x = y + z$, where $y \in M$ and $z \in N$.

Theorem

Let M be a closed linear subspace of a Hilbert space H. Then, $H = M \oplus M^{\perp}$.

Proof. Let $x \in H$. Assume that $M \neq \emptyset$ (otherwise the result is trivial). Take $y \in M$ as the unique minimizer of $\inf_{m \in M} \|x - m\|$, and $z := x - y$. By the projection theorem, $z \in M^{\perp}$.
If $x = y' + z'$, with $y' \in M$ and $z' \in M^{\perp}$, then $(x - y') \perp M$, so by the projection theorem, $y' = y$, which proves the uniqueness of the decomposition. □

Corollary

If M is a closed linear subspace of a Hilbert space H, then $(M^{\perp})^{\perp} = M$.

Proof. By definition, $M \subseteq (M^{\perp})^{\perp}$. Let $x \in (M^{\perp})^{\perp}$, and write it as $x = y + z$ with $y \in M$ and $z \in M^{\perp}$. Since $x \perp M^{\perp}$, $0 = (x, z) = (y + z, z) = (y, z) + \|z\|^2 = \|z\|^2$, so $z = 0$ and $x \in M$. □

Outline

Let $e_n(t) := (2\pi)^{-1/2} e^{int}$, $t \in [-\pi, \pi]$, $n \in \mathbb{Z}$. We want to prove that $(e_n)_{n \in \mathbb{Z}}$ is total in $L_2[-\pi, \pi]$.

We need to show that $\text{clin}\{e_n : n \in \mathbb{Z}\} = L_2[-\pi, \pi]$. It is known that the closure of $C[-\pi, \pi]$ is $L_2[-\pi, \pi]$, so it is enough to show that for every $f \in C[-\pi, \pi]$ there is a sequence in $\text{clin}\{e_n : n \in \mathbb{Z}\}$ converging to f. An obvious choice is $f_N = \sum_{n=-N}^{N} (f, e_n) e_n$, but it is easier to work with

$$F_m = \frac{1}{m+1}(f_0 + f_1 + \cdots + f_m), \quad m = 0, 1, \ldots \qquad \text{(\textit{Césaro sum of the } } f_N\text{\textit{'s})}$$

Since $(f, e_n) = (2\pi)^{-1/2} \int_{-\pi}^{\pi} f(t) e^{-int} dt$, we have

$$F_m(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\tau) K_m(t - \tau) d\tau, \quad \text{where} \quad K_m(x) := \frac{1}{m+1} \sum_{N=0}^{m} \sum_{n=-N}^{N} e^{-inx}. \quad \text{(\textit{Fejér kernel})}$$

Classical Fourier Series (cont.)

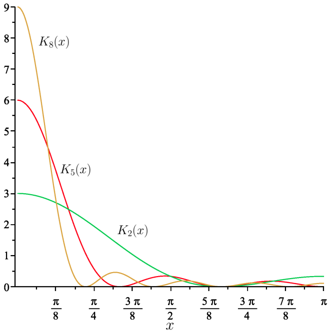
Fejér Kernel properties:

(1) $K_m(x) \geqslant 0$ for all $x \in \mathbb{R}$, $m = 0, 1, 2, \ldots$

(2) $\displaystyle\int_{-\pi}^{\pi} K_m(x)dx = 2\pi$, for $m = 0, 1, 2, \ldots$

(3) For all $0 < \delta < \pi$, $\displaystyle\left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}\right) K_m(x)dx \to 0$ as $m \to \infty$. *(see bonus slides for proofs)*

Therefore, $(K_m/2\pi)$ is a *Delta sequence* (it "converges" to a Dirac delta).



We will prove a strong result: $\displaystyle\lim_{m\to\infty} \sup_{t\in[-\pi,\pi]} |f(t) - F_m(t)| = 0.$

$$\Rightarrow \|f - F_m\|_2^2 = \int_{-\pi}^{\pi} |f(t) - F_m(t)|^2 dt \leqslant 2\pi \sup_{t\in[-\pi,\pi]} |f(t) - F_m(t)|^2 \to 0 \quad \text{as } m \to \infty. \text{ (L_2 convergence)}$$

Classical Fourier Series (cont.)

Take a $\delta > 0$ (to be defined more precisely later):

$$\begin{aligned}
|f(t) - F_m(t)| &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(t) - f(\tau)] K_m(t-\tau) d\tau \right| \\
&\leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t) - f(\tau)| K_m(t-\tau) d\tau \\
&= \frac{1}{2\pi} \left(\int_{\substack{|t-\tau|>\delta \\ -\pi\leqslant\tau\leqslant\pi}} + \int_{t-\delta}^{t+\delta} \right) |f(t) - f(\tau)| K_m(t-\tau) d\tau.
\end{aligned}$$

For the first integral, we use the fact that f is bounded, *i.e.*, there is an $M > 0$ s.t. $\sup_{t\in[-\pi,\pi]} |f(t)| \leqslant M$, hence

$$\frac{1}{2\pi} \int_{\substack{|t-\tau|>\delta \\ -\pi\leqslant\tau\leqslant\pi}} |f(t) - f(\tau)| K_m(t-\tau) d\tau \leqslant \frac{2M}{2\pi} \left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} \right) K_m(\tau) d\tau. \quad \text{(This is negligible as } m \to \infty.\text{)}$$

Classical Fourier Series (cont.)

For the second integral, we need to recall *uniform continuity*:

Definition (reminder). Given metric spaces (X, d_X) and (Y, d_Y), $f : X \to Y$ is *uniformly continuous* if for every $\varepsilon > 0$ there is a $\delta > 0$ s.t. for all $x, y \in X$, $d_X(x, y) < \delta$ implies $d_Y(f(x), f(y)) < \varepsilon$.

Reminder. By Heine-Cantor's theorem, given metric spaces (X, d_X) and (Y, d_Y), if X is compact and $f : X \to Y$ is continuous, then f is uniformly continuous.

Let $\varepsilon > 0$. Then, take δ as in the definition of uniform continuity, so

$$\frac{1}{2\pi}\int_{\substack{|t-\tau|<\delta \\ -\pi\leqslant\tau\leqslant\pi}} |f(t)-f(\tau)|K_m(t-\tau)d\tau \leqslant \frac{\varepsilon}{2\pi}\int_{\substack{|t-\tau|<\delta \\ -\pi\leqslant\tau\leqslant\pi}} K_m(t-\tau)d\tau \leqslant \varepsilon.$$

Therefore: $\sup_{t\in[-\pi,\pi]} |f(t)-F_m(t)| < \frac{M}{2\pi}\left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}\right)K_m(\tau)d\tau + \varepsilon \to \varepsilon \quad \text{as} \quad m \to \infty.$

and since $\varepsilon > 0$ was arbitrary, taking $\varepsilon \to 0$ gives $\lim_{m\to\infty} \sup_{t\in[-\pi,\pi]} |f(t)-F_m(t)| = 0.$ $\square$

Classical Fourier Series (cont.)

We have actually proved

Theorem (Fejér)
Let $f : [-\pi, \pi] \to \mathbb{C}$ be continuous, $s_n(f)$ be the n-th partial sum of its Fourier series, and $\sigma_n(f)$ be the arithmetic mean of $s_0(f), \dots, s_n(f)$. Then $\sigma_n(f) \to f$ uniformly as $n \to \infty$.

Notice that $s_n(f)$ does not always converge point-wisely to continuous f. (*An example is provided in the bonus slides of Topic 9!*)

A similar result (proven analogously, with a different kernel) is

Theorem (Weierstrass theorem)
Let $f : [a, b] \to \mathbb{R}$ be continuous, where $-\infty < a < b < \infty$. For every $\varepsilon > 0$ there is a polynomial p s.t. $\sup_{t \in [a,b]} |f(t) - p(t)| < \varepsilon$.

Least Squares Estimation

Outline

Bonus: Example of Maximal Non-Total Orthonormal Sets

By Zorn's Lemma, every inner product space V has a maximal orthonormal set A (*why?*). If V is complete, $\text{clin } A = V$; otherwise $V = \text{clin } A \oplus (\text{clin } A)^{\perp}$, so $(\text{clin } A)^{\perp} \neq \{0\}$ and there is an $x \in (\text{clin } A)^{\perp}$ of unit norm, so $A \cup \{x\}$ is also orthonormal, contradicting the maximality of A.

However, in every incomplete inner product space V there are maximal orthonormal sets which are not total, *i.e.*, whose closed linear span is not the entire space:

Proof. First note that if every proper, closed subspace M of V is s.t. $M^{\perp} \neq \{0\}$, then V is complete. Indeed, assume that V is incomplete, and let $\hat{V}$ be the completion of V. Pick an $x \in \hat{V} \setminus V$, and let $\hat{M} = \{x\}^{\perp}$ in $\hat{V}$. Then, $\hat{M} \cap V$ is closed in V (because $\hat{M}$ is closed in $\hat{V}$). If $x \perp V$, then $d(x, V) = \|x\| > 0$, and V would not be dense in $\hat{V}$; thus, $\hat{M} \cap V \neq V$, and there is a $y \in V$ s.t. $(x, y) \neq 0$, which we can normalize so that $(x, y) = 1$.
Note that $\hat{M} \cap V$ is dense in $\hat{M}$. Indeed, let $z \in \hat{M}$ and let (x_n) be a sequence in V s.t. $x_n \to z$ (which exists because $\overline{V} = \hat{V}$). Let $x'_n = x_n - (x_n, x)y$; then $x'_n \in V$, $(x'_n, x) = (x_n, x) - (x_n, x)(y, x) = 0$ so that $x'_n \in \hat{M}$, and $\|x'_n - z\| \leqslant \|x_n - z\| + |(x_n, x)| \|y\| \to 0 + |(z, x)| \|y\| = 0$, thus $x'_n \to z$. Then, $(\hat{M} \cap V)^{\perp} \cap V = (\overline{\hat{M} \cap V})^{\perp} \cap V$ $= \hat{M}^{\perp} \cap V = \text{lin}\{x\} \cap V = \emptyset$, so $M = \hat{M} \cap V$ is the sought proper, closed subspace of V.
Now, assume every maximal orthonormal set in an incomplete V is a basis, and let M be a closed, proper subspace of V s.t. $M^{\perp} = \{0\}$. Let B be a maximal orthonormal set in M, and extend it to a maximal orthonormal set $B \cup B_1$ for V. Assume $B_1 \neq \emptyset$, and let $x_1 \in B_1$; since $M^{\perp} = \{0\}$, there is a $y \in M$ s.t. $(y, x_1) \neq 0$. As $B \cup B_1$ is a basis, $y = \sum_k c_k y_k + \sum_k d_k x_k$ ($y_k \in B$, $x_k \in B_1$). Now, $z = \sum_k d_k x_k = y - \sum_k c_k y_k \in M$, but $x_k \perp B$ for all k, hence $z \perp B$. As B is maximal in M, $z = 0$, so $(y, x_1) = d_1 = 0$, a contradiction. Hence, $B_1 = \emptyset$, B is a maximal orthonormal set for V, so B is a basis for V, *i.e.*, $M = V$, a contradiction. Thus, V contains a non-total maximal orthonormal set. □

Bonus: Example of Maximal Non-Total Orthonormal Sets (cont.)

From this result, every incomplete inner product space has maximal non-total orthonormal sets. Here is a specific example:

Let $V = \ell_2$, and denote by (e_n) its standard orthonormal basis. Consider the linear subspace $Y \subseteq V$ spanned by $A = \{a, e_2, e_3, \ldots\}$, where $a := \sum_{k=1}^{\infty} (1/k) e_k$. Then, $B = \{e_2, e_3, \ldots\}$ is a maximal orthonormal set in Y, because if $x = \alpha_1 a + \sum_{k=2}^{N} \alpha_k e_k \in Y$ is orthogonal to B (*why is it enough to consider such an x?*), then $0 = (x, e_{N+1}) = \alpha_1/(N+1)$, and $0 = (x, e_k) = \alpha_k$ for $k = 2, \ldots, N$, hence $x = 0$. However, clin B does not include a, so B is a maximal orthonormal set for Y which is not total in Y.
Note, however, that Y does have an orthonormal basis, which can be obtained by applying Gram-Schmidt to A (*see exercise set 3!*).

Definition. Two Hilbert spaces H, K are *isomorphic* if there is a bijective mapping $U: H \to K$ s.t., for all $x, y \in H$ and $\alpha \in \mathbb{C}$, $U(x+y) = U(x) + U(y)$, $U(\alpha x) = \alpha U(x)$ and $(U(x), U(y)) = (x, y)$. Such a mapping is a *unitary linear operator*.

Theorem. Every separable Hilbert space is isomorphic to $\mathbb{C}^n$ for some $n \in \mathbb{N}$, or to ℓ_2.

Proof. Assume H is a separable Hilbert space, so it has a total orthonormal sequence. Suppose first that such sequence is finite, say, $\{e_1, \dots, e_n\}$. Then, $x = \sum_{k=1}^{n} (x, e_k) e_k$ for each $x \in H$. Let $U: H \to \mathbb{C}^n$ be given by $U\left(\sum_{k=1}^{n} \lambda_k e_k\right) = (\lambda_1, \dots, \lambda_n)$; U is bijective and linear, and if $x = \sum_{k=1}^{n} x_k e_k$, $y = \sum_{k=1}^{n} y_k e_k$, we have that $(x, y) = \sum_{k=1}^{n} x_k \overline{y_k} = (U(x), U(y))$, so U is unitary and H is isomorphic to $\mathbb{C}^n$.
If the total orthonormal sequence is infinite, say, $(e_k)_{k \in \mathbb{N}}$, define the mapping $U: H \to \ell_2$ by $U(x) = (\lambda_k)_{k \in \mathbb{N}}$, where $x = \sum_{k=1}^{\infty} \lambda_k e_k$. U is linear and unitary (as in the finite case), hence injective. By the characterization of total orthonormal sequences, $U(x) \in \ell_2$, and if $(\lambda_k)_{k \in \mathbb{N}} \in \ell_2$, $\sum_{k=1}^{\infty} \lambda_k e_k$ converges to an $x \in \ell_2$, so U is surjective. Thus, H is isomorphic to ℓ_2. □

Bonus: Examples of Non-Separable Hilbert Spaces

1. $\ell_2(\mathbb{R})$**:** The space of all $f: \mathbb{R} \to \mathbb{R}$ s.t. $E_f = \{x \in \mathbb{R}: f(x) \neq 0\}$ is countable and $\sum_{x \in E_f} f^2(x) < \infty$ (this sum is well defined, *why?*), with inner product $(f,g) = \sum_{x \in E_f \cap E_g} f(x)\overline{g(x)}$. $\ell_2(\mathbb{R})$ is a Hilbert space (*Exercise! Hint: countable unions of countable sets are countable*). Also, the functions $f_y \in \ell_2(\mathbb{R})$, with $f_y(x) = 1$ if $x = y$ and $f_y(x) = 0$ otherwise, are an uncountable orthonormal system, so $\ell_2(\mathbb{R})$ is non-separable.
2. **Almost-periodic functions:** In an attempt to extend the classical Fourier series to non-periodic functions in $\mathbb{R}$, the following definition has been coined:

   $f: \mathbb{R} \to \mathbb{C}$ is *almost-periodic* (AP) if it is the uniform limit of functions $\sum_{k=1}^n a_k e^{i\lambda_k t}$, with $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. The set E of AP functions is a vector space, with inner product $(f,g) = \lim_{T \to \infty} (2T)^{-1} \int_{-T}^{T} f(t)\overline{g(t)} dt$ (modulo an equivalence relation). The completion of E is a Hilbert space, but not all its elements can be identified as functions (*e.g.*, $\sum_{k=1}^{\infty} (1/k) e^{it/k}$). Also, $(e^{\lambda t})_{\lambda \in \mathbb{R}}$ is an uncountable orthonormal system in E, so E is non-separable.

Bonus: Proofs of Properties of Fejér Kernels

Letting $z = e^{ix}$, the Fejér kernel can be written, for every x not a multiple of 2π, as

$$K_m(x) = \frac{1}{m+1}\sum_{N=0}^{m}\sum_{n=-N}^{N} z^{-n} = \frac{1}{m+1}\sum_{N=0}^{m}\frac{z^N - z^{-N-1}}{1-z^{-1}} = \frac{1}{(m+1)(1-z^{-1})}\left[\frac{1-z^{m+1}}{1-z} - \frac{z^{-1}-z^{-m-2}}{1-z^{-1}}\right]$$

$$= \frac{1}{(m+1)(1-z^{-1})}\left[\frac{1-z^{m+1}}{1-z} + \frac{1-z^{-m-1}}{1-z}\right] = \frac{2-z^{m+1}-z^{-m-1}}{(m+1)(|1-z|^2)} = \frac{\sin^2\left(\frac{(m+1)x}{2}\right)}{(m+1)\sin^2\left(\frac{x}{2}\right)}. \qquad (*)$$

This, and the continuity of K_m, directly proves Property 1.

Since $\int_{-\pi}^{\pi} e^{inx}dx = 2\pi$ if $n = 0$ and $= 0$ otherwise, $\int_{-\pi}^{\pi} K_m(x)dx = (m+1)^{-1}\sum_{N=0}^{m}\sum_{n=-N}^{N}\int_{-\pi}^{\pi} e^{inx}dx = (m+1)^{-1}\sum_{N=0}^{m} 2\pi = 2\pi$, which establishes Property 2.

Finally, note that if $x \in [-\pi,-\delta)\cup(\delta,\pi]$, then $\sin^2(x/2) \geqslant \sin^2(\delta/2) > 0$. Thus, by $(*)$, for this range of values of x, $0 \leqslant K_m(x) \leqslant (m+1)^{-1}\sin^{-2}(\delta/2)$, so

$$0 \leqslant \left(\int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}\right) K_m(x)dx \leqslant \frac{2\pi}{m+1}\sin^{-2}(\delta/2) \to 0 \quad \text{as } m \to \infty,$$

which proves Property 3. □