Introduction to Functional Analysis

George Voutsadakis[1]

[1] Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

Subsection 1

Inner Product Spaces. Hilbert Spaces

Inner Product Spaces

Definition (Inner Product Space)

An **inner product** on a vector space X is a mapping of $X \times X$ into the scalar field K of X, i.e., with every pair of vectors x and y, there is associated a scalar which is written $\langle x, y \rangle$ and is called the **inner product** of x and y, such that, for all vectors x, y, z and scalars α, we have:

(IP1) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;

(IP2) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$;

(IP3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$;

(IP4) $\langle x, x \rangle \geq 0$
$\langle x, x \rangle = 0 \Longleftrightarrow x = 0$

An **inner product space** (or **pre-Hilbert space**) is a vector space X with an inner product defined on X.

- If X is a real vector space, we have $\langle x, y \rangle = \langle y, x \rangle$. (**Symmetry**)

Hilbert Spaces

Definition (Hilbert Spaces)

An inner product on X defines a **norm** on X given by

$$\|x\| = \sqrt{\langle x, x \rangle}$$

and a **metric** on X given by

$$d(x,y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}.$$

A **Hilbert space** is a complete inner product space (complete in the metric defined by the inner product).

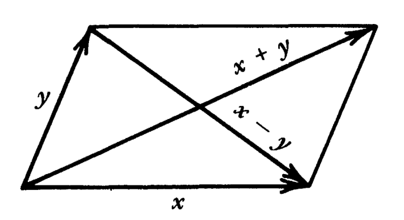
- Hence, inner product spaces are normed spaces, and Hilbert spaces are Banach spaces.
- The proof that $\|x\|$ above satisfies the axioms (N1) to (N4) of a norm will be given later.

Sesquilinearity of the Inner Product

- From (IP1) to (IP3) we obtain the formulas
  (a) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$;
  (b) $\langle x, \alpha y\rangle = \overline{\alpha}\langle x, y\rangle$;
  (c) $\langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle$;
- (a) shows that the inner product is linear in the first factor.
- Since in (c) we have complex conjugates $\overline{\alpha}$ and $\overline{\beta}$ on the right, we say that the inner product is **conjugate linear** in the second factor.
- Expressing both properties together, we say that the inner product is **sesquilinear**.

The Parallelogram Equality

- The reader may show by a simple straightforward calculation that a norm on an inner product space satisfies the important **parallelogram equality**
$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2).$$



- If a norm does not satisfy the parallelogram equality, it cannot be obtained from an inner product.
- Such norms do exist:

  Not all normed spaces are inner product spaces.

Orthogonality

- Recall that if the dot product of two vectors in three dimensional spaces is zero, the vectors are orthogonal, i.e., they are perpendicular or at least one of them is the zero vector.

Definition (Orthogonality)

An element x of an inner product space X is said to be **orthogonal** to an element $y \in X$ if $\langle x, y \rangle = 0$. We also say that x and y are **orthogonal**, and we write $x \perp y$.

Similarly, for subsets $A, B \subseteq X$ we write:

- $x \perp A$ if $x \perp a$, for all $a \in A$;
- $A \perp B$ if $a \perp b$, for all $a \in A$ and all $b \in B$.

The Euclidean Space $\mathbb{R}^n$

- The space $\mathbb{R}^n$ is a Hilbert space with inner product defined by
$$\langle x,y\rangle = \xi_1\eta_1 + \cdots + \xi_n\eta_n,$$
where $x=(\xi_j)=(\xi_1,\ldots,\xi_n)$ and $y=(\eta_j)=(\eta_1,\ldots,\eta_n)$.
- By definition, we obtain
$$\|x\| = \langle x,x\rangle^{1/2} = (\xi_1^2+\cdots+\xi_n^2)^{1/2}.$$
- And, also, the Euclidean metric defined by
$$d(x,y)=\|x-y\|=\langle x-y,x-y\rangle^{1/2}=[(\xi_1-\eta_1)^2+\cdots+(\xi_n-\eta_n)^2]^{1/2}.$$
- Completeness has been proved.
- If $n=3$, we get the usual dot product $\langle x,y\rangle = x\cdot y = \xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3$ of $x=(\xi_1,\xi_2,\xi_3)$ and $y=(\eta_1,\eta_2,\eta_3)$.
The orthogonality $\langle x,y\rangle = x\cdot y=0$ agrees with the elementary concept of perpendicularity.

The Unitary Space $\mathbb{C}^n$

- The space $\mathbb{C}^n$ is a Hilbert space with inner product given by

$$\langle x, y \rangle = \xi_1 \overline{\eta_1} + \cdots + \xi_n \overline{\eta_n}.$$

- By definition, we obtain the norm defined by

$$\|x\| = (\xi_1 \overline{\xi_1} + \cdots + \xi_n \overline{\xi_n})^{1/2} = (|\xi_1|^2 + \cdots + |\xi_n|^2)^{1/2}.$$

- We see why we have to take complex conjugates $\overline{\eta_j}$ in the formula.
  - It entails $\langle y, x \rangle = \overline{\langle x, y \rangle}$, which is (IP3);
  - It implies that $\langle x, x \rangle$ is real.

The Real Space $C[a,b]$

- The norm is defined by

$$\|x\| = \left(\int_a^b x(t)^2 dt\right)^{1/2}.$$

- It can be obtained from the inner product

$$\langle x, y\rangle = \int_a^b x(t)y(t)dt.$$

The Complex Space $C[a,b]$

- We may consider **complex-valued functions** (keeping $t \in [a,b]$ real). These functions form a complex vector space, which becomes an inner product space if we define
$$\langle x, y\rangle = \int_a^b x(t)\overline{y(t)}dt.$$
- The complex conjugate has the effect that (IP3) holds, so that $\langle x, x\rangle$ is still real.
- This property is again needed in connection with the norm, which is now defined by
$$\|x\| = \left(\int_a^b |x(t)|^2 dt\right)^{1/2}$$
since $x(t)\overline{x(t)} = |x(t)|^2$.
- The completion of the metric space corresponding to
  - real functions is the real space $L^2[a,b]$;
  - complex functions is the complex space $L^2[a,b]$.

The Hilbert Sequence Space ℓ^2

- The space ℓ^2 is a Hilbert space with inner product defined by

$$\langle x, y\rangle = \sum_{j=1}^{\infty} \xi_j \overline{\eta_j}.$$

- Convergence of this series follows from the Cauchy-Schwarz inequality.
- This inner products generalizes $\langle x, y\rangle = \xi_1\overline{\eta_1} + \cdots \xi_n\overline{\eta_n}$.
- The norm is defined by

$$\|x\| = \langle x, x\rangle^{1/2} = \left(\sum_{j=1}^{\infty} |\xi_j|^2\right)^{1/2}.$$

- Completeness has already been established.

The Space $\ell^p, p \neq 2$

- The space ℓ^p, with $p \neq 2$, is not an inner product space, hence not a Hilbert space.

  This means that the norm of ℓ^p, with $p \neq 2$, cannot be obtained from an inner product.

  We prove this by showing that the norm does not satisfy the parallelogram equality: Take $x = (1, 1, 0, 0, \ldots) \in \ell^p$ and $y = (1, -1, 0, 0, \ldots) \in \ell^p$. Calculate

  $$\|x\| = \|y\| = 2^{1/p}, \qquad \|x+y\| = \|x-y\| = 2.$$

  Since, if $p \neq 2$, $2^2 + 2^2 \neq 2(2^{2/p} + 2^{2/p})$, we see that the parallelogram equality is not satisfied if $p \neq 2$.
- Since ℓ^p is complete, ℓ^p, with $p \neq 2$, is a Banach space which is not a Hilbert space.

The Space $C[a,b]$

- The space $C[a,b]$ is not an inner product space, hence not a Hilbert space.

  We show that the norm defined by

  $$\|x\| = \max_{t \in J} |x(t)|, \quad J = [a,b],$$

  cannot be obtained from an inner product since this norm does not satisfy the parallelogram equality.

  Indeed, if we take $x(t) = 1$ and $y(t) = \frac{t-a}{b-a}$, we have $\|x\| = 1$, $\|y\| = 1$ and

  $$x(t) + y(t) = 1 + \frac{t-a}{b-a}, \quad x(t) - y(t) = 1 - \frac{t-a}{b-a}.$$

  Hence $\|x+y\| = 2$, $\|x-y\| = 1$ and

  $$\|x+y\|^2 + \|x-y\|^2 = 5, \qquad 2(\|x\|^2 + \|y\|^2) = 4.$$

The Polarization Identity

- To an inner product there corresponds a norm which is given by

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

- It is remarkable that, conversely, we can "rediscover" the inner product from the corresponding norm:
  - For a real inner product space

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

  - For a complex inner product space

$$\begin{aligned} \mathrm{Re}\langle x, y\rangle &= \tfrac{1}{4}(\|x+y\|^2 - \|x-y\|^2) \\ \mathrm{Im}\langle x, y\rangle &= \tfrac{1}{4}(\|x+iy\|^2 - \|x-iy\|^2). \end{aligned}$$

  This formula is sometimes called the **polarization identity**.

Subsection 2

Further Properties of Inner Product Spaces

The Norm Induced by an Inner Product

- Given an inner product $\langle\cdot,\cdot\rangle$, $\|x\| = \sqrt{\langle x,x\rangle}$ defines a norm:
  - (N1) and (N2) follow from (IP4).
  - (N3) is obtained by the use of (IP2) and (IP3): In fact, $\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\overline{\alpha}\langle x,x\rangle = |\alpha|^2\|x\|^2$.
  - (N4) is included in the following:

Lemma (Schwarz Inequality, Triangle Inequality)

An inner product and the corresponding norm satisfy the Schwarz inequality and the triangle inequality:

(a) We have

$$|\langle x,y\rangle| \leq \|x\|\|y\| \qquad \text{(\textbf{Schwarz Inequality})}$$

where the equality sign holds if and only if $\{x,y\}$ is a linearly dependent set.

(b) That norm also satisfies

$$\|x+y\| \leq \|x\| + \|y\| \qquad \text{(\textbf{Triangle Inequality})}$$

where the equality sign holds if and only it $y = 0$ or $x = cy$ (c real and ≥ 0).

Proof of the Schwarz Inequality

- If $y=0$, then $|\langle x,y\rangle| \leq \|x\|\|y\|$ holds, since $\langle x,0\rangle = 0$.

  Let $y \neq 0$. For every scalar α, we have

$$\begin{aligned} 0 \leq \|x-\alpha y\|^2 &= \langle x-\alpha y, x-\alpha y\rangle \\ &= \langle x,x\rangle - \overline{\alpha}\langle x,y\rangle - \alpha[\langle y,x\rangle - \overline{\alpha}\langle y,y\rangle]. \end{aligned}$$

  The expression in the brackets is zero if we choose $\overline{\alpha} = \frac{\langle y,x\rangle}{\langle y,y\rangle}$. The remaining inequality is

$$0 \leq \langle x,x\rangle - \frac{\langle y,x\rangle}{\langle y,y\rangle}\langle x,y\rangle = \|x\|^2 - \frac{|\langle x,y\rangle|^2}{\|y\|^2},$$

  where we used $\langle y,x\rangle = \overline{\langle x,y\rangle}$. Multiplying by $\|y\|^2$, transferring the last term to the left and taking square roots, we get the inequality.

  Equality holds in this derivation if and only if $y=0$ or $0 = \|x-\alpha y\|^2$. Hence $x-\alpha y = 0$, so that $x = \alpha y$, proving linear dependence.

Proof of the Triangle Inequality

- We have

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

By the Schwarz inequality,

$$|\langle x,y\rangle| = |\langle y,x\rangle| \le \|x\|\|y\|.$$

By the triangle inequality for numbers, we thus obtain

$$\begin{array}{rcl} \|x+y\|^2 & \le & \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ & \le & \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ & = & (\|x\|+\|y\|)^2. \end{array}$$

Taking square roots on both sides, we have the inequality.

Equality in the Triangle Inequality

- Equality holds in this derivation if and only if $\langle x,y\rangle + \langle y,x\rangle = 2\|x\|\|y\|$. The left-hand side is $2\text{Re}\langle x,y\rangle$, where Re denotes the real part. Thus, $\text{Re}\langle x,y\rangle = \|x\|\|y\| \geq |\langle x,y\rangle|$. Since the real part of a complex number cannot exceed the absolute value, we must have equality. This implies linear dependence by part (a), say, $y=0$ or $x=cy$.

  We show that c is real and ≥ 0.

  We have $\text{Re}\langle x,y\rangle = |\langle x,y\rangle|$. But if the real part of a complex number equals the absolute value, the imaginary part must be zero. Hence $\langle x,y\rangle = \text{Re}\langle x,y\rangle \geq 0$. Now we get

$$0 \leq \langle x,y\rangle = \langle cy,y\rangle = c\|y\|^2.$$

  Therefore, $c \geq 0$.

Continuity of the Inner Product

Lemma (Continuity of Inner Product)

If in an inner product space, $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

- Subtracting and adding a term, using the triangle inequality for numbers and, finally, the Schwarz inequality, we obtain

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n, y_n \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle| \\ &\leq |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle| \\ &\leq \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\| \stackrel{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

since $y_n - y \to 0$ and $x_n - x \to 0$ as $n \to \infty$.

Isomorphisms of Inner Product Spaces

- An **isomorphism** T of an inner product space X onto an inner product space $\widetilde{X}$ over the same field is a bijective linear operator $T: X \to \widetilde{X}$ which preserves the inner product, i.e., for all $x, y \in X$,

$$\langle Tx, Ty \rangle = \langle x, y \rangle,$$

  where we denoted inner products on X and $\widetilde{X}$ by the same symbol.
- $\widetilde{X}$ is then called **isomorphic** with X.
- X and $\widetilde{X}$ are called **isomorphic inner product spaces**.
- The bijectivity and linearity guarantees that T is a vector space isomorphism of X onto $\widetilde{X}$; So T preserves the whole structure of inner product space.
- T is also an isometry of X onto $\widetilde{X}$ because distances in X and $\widetilde{X}$ are determined by the norms defined by the inner products on X and $\widetilde{X}$.

Completion of an Inner Product Space

Theorem (Completion)

For any inner product space X, there exists a Hilbert space H and an isomorphism A from X onto a dense subspace $W \subseteq H$. The space H is unique except for isomorphisms.

- We know there exists a Banach space H and an isometry A from X onto a subspace W of H which is dense in H. For reasons of continuity, under such an isometry, sums and scalar multiples of elements in X and W correspond to each other. So A is even an isomorphism of X onto W, both regarded as normed spaces.

  The preceding lemma shows that we can define an inner product on H by setting

  $$\langle \widehat{x}, \widehat{y} \rangle = \lim_{n \to \infty} \langle x_n, y_n \rangle,$$

  with (x_n) and (y_n) representatives of $\widehat{x} \in H$ and $\widehat{y} \in H$, respectively.

Completion of an Inner Product Space

- We can see that A is an isomorphism of X onto W, both regarded as inner product spaces.

  The normed space completion theorem also guarantees that H is unique except for isometries, that is, two completions H and $\widetilde{H}$ of X are related by an isometry $T : H \to \widetilde{H}$.

  Reasoning as in the case of A, we conclude that T must be an isomorphism of the Hilbert space H onto the Hilbert space $\widetilde{H}$.

Subspaces

- A **subspace** Y of an inner product space X is defined to be a vector subspace of X taken with the inner product on X restricted to $Y \times Y$.
- Similarly, a **subspace** Y of a Hilbert space H is defined to be a subspace of H, regarded as an inner product space.

  Note that Y need not be a Hilbert space because Y may not be complete.

Theorem (Subspace)

Let Y be a subspace of a Hilbert space H. Then:

(a) Y is complete if and only if Y is closed in H.

(b) If Y is finite dimensional, then Y is complete.

(c) If H is separable, so is Y. More generally, every subset of a separable inner product space is separable.

Subspaces (Cont'd)

- Parts (a) and (b) follow from results on normed spaces.

  Suppose H is separable. Then H has a countable dense subset $\{h_n\}_{n=1}^{\infty}$. Let $Y \subseteq H$. For all integers $m, n > 0$, let $y_{mn} \in Y$ be such that $\|y_{mn} - h_n\| < \frac{1}{m}$, if such an element exists. Clearly $\{y_{mn}\}$ is countable. It suffices to show that it is dense in Y.

  Let $y \in Y$ and $\varepsilon > 0$. Let m be such that $\frac{1}{m} < \frac{\varepsilon}{2}$.

  Since $\{h_n\}$ is dense in H, there exists $n > 0$, such that $\|y - h_n\| < \frac{1}{m}$. Note that this shows that $y_{mn} \in Y$ is defined. Now we get

  $$\|y - y_{mn}\| \leq \|y - h_n\| + \|y_{mn} - h_n\| < \frac{1}{m} + \frac{1}{m} < \varepsilon.$$

  Hence $\{y_{mn}\}$ is dense in Y.

Subsection 3

Orthogonal Complements and Direct Sums

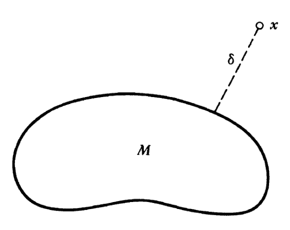
Distance

- In a metric space X, the **distance** δ from an element $x \in X$ to a nonempty subset $M \subseteq X$ is defined to be

$$\delta = \inf_{\tilde{y} \in M} d(x, \tilde{y}), \quad M \neq \emptyset.$$

- In a normed space this becomes

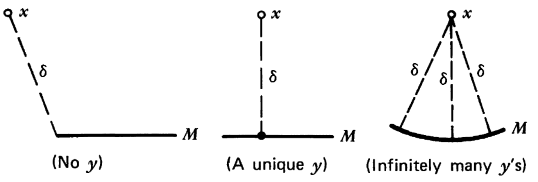
$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\|, \quad M \neq \emptyset.$$



- It is important to know whether there is a $y \in M$, such that $\delta = \|x - y\|$, i.e., intuitively speaking, a point $y \in M$ which is closest to the given x, and if such an element exists, whether it is unique.

Discussion on Distance

- Even in a very simple space such as the Euclidean plane $\mathbb{R}^2$, there may be no y satisfying $\delta = \|x - y\|$, or precisely one such y, or more than one y:



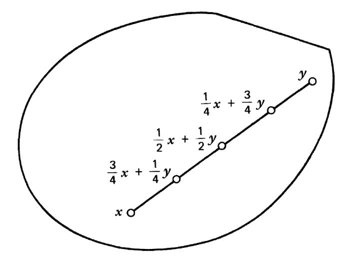
- And we may expect that other spaces, in particular infinite dimensional ones, will be much more complicated in that respect.
- For general normed spaces this is the case, but for Hilbert spaces the situation remains relatively simple.

Segments and Convexity

- The **segment** joining two given elements x and y of a vector space X is defined to be the set of all $z \in X$ of the form

$$z = \alpha x + (1-\alpha)y, \quad \alpha \in \mathbb{R},\ 0 \le \alpha \le 1.$$

- A subset M of X is said to be **convex** if for every $x, y \in M$, the segment joining x and y is contained in M.



- Every subspace Y of X is convex.
- The intersection of convex sets is a convex set.

Minimizing Vector

Theorem (Minimizing Vector)

Let X be an inner product space and $M \neq \emptyset$ a convex subset which is complete (in the metric induced by the inner product). Then, for every $x \in X$, there exists unique $y \in M$, such that $\delta = \inf_{\widetilde{y} \in M} \|x - \widetilde{y}\| = \|x - y\|$.

(a) **Existence** By the definition of an infimum, there is a sequence (y_n) in M, such that $\delta_n \to \delta$ where $\delta_n = \|x - y_n\|$. We show that (y_n) is Cauchy. Let $y_n - x = v_n$. Then $\|v_n\| = \delta_n$ and

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2\left\|\frac{1}{2}(y_n + y_m) - x\right\| \geq 2\delta$$

because M is convex, so that $\frac{1}{2}(y_n + y_m) \in M$. Furthermore, we have $y_n - y_m = v_n - v_m$. Hence by the parallelogram equality,

$$\begin{aligned} \|y_n - y_m\|^2 = \|v_n - v_m\|^2 &= -\|v_n + v_m\|^2 + 2(\|v_n\|^2 + \|v_m\|^2) \\ &\leq -(2\delta)^2 + 2(\delta_n^2 + \delta_m^2). \end{aligned}$$

Hence (y_n) is Cauchy.

Minimizing Vector (Cont'd)

- Since (y_n) is Cauchy and M is complete, (y_n) converges, say, $y_n \to y \in M$. Since $y \in M$, we have $\|x-y\| \geq \delta$. Also, $\|x-y\| \leq \|x-y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \to \delta$. So $\|x-y\| = \delta$.

(b) **Uniqueness** We assume that $y \in M$ and $y_0 \in M$ both satisfy $\|x-y\| = \delta$ and $\|x-y_0\| = \delta$ and show that then $y_0 = y$. By the parallelogram equality,

$$\begin{aligned}\|y-y_0\|^2 &= \|(y-x)-(y_0-x)\|^2 \\ &= 2\|y-x\|^2 + 2\|y_0-x\|^2 - \|(y-x)+(y_0-x)\|^2 \\ &= 2\delta^2 + 2\delta^2 - 2^2\|\tfrac{1}{2}(y+y_0)-x\|^2.\end{aligned}$$

On the right, $\frac{1}{2}(y+y_0) \in M$, so that $\|\frac{1}{2}(y+y_0)-x\| \geq \delta$. This implies that the right-hand side is less than or equal to $2\delta^2 + 2\delta^2 - 4\delta^2 = 0$. Hence, we have the inequality $\|y-y_0\| \leq 0$. Clearly, $\|y-y_0\| \geq 0$. So we must have equality, and $y_0 = y$.

Orthogonality

- In geometry, the unique point y in a given subspace Y closest to a given x is found by "dropping a perpendicular from x to Y".

Lemma (Orthogonality)

Let X be an inner product space and $Y \neq \emptyset$ a complete subspace and $x \in X$ fixed. Then $z = x - y$ is orthogonal to Y.

- If $z \perp Y$ were false, there would be a $y_1 \in Y$ such that $\langle z, y_1 \rangle = \beta \neq 0$. Clearly, $y_1 \neq 0$, since otherwise $\langle z, y_1 \rangle = 0$. Furthermore, for any scalar α,

$$\begin{aligned} \|z - \alpha y_1\|^2 &= \langle z - \alpha y_1, z - \alpha y_1 \rangle \\ &= \langle z, z \rangle - \overline{\alpha} \langle z, y_1 \rangle - \alpha [\langle y_1, z \rangle - \overline{\alpha} \langle y_1, y_1 \rangle] \\ &= \langle z, z \rangle - \overline{\alpha} \beta - \alpha [\overline{\beta} - \overline{\alpha} \langle y_1, y_1 \rangle]. \end{aligned}$$

The expression in the brackets is zero if we choose $\overline{\alpha} = \frac{\overline{\beta}}{\langle y_1, y_1 \rangle}$. We now have $\|z\| = \|x - y\| = \delta$. So $\|z - \alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle y_1, y_1 \rangle} < \delta^2$. This is impossible, since $z - \alpha y_1 = x - (y + \alpha y_1)$ implies $\|z - \alpha y_1\| \geq \delta$ by the definition of δ.

Direct Sum of Vector Spaces

Definition (Direct Sum)

A vector space X is said to be the **direct sum** of two subspaces Y and Z of X, written $X = Y \bigoplus Z$, if each $x \in X$ has a unique representation

$$x = y + z, \qquad y \in Y, \ z \in Z.$$

Then Z is called an **algebraic complement** of Y in X and vice versa. Y, Z is called a **complementary pair** of subspaces in X.

Example: $Y = \mathbb{R}$ is a subspace of the Euclidean plane $\mathbb{R}^2$. Clearly, Y has infinitely many algebraic complements in $\mathbb{R}^2$, each of which is a real line. But most convenient is a complement that is perpendicular. In $\mathbb{R}^3$ the situation is the same in principle.

- In the case of a general Hilbert space H, the main interest concerns representations of H as a direct sum of a closed subspace Y and its **orthogonal complement** $Y^\perp = \{z \in H : z \perp Y\}$, which is the set of all vectors orthogonal to Y.

Direct Sum or Projection Theorem

Theorem (Direct Sum)

Let Y be any closed subspace of a Hilbert space H. Then $H = Y \bigoplus Z$, $Z = Y^{\perp}$.

- Since H is complete and Y is closed, Y is complete.

  Since Y is convex, by the preceding theorem and lemma, for every $x \in H$, there is a $y \in Y$, such that $x = y + z$, $z \in Z = Y^{\perp}$.

  To prove uniqueness, we assume that

  $$x = y + z = y_1 + z_1,$$

  where $y, y_1 \in Y$ and $z, z_1 \in Z$. Then $y - y_1 = z_1 - z$. Since $y - y_1 \in Y$ whereas $z_1 - z \in Z = Y^{\perp}$, we see that $y - y_1 \in Y \cap Y^{\perp} = \{0\}$. This implies $y = y_1$. Hence also $z = z_1$.

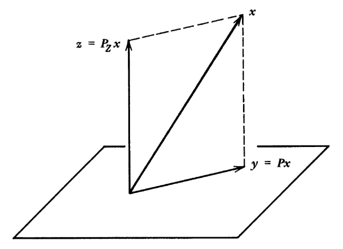
The Canonical Projection

- In $x = y + z$, $z \in Z = Y^{\perp}$, the element y is called the **orthogonal projection** of x on Y (or, briefly, the **projection** of x on Y).
- The equation defines a mapping

$$P : H \to Y;$$

$$x \mapsto y = Px.$$

  P is called the (**orthogonal**) **projection** (or **projection operator**) of H onto Y.



The Canonical Projection and the Null Space Lemma

- Obviously, P is a bounded linear operator.
- P maps H onto Y, Y onto itself, $Z = Y^{\perp}$ onto $\{0\}$.
- P is **idempotent**, that is, $P^2 = P$, i.e., for every $x \in H$, $P^2 x = P(Px) = Px$. Hence $P|_Y$ is the identity operator on Y.
- For $Z = Y^{\perp}$ our discussion yields

Lemma (Null Space)

The orthogonal complement $Y^{\perp}$ of a closed subspace Y of a Hilbert space H is the null space $\mathcal{N}(P)$ of the orthogonal projection P of H onto Y.

Annihilators

- The **annihilator** $M^{\perp}$ of a set $M \neq \emptyset$ in an inner product space X is the set
$$M^{\perp} = \{x \in X : x \perp M\}.$$
Thus, $x \in M^{\perp}$ if and only if $\langle x, v \rangle = 0$, for all $v \in M$.
- $M^{\perp}$ is a vector space since $x, y \in M^{\perp}$ implies, for all $v \in M$ and all scalars α, β,
$$\langle ax + \beta y, v \rangle = \alpha \langle x, v \rangle + \beta \langle y, v \rangle = 0.$$
Hence, $\alpha x + \beta y \in M^{\perp}$.
- $(M^{\perp})^{\perp}$ is written $M^{\perp\perp}$, etc.

Annihilators and Closure

Proposition

The annihilator $M^{\perp}$ of a set $M \neq \emptyset$ in an inner product space X is a closed subspace of X.

- Let x be a limit point of $M^{\perp}$.

  Then there exists a sequence $\{x_n\}$ in $M^{\perp}$ such that $x_n \to x$.

  Since the inner product is continuous, it follows that for any fixed y, $\langle x_n, y \rangle \to \langle x, y \rangle$.

  But $\langle x_n, y \rangle = 0$, for all $y \in M$.

  Therefore, $\langle x, y \rangle = 0$, for all $y \in M$.

  Hence, $x \in M^{\perp}$ and $M^{\perp}$ is closed.

A Space and its Double Complement

- In general we have $M \subseteq M^{\perp\perp}$: $x \in M$ implies $x \perp M^{\perp}$ implies $x \in (M^{\perp})^{\perp}$.

Lemma (Closed Subspace)

If Y is a closed subspace of a Hilbert space H, then $Y = Y^{\perp\perp}$.

- In general $Y \subseteq Y^{\perp\perp}$. We show $Y \supseteq Y^{\perp\perp}$. Let $x \in Y^{\perp\perp}$. Then $x = y + z$, where $y \in Y \subseteq Y^{\perp\perp}$. Since $Y^{\perp\perp}$ is a vector space and $x \in Y^{\perp\perp}$ by assumption, we also have $z = x - y \in Y^{\perp\perp}$. Hence, $z \perp Y^{\perp}$. But $z \in Y^{\perp}$. Together $z \perp z$, hence $z = 0$, so that $x = y$, that is, $x \in Y$. Since $x \in Y^{\perp\perp}$ was arbitrary, this proves $Y \supseteq Y^{\perp\perp}$.
- Since $Z^{\perp} = Y^{\perp\perp} = Y$, we get $H = Z \bigoplus Z^{\perp}$.

  So $x \mapsto z$ defines a projection $P_Z : H \to Z$ of H onto Z, whose properties are quite similar to those of the projection $P : H \to Y$.

Sets with Dense Span

Lemma (Dense Set)

For any subset $M \neq \emptyset$ of a Hilbert space H, the span of M is dense in H if and only if $M^\perp = \{0\}$.

- Let $x \in M^\perp$ and assume $V = \operatorname{span} M$ is dense in H. Then $x \in \overline{V} = H$. So, there is a sequence (x_n) in V, such that $x_n \to x$. Since $x \in M^\perp$ and $M^\perp \perp V$, we have $\langle x_n, x \rangle = 0$. The continuity of the inner product implies that $\langle x_n, x \rangle \to \langle x, x \rangle$. Together, $\langle x, x \rangle = \|x\|^2 = 0$, so that $x = 0$. Since $x \in M^\perp$ was arbitrary, this shows that $M^\perp = \{0\}$.

  Conversely, suppose that $M^\perp = \{0\}$. If $x \perp V$, then $x \perp M$, so that $x \in M^\perp$ and $x = 0$. Hence $V^\perp = \{0\}$. Noting that V is a subspace of H, we thus obtain $\overline{V} = H$.

Subsection 4

Orthonormal Sets and Sequences

Orthogonal Families of Vectors

- Of particular interest are sets whose elements are orthogonal in pairs.
- In the space $\mathbb{R}^3$, a set of that kind is the set of the three unit vectors in the positive directions of the axes of a rectangular coordinate system, say e_1, e_2, e_3.

  These vectors form a basis for $\mathbb{R}^3$, so that every $x \in \mathbb{R}^3$ has a unique representation

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3.$$

- A great advantage of the orthogonality is that, given x, we can readily determine the unknown coefficients $\alpha_1, \alpha_2, \alpha_3$ by taking inner products (dot products): To obtain α_1, we must multiply that representation of x by e_1:

$$\langle x, e_1 \rangle = \alpha_1 \langle e_1, e_1 \rangle + \alpha_2 \langle e_2, e_1 \rangle + \alpha_3 \langle e_3, e_1 \rangle = \alpha_1.$$

  and so on.

Orthonormal Sets and Sequences

Definition (Orthonormal Sets and Sequences)

An **orthogonal set** M in an inner product space X is a subset $M \subseteq X$ whose elements are pairwise orthogonal.

An **orthonormal set** $M \subseteq X$ is an orthogonal set in X whose elements have norm 1, that is, for all $x, y \in M$, $\langle x, y \rangle = \begin{cases} 0, & \text{if } x \neq y \\ 1, & \text{if } x = y \end{cases}$.

If an orthogonal or orthonormal set M is countable, we can arrange it in a sequence (x_n) and call it an **orthogonal** or **orthonormal sequence**, respectively.

More generally, an indexed set, or family, $(x_\alpha), \alpha \in I$, is called **orthogonal** if $x_\alpha \perp x_\beta$, for all $\alpha, \beta \in I, \alpha \neq \beta$. The family is called **orthonormal** if it is orthogonal and all x_α have norm 1, so that for all $\alpha, \beta \in I$, we have

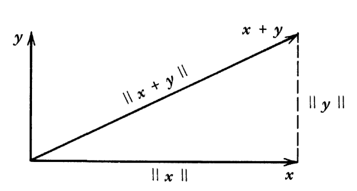
$$\langle x_\alpha, x_\beta \rangle = \delta_{\alpha\beta} = \begin{cases} 0, & \text{if } \alpha \neq \beta \\ 1, & \text{if } \alpha = \beta \end{cases}.$$

The Pythagorean Relation

- For orthogonal elements x, y we have $\langle x, y\rangle = 0$.

  So we obtain the **Pythagorean relation**

  $$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$



- More generally, if $\{x_1, \ldots, x_n\}$ is an orthogonal set, then

  $$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

  In fact, $\langle x_j, x_k\rangle = 0$ if $j \neq k$; consequently,

  $$\left\|\sum_j x_j\right\|^2 = \left\langle \sum_j x_j, \sum_k x_k \right\rangle = \sum_j \sum_k \langle x_j, x_k\rangle = \sum_j \langle x_j, x_j\rangle = \sum_j \|x_j\|^2.$$

Linear Independence

Lemma (Linear Independence)

An orthonormal set is linearly independent.

- Let $\{e_1, \ldots, e_n\}$ be orthonormal and consider the equation

$$\alpha_1 e_1 + \cdots + \alpha_n e_n = 0.$$

Multiplication by a fixed e_j gives

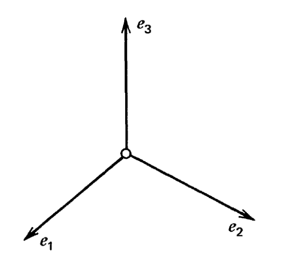
$$\left\langle \sum_k \alpha_k e_k, e_j \right\rangle = \sum_k \alpha_k \langle e_k, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j = 0.$$

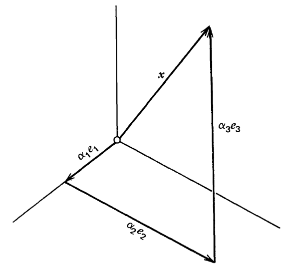
This proves linear independence for any finite orthonormal set.

It also implies linear independence if the given orthonormal set is infinite, by the definition of linear independence.

The Spaces $\mathbb{R}^3$ and ℓ^2

- (**Euclidean Space** $\mathbb{R}^3$) In the space $\mathbb{R}^3$, the three unit vectors $(1,0,0),(0,1,0),(0,0,1)$ in the direction of the three axes of a rectangular coordinate system form an orthonormal set.





- (**Space** ℓ^2) In the space ℓ^2, an orthonormal sequence is (e_n), where $e_n = (\delta_{nj})$ has the n-th element 1 and all others zero.

Continuous Functions

- Let X be the inner product space of all real-valued continuous functions on $[0,2\pi]$ with inner product $\langle x,y\rangle = \int_0^{2\pi} x(t)y(t)dt$.
- An orthogonal sequence in X is (u_n), where $u_n(t) = \cos nt$, $n = 0,1,\ldots$.
- Another orthogonal sequence is (v_n), where $v_n(t) = \sin nt$, $n = 1,2,\ldots$.
- By integration we obtain

$$\langle u_m,u_n\rangle = \int_0^{2\pi} \cos mt \cos mt = \begin{cases} 0, & \text{if } m \neq n \\ \pi, & \text{if } m = n = 1,2,\ldots \\ 2\pi, & \text{if } m = n = 0 \end{cases}$$

- Similarly for (v_n).
- An orthonormal sequence is (e_n), where

$$e_0 = \frac{1}{\sqrt{2\pi}}, \quad e_n(t) = \frac{u_n(t)}{\|u_n\|} = \frac{\cos t}{\sqrt{\pi}}, \quad n = 1,2,\ldots.$$

- From (v_n) we obtain the orthonormal sequence $(\widetilde{e}_n)$, where $\widetilde{e}_n(t) = \frac{v_n(t)}{\|v_n\|} = \frac{\sin nt}{\sqrt{\pi}}$, $n = 1,2,\ldots$.

Determination of the Coefficients

- If $(e_1, e_2, \ldots)$ is an orthonormal sequence in an inner product space X and we have $x \in \text{span}\{e_1, \ldots, e_n\}$, where n is fixed, then by the definition of the span $x = \sum_{k=1}^{n} \alpha_k e_k$.

  If we take the inner product by a fixed e_j, we obtain

  $$\langle x, e_j \rangle = \langle \sum \alpha_k e_k, e_j \rangle = \sum \alpha_k \langle e_k, e_j \rangle = \alpha_j.$$

  With these coefficients, $x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$.
- Another advantage of orthonormality becomes apparent if we want to add another term $\alpha_{n+1} e_{n+1}$ to take care of an

  $$\widetilde{x} = x + \alpha_{n+1} e_{n+1} \in \text{span}\{e_1, \ldots, e_{n+1}\}.$$

  Then we need to calculate only one more coefficient since the other coefficients remain unchanged.

A Perpendicularity Relation

- Consider any $x \in X$, not necessarily in $Y_n = \text{span}\{e_1, \ldots, e_n\}$.
  Define $y \in Y_n$ by setting $y = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$, where n is fixed.
  Then define z by setting $x = y + z$, i.e., $z = x - y$.
  Claim: $z = x - y \perp y$.
  By the orthonormality,

$$\|y\|^2 = \langle \sum \langle x, e_k \rangle e_k, \sum \langle x, e_m \rangle e_m \rangle = \sum |\langle x, e_k \rangle|^2.$$

  Using this, we can now show that $z \perp y$:

$$\begin{array}{rcl} \langle z, y \rangle = \langle x - y, y \rangle & = & \langle x, y \rangle - \langle y, y \rangle \\ & = & \langle x, \sum \langle x, e_k \rangle e_k \rangle - \|y\|^2 \\ & = & \sum \langle x, e_k \rangle \overline{\langle x, e_k \rangle} - \sum |\langle x, e_k \rangle|^2 \\ & = & 0. \end{array}$$

The Bessel Inequality and the Fourier Coefficients

- We showed that, if $y = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$ and $z = x - y$, then $z \perp y$. The Pythagorean relation gives $\|x\|^2 = \|y\|^2 + \|z\|^2$. It follows that

$$\|z\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum |\langle x, e_k \rangle|^2.$$

Since $\|z\| \geq 0$, we have, for every $n = 1, 2, \ldots$, $\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq \|x\|^2$. These sums have nonnegative terms, so that they form a monotone increasing sequence. This sequence converges because it is bounded by $\|x\|^2$. Thus, the infinite series converges.

Theorem (Bessel Inequality)

Let (e_k) be an orthonormal sequence in an inner product space X. Then, for every $x \in X$,

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2. \qquad \textbf{(Bessel Inequality)}.$$

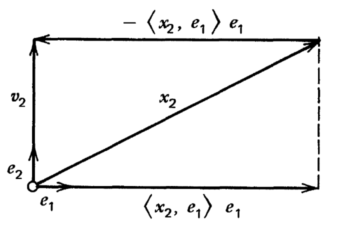
- The inner products $\langle x, e_k \rangle$ are called the **Fourier coefficients** of x with respect to the orthonormal sequence (e_k).

The Gram-Schmidt Orthonormalization Process

- Let (x_j) be a linearly independent sequence in an inner product space. The **Gram-Schmidt orthonormalization process** produces an orthonormal sequence (e_j), such that, for every n, $\text{span}\{e_1,\ldots,e_n\} = \text{span}\{x_1,\ldots,x_n\}$.
  1. The first element of (e_k) is $e_1 = \frac{1}{\|x_1\|} x_1$.
  2. x_2 can be written $x_2 = \langle x_2, e_1\rangle e_1 + v_2$.

     Then $v_2 = x_2 - \langle x_2, e_1\rangle e_1$ is not the zero vector, since (x_j) is linearly independent. Also $v_2 \perp e_1$ since $\langle v_2, e_1\rangle = 0$. We can take, then,

$$e_2 = \frac{1}{\|v_2\|} v_2.$$

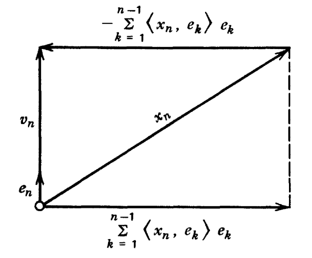


The Gram-Schmidt Orthonormalization Process

3. The vector $v_3 = x_3 - \langle x_3, e_1 \rangle e_1 - \langle x_3, e_2 \rangle e_2$ is not the zero vector. $v_3 \perp e_1$ as well as $v_3 \perp e_2$. We take $e_3 = \frac{1}{\|v_3\|} v_3$.

n. The vector $v_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k$ is not the zero vector. It is orthogonal to $e_1, \ldots, e_{n-1}$. We now obtain

$$e_n = \frac{1}{\|v_n\|} v_n.$$



- In each step, the sum subtracted from x_n is the projection of x_n on $\text{span}\{e_1, \ldots, e_{n-1}\}$. This gives v_n, which is then multiplied by $\frac{1}{\|v_n\|}$ so that we get a vector of norm one.

  v_n cannot be the zero vector for any n: Otherwise, the smallest n for which $v_n = 0$ would give that x_n is a linear combination of $e_1, \ldots, e_{n-1}$. Hence, x_n would be a linear combination of $x_1, \ldots, x_{n-1}$. But $\{x_1, \ldots, x_n\}$ are linearly independent.

Subsection 5

Series Related to Orthonormal Sequences and Sets

Fourier Series

- A **trigonometric series** is a series of the form

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt).$$

- A real-valued function x on $\mathbb{R}$ is said to be **periodic** if there is a positive number p (called a **period** of x), such that $x(t+p) = x(t)$, for all $t \in \mathbb{R}$.
- Let x be of period 2π and continuous. By definition, the **Fourier series** of x is the trigonometric series $a_0 + \sum_{k=1}^{\infty}(a_k \cos kt + b_k \sin kt)$ with coefficients a_k and b_k given by the **Euler formulas**:

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_0^{2\pi} x(t)dt \\ a_k &= \frac{1}{\pi}\int_0^{2\pi} x(t)\cos kt\,dt, \quad k = 1, 2, \ldots \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} x(t)\sin kt\,dt, \quad k = 1, 2, \ldots \end{aligned}$$

These coefficients are called the **Fourier coefficients** of x.

Fourier Series Expansion: An Example

- If the Fourier series of x converges for each t and has the sum $x(t)$, then we write

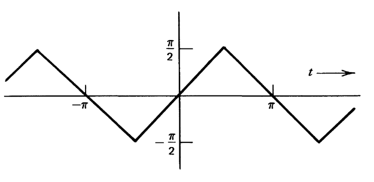
$$x(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt).$$

- Since x is periodic of period 2π, we may replace the interval of integration $[0, 2\pi]$ by any other interval of length 2π.

  Example: Let

$$x(t) = \begin{cases} t, & \text{if } -\frac{\pi}{2} \leq t < \frac{\pi}{2} \\ \pi - t, & \text{if } \frac{\pi}{2} \leq t < \frac{3\pi}{2} \end{cases}$$

  and $x(t + 2\pi) = x(t)$.



Example (Cont'd)

- $x(t) = \begin{cases} t, & \text{if } -\frac{\pi}{2} \leq t < \frac{\pi}{2} \\ \pi - t, & \text{if } \frac{\pi}{2} \leq t < \frac{3\pi}{2} \end{cases}$.

  Since $x(t)$ is odd and $\cos kt$ is even, $a_k = 0$, for $k = 0, 1, \ldots$.

  Choosing $[-\frac{\pi}{2}, \frac{3\pi}{2}]$ as a convenient interval of integration and integrating by parts,

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} t \sin kt dt + \frac{1}{\pi}\int_{\pi/2}^{3\pi/2} (\pi - t)\sin kt dt \\ &= -\frac{1}{\pi k}[t \cos kt] \Big|_{-\pi/2}^{\pi/2} + \frac{1}{\pi k}\int_{-\pi/2}^{\pi/2} \cos kt dt \\ &\quad -\frac{1}{\pi k}[(\pi - t)\cos kt] \Big|_{\pi/2}^{3\pi/2} - \frac{1}{\pi k}\int_{\pi/2}^{3\pi/2} \cos kt dt \\ &= \frac{4}{\pi k^2} \sin \frac{k\pi}{2}, \quad k = 1, 2, \ldots. \end{aligned}$$

  Hence

$$x(t) = \frac{4}{\pi}\left(\sin t - \frac{1}{3^2}\sin 3t + \frac{1}{5^2}\sin 5t \mp \cdots\right).$$

Fourier Series in the General Setting

- We set $u_k(t) = \cos kt$ and $v_k(t) = \sin kt$.
  Hence the series take the form
$$x(t) = a_0 u_0(t) + \sum_{k=1}^{\infty} [a_k u_k(t) + b_k v_k(t)].$$
- We multiply the series by a fixed u_j and integrate over t from 0 to 2π, which amounts to taking the inner product by u_j.
- We assume that term wise integration is permissible (uniform convergence would suffice) and use the orthogonality of (u_k) and (v_k) as well as the fact that $u_j \perp v_k$, for all j, k.
- Then we obtain
$$\begin{array}{rcl}\langle x, u_j \rangle & = & a_0 \langle u_0, u_j \rangle + \sum [a_k \langle u_k, u_j \rangle + b_k \langle v_k, u_j \rangle] \\ & = & a_j \langle u_j, u_j \rangle \\ & = & a_j \|u_j\|^2 = \begin{cases} 2\pi a_0, & \text{if } j = 0 \\ \pi a_j, & \text{if } j = 1, 2, \ldots \end{cases}.\end{array}$$

Fourier Series in the General Setting (Cont'd)

- We found

$$\langle x, u_j \rangle = a_j \|u_j\|^2 = \begin{cases} 2\pi a_0, & \text{if } j = 0 \\ \pi a_j, & \text{if } j = 1, 2, \ldots \end{cases}.$$

- Similarly, if we multiply by v_j and proceed as before, we arrive at

$$\langle x, v_j \rangle = b_j \|v_j\|^2 = \pi b_j.$$

- Solving for a_j and b_j and using the orthonormal sequences (e_j) and $(\widetilde{e}_j)$, where $e_j = \frac{1}{\|u_j\|} u_j$ and $\widetilde{e}_j = \frac{1}{\|v_j\|} v_j$ we obtain

$$\begin{aligned} a_j &= \frac{1}{\|u_j\|^2} \langle x, u_j \rangle = \frac{1}{\|u_j\|} \langle x, e_j \rangle; \\ b_j &= \frac{1}{\|v_j\|^2} \langle x, v_j \rangle = \frac{1}{\|v_j\|} \langle x, \widetilde{e}_j \rangle. \end{aligned}$$

Fourier Series in the General Setting (Cont'd)

- Thus, in the series

$$\begin{aligned} a_k u_k(t) &= \frac{1}{\|u_k\|}\langle x, e_k\rangle u_k(t) = \langle x, e_k\rangle e_k(t); \\ b_k v_k(t) &= \frac{1}{\|v_k\|}\langle x, \widetilde{e}_k\rangle v_k(t) = \langle x, \widetilde{e}_k\rangle \widetilde{e}_k(t). \end{aligned}$$

Hence we may write the Fourier series in the form

$$x = \langle x, e_0\rangle e_0 + \sum_{k=1}^{\infty} [\langle x, e_k\rangle e_k + \langle x, \widetilde{e}_k\rangle \widetilde{e}_k].$$

This justifies using the term "Fourier coefficients".

Convergence of a Series in a Hilbert Space

- Given any orthonormal sequence (e_k) in a Hilbert space H, we may consider series of the form $\sum_{k=1}^{\infty} \alpha_k e_k$, where $\alpha_1, \alpha_2, \ldots$ are any scalars.
- Such a series **converges** and has the **sum** s if there exists an $s \in H$, such that the sequence (s_n) of the partial sums $s_n = \alpha_1 e_1 + \cdots \alpha_n e_n$ converges to s, that is, $\|s_n - s\| \xrightarrow{n\to\infty} 0$.

Theorem (Convergence)

Let (e_k) be an orthonormal sequence in a Hilbert space H. Then:

(a) The series converges (in the norm on H) if and only if $\sum_{k=1}^{\infty} |\alpha_k|^2$ converges.

(b) If the series converges, then the coefficients α_k are the Fourier coefficients $\langle x, e_k \rangle$, where x is the sum of $\sum_{k=1}^{\infty} \alpha_k e_k$. In this case, $x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$.

(c) For any $x \in H$, the series $\sum_{k=1}^{\infty} \alpha_k e_k$, with $\alpha_k = \langle x, e_k \rangle$ converges (in the norm of H).

Proof of the Convergence Theorem

(a) Let $s_n = \alpha_1 e_1 + \cdots + \alpha_n e_n$ and $\sigma_n = |\alpha_1|^2 + \cdots + |\alpha_n|^2$. Then, because of the orthonormality, for any m and $n > m$,

$$\begin{aligned} \|s_n - s_m\|^2 &= \|\alpha_{m+1} e_{m+1} + \cdots + \alpha_n e_n\|^2 \\ &= |\alpha_{m+1}|^2 + \cdots + |\alpha_n|^2 = \sigma_n - \sigma_m. \end{aligned}$$

Hence (s_n) is Cauchy in H if and only if (σ_n) is Cauchy in $\mathbb{R}$. Since H and $\mathbb{R}$ are complete, the first statement of the theorem follows.

(b) Taking the inner product of s_n and e_j and using the orthonormality, we have $\langle s_n, e_j \rangle = \alpha_j$, for $j = 1, \ldots, k$, $k \leq n$ fixed. By assumption, $s_n \to x$. Since the inner product is continuous, $\alpha_j = \langle s_n, e_j \rangle \to \langle x, e_j \rangle$, $j \leq k$. Here we can take $k (\leq n)$ as large as we please because $n \to \infty$. So we have $\alpha_j = \langle x, e_j \rangle$, for every $j = 1, 2, \ldots$.

(c) From the Bessel inequality, we see that the series $\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2$ converges. From this and Part (a), we conclude that (c) must hold.

Countability of Nonzero Fourier Coefficients

Lemma (Fourier Coefficients)

Any x in an inner product space X can have at most countably many nonzero Fourier coefficients $\langle x, e_\kappa \rangle$ with respect to an orthonormal family (e_κ), $\kappa \in I$, in X.

- If an orthonormal family (e_κ), $\kappa \in I$, in an inner product space X is uncountable (since the index set I is uncountable), we can still form the Fourier coefficients $\langle x, e_\kappa \rangle$ of an $x \in X$, where $\kappa \in I$. Now we use $\sum_{\kappa \in I} |\langle x, e_\kappa \rangle|^2 \leq \|x\|^2$ to conclude that, for each fixed $m = 1, 2, \ldots$, the number of Fourier coefficients such that $|\langle x, e_\kappa \rangle| > \frac{1}{m}$ must be finite.
- Hence with any fixed $x \in H$ we can associate a series $\sum_{\kappa \in I} \langle x, e_\kappa \rangle e_\kappa$.

  We can arrange the e_κ with $\langle x, e_\kappa \rangle \neq 0$ in a sequence $(e_1, e_2, \ldots)$, so as to get the series $\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$.

  Convergence follows from the Convergence Theorem.

Order-Independence of the Sum

- We show that the sum does not depend on the order in which the e_κ are arranged in a sequence.
- Let (w_m) be a rearrangement of (e_n). By definition this means that there is a bijective mapping $n \mapsto m(n)$ of $\mathbb{N}$ onto itself such that corresponding terms of the two sequences are equal, i.e., $w_{m(n)} = e_n$.
- Let $\alpha_n = \langle x, e_n \rangle$, $\beta_m = \langle x, w_m \rangle$, $x_1 = \sum_{n=1}^\infty \alpha_n e_n$, $x_2 = \sum_{m=1}^\infty \beta_m w_m$.
- Then by Part (b) of the Convergence Theorem, $\alpha_n = \langle x, e_n \rangle = \langle x_1, e_n \rangle$, $\beta_m = \langle x, w_m \rangle = \langle x_2, w_m \rangle$. But $e_n = w_{m(n)}$. Hence, $\langle x_1 - x_2, e_n \rangle = \langle x_1, e_n \rangle - \langle x_2, w_{m(n)} \rangle = \langle x, e_n \rangle - \langle x, w_{m(n)} \rangle = 0$. Similarly, $\langle x_1 - x_2, w_m \rangle = 0$. This implies

$$\begin{aligned} \|x_1 - x_2\|^2 &= \langle x_1 - x_2, \sum \alpha_n e_n - \sum \beta_m w_m \rangle \\ &= \sum \overline{\alpha}_n \langle x_1 - x_2, e_n \rangle - \sum \overline{\beta}_m \langle x_1 - x_2, w_m \rangle = 0. \end{aligned}$$

Consequently, $x_1 - x_2 = 0$ and $x_1 = x_2$. Since the rearrangement (w_m) of (e_n) was arbitrary, this completes the proof.

Subsection 6

Total Orthonormal Sets and Sequences

Total Orthonormal Sets

Definition (Total Orthonormal Set)

A **total set** (or **fundamental set**) in a normed space X is a subset $M \subseteq X$ whose span is dense in X. Accordingly, an orthonormal set (or sequence or family) in an inner product space X which is total in X is called a **total orthonormal set** (or sequence or family, respectively) in X.

- M is total in X if and only if $\overline{\text{span} M} = X$.
- A total orthonormal family in X is sometimes called an **orthonormal basis** for X.

  This is not a basis for X as a vector space, in the sense of algebra, unless X is finite dimensional.

Existence and Cardinality

- In every Hilbert space $H \neq \{0\}$, there exists a total orthonormal set.
  - For a finite dimensional H this is clear.
  - For an infinite dimensional separable H, it follows from the Gram-Schmidt process by (ordinary) induction.
  - For a nonseparable H a (nonconstructive) proof requires Zorn's lemma.
- All total orthonormal sets in a given Hilbert space $H \neq \{0\}$ have the same cardinality. The latter is called the **Hilbert dimension** or **orthogonal dimension** of H. (If $H = \{0\}$, this dimension is defined to be 0.)
  - For a finite dimensional H the statement is clear since then the Hilbert dimension is the dimension in the sense of algebra.
  - For an infinite dimensional separable H the statement will be proven shortly.
  - For a general H the proof would require somewhat more advanced tools from set theory.

The Totality Theorem

- A total orthonormal set cannot be augmented to a more extensive orthonormal set by the adjunction of new elements.

Theorem (Totality)

Let M be a subset of an inner product space X. Then:

(a) If M is total in X, then there does not exist a nonzero $x \in X$ which is orthogonal to every element of M; briefly, $x \perp M$ implies $x = 0$.

(b) If X is complete, that condition is also sufficient for the totality of M in X.

(a) Let H be the completion of X. Then X, regarded as a subspace of H, is dense in H. By assumption, M is total in X. So $\text{span} M$ is dense in X. Hence, it is dense in H. Thus, the orthogonal complement of M in H is $\{0\}$. A fortiori, if $x \in X$ and $x \perp M$, then $x = 0$.

(b) If X is a Hilbert space and M satisfies that condition, so that $M^{\perp} = \{0\}$, then M is total in X.

- The completeness of X in (b) is essential. If X is not complete, there may not exist an orthonormal set $M \subseteq X$, such that M is total in X.

Parseval's Relation

- Consider any given orthonormal set M in a Hilbert space H.
- We know that each fixed $x \in H$ has at most countably many nonzero Fourier coefficients. So we can arrange these coefficients in a sequence, say, $\langle x, e_1 \rangle, \langle x, e_2 \rangle, \ldots$.
- The Bessel inequality is

$$\sum_k |\langle x, e_k \rangle|^2 \leq \|x\|^2 \quad (\textbf{Bessel Inequality})$$

  where the left-hand side is an infinite series or a finite sum.
- With the equality sign this becomes

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2 \quad (\textbf{Parseval Relation}).$$

- Parseval's Relation helps to characterize totality.

A Second Totality Criterion

Theorem (Totality)

An orthonormal set M in a Hilbert space H is total in H if and only if, for all $x \in H$, the Parseval relation

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2$$

holds, where the summation is over all nonzero Fourier coefficients of x with respect to M.

(a) If M is not total, there is a nonzero $x \perp M$ in H. Since $x \perp M$, $\langle x, e_k \rangle = 0$, for all k. Thus, the left-hand side is zero, whereas $\|x\|^2 \neq 0$. This shows that Parseval's relation does not hold.

Hence if Parseval's relation holds for all $x \in H$, then M must be total in H.

Proof of the Totality Criterion

(b) Conversely, assume M to be total in H. Consider any $x \in H$ and its nonzero Fourier coefficients arranged in a sequence $\langle x, e_1 \rangle, \langle x, e_2 \rangle, \ldots$, or in some definite order if there are only finitely many of them. Define y by $y = \sum_k \langle x, e_k \rangle e_k$, noting that in the case of an infinite series, convergence follows from the convergence theorem.
We show that $x - y \perp M$: For every e_j occurring in the sum y, we have, using orthonormality,

$$\langle x - y, e_j \rangle = \langle x, e_j \rangle - \sum_k \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0.$$

And for every $v \in M$ not contained in y, we have $\langle x, v \rangle = 0$. So

$$\langle x - y, v \rangle = \langle x, v \rangle - \sum_k \langle x, e_k \rangle \langle e_k, v \rangle = 0 - 0 = 0.$$

Hence $x - y \perp M$, that is, $x - y \in M^\perp$. Since M is total in H, we have $M^\perp = \{0\}$. Together, $x - y = 0$, i.e., $x = y$. Thus, $\|x\|^2 = \langle \sum_k \langle x, e_k \rangle e_k, \sum_m \langle x, e_m \rangle e_m \rangle = \sum_k \langle x, e_k \rangle \overline{\langle x, e_k \rangle}$.

Separable Hilbert Spaces

Theorem (Separable Hilbert Spaces)

Let H be a Hilbert space. Then:

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

(b) If H contains an orthonormal sequence which is total in H, then H is separable.

(a) Let H be separable, B any dense set in H and M any orthonormal set. Then, for any $x \neq y$ in M, $\|x-y\|^2 = \langle x-y, x-y \rangle = \langle x, x \rangle + \langle y, y \rangle = 2$. Hence spherical neighborhoods N_x of x and N_y of y of radius $\frac{\sqrt{2}}{3}$ are disjoint.

Since B is dense in H, there is a $b \in B$ in N_x and a $\widetilde{b} \in B$ in N_y and $b \neq \widetilde{b}$, since $N_x \cap N_y = \varnothing$. Hence, if M were uncountable, we would have uncountably many such pairwise disjoint spherical neighborhoods, so that B would be uncountable. Thus, H would not contain a dense set which is countable, contradicting separability.

Separable Hilbert Spaces (The Converse)

(b) Let (e_k) be a total orthonormal sequence in H. Let A be the set of all linear combinations $\gamma_1^{(n)}e_1+\cdots+\gamma_n^{(n)}e_n$, $n=1,2,\ldots$, where $\gamma_k^{(n)}=\alpha_k^{(n)}+ib_k^{(n)}$ and $a_k^{(n)}$ and $b_k^{(n)}$ are rational ($b_k^{(n)}=0$ if H is real). Clearly, A is countable. We prove that A is dense in H by showing that, for every $x\in H$ and $\varepsilon>0$, there is a $v\in A$, such that $\|x-v\|<\varepsilon$.

Since (e_k) is total in H, there is an n, such that $Y_n=\text{span}\{e_1,\ldots,e_n\}$ contains a point whose distance from x is less than $\frac{\varepsilon}{2}$. In particular, $\|x-y\|<\frac{\varepsilon}{2}$, for the orthogonal projection y of x on Y_n, which is given by $y=\sum_{k=1}^n\langle x,e_k\rangle e_k$. Hence, $\|x-\sum_{k=1}^n\langle x,e_k\rangle e_k\|<\frac{\varepsilon}{2}$. Since the rationals are dense in $\mathbb{R}$, for each $\langle x,e_k\rangle$, there is a $\gamma_k^{(n)}$ (with rational real and imaginary parts) s.t. $\|\sum_{k=1}^n[\langle x,e_k\rangle-\gamma_k^{(n)}]e_k\|<\frac{\varepsilon}{2}$. Hence $v\in A$ defined by $v=\sum_{k=1}^n\gamma_k^{(n)}e_k$ satisfies $\|x-v\|=\|x-\sum\gamma_k^{(n)}e_k\|\leq \|x-\sum\langle x,e_k\rangle e_k\|+\|\sum\langle x,e_k\rangle e_k-\sum\gamma_k^{(n)}e_k\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$. This proves that A is dense in H. Since A is countable, H is separable.

Isomorphisms of Hilbert Spaces

- An **isomorphism** of a Hilbert space H onto a Hilbert space $\widetilde{H}$ over the same field is a bijective linear operator $T : H \to \widetilde{H}$, such that for all $x, y \in H$,
$$\langle Tx, Ty \rangle = \langle x, y \rangle.$$
  H and $\widetilde{H}$ are then called **isomorphic Hilbert spaces**.
- Since T is linear, it preserves the vector space structure, and the displayed condition shows that T is isometric.

  From this and the bijectivity of T it follows that H and $\widetilde{H}$ are algebraically as well as metrically indistinguishable.

  They are essentially the same, except for the nature of their elements, so that we may think of $\widetilde{H}$ as being essentially H with a "tag" T attached to each vector x.

Isomorphism and Hilbert Dimension

Theorem (Isomorphism and Hilbert Dimension)

Two Hilbert spaces H and $\widetilde{H}$, both real or both complex, are isomorphic if and only if they have the same Hilbert dimension.

(a) If H is isomorphic with $\widetilde{H}$ and $T: H \to \widetilde{H}$ is an isomorphism, then orthonormal elements in H have orthonormal images under T. Since T is bijective, we thus conclude that T maps every total orthonormal set in H onto a total orthonormal set in $\widetilde{H}$. Hence H and $\widetilde{H}$ have the same Hilbert dimension.

(b) Conversely, suppose that H and $\widetilde{H}$ have the same Hilbert dimension.

The case $H = \{0\}$ and $\widetilde{H} = \{0\}$ is trivial.

Let $H \neq \{0\}$. Then $\widetilde{H} \neq \{0\}$, and any total orthonormal sets M in H and $\widetilde{M}$ in $\widetilde{H}$ have the same cardinality. So we can index them by the same index set $\{k\}$ and write $M = (e_k)$ and $\widetilde{M} = (\widetilde{e}_k)$. To show that H and $\widetilde{H}$ are isomorphic, we construct an isomorphism of H onto $\widetilde{H}$.

Isomorphism and Hilbert Dimension (Cont'd)

- For every $x \in H$, we have $x = \sum_k \langle x, e_k \rangle e_k$, where the right-hand side is a finite sum or an infinite series, and $\sum_k |\langle x, e_k \rangle|^2 < \infty$, by the Bessel inequality. Define

$$\widetilde{x} = Tx = \sum_k \langle x, e_k \rangle \widetilde{e}_k.$$

  We have convergence by the Convergence Theorem. So $\widetilde{x} \in \widetilde{H}$.
  - The operator T is linear since the inner product is linear with respect to the first factor.
  - T is isometric: $\|\widetilde{x}\|^2 = \|Tx\|^2 = \sum_k |\langle x, e_k \rangle|^2 = \|x\|^2$. Thus, T preserves the inner product.
  - Injectivity: If $Tx = Ty$, then $\|x - y\| = \|T(x - y)\| = \|Tx - Ty\| = 0$. So $x = y$ and T is injective.
  - T is surjective: Given any $\widetilde{x} = \sum_k \alpha_k \widetilde{e}_k$ in $\widetilde{H}$, we have $\sum |\alpha_k|^2 < \infty$ by the Bessel inequality. Hence $\sum_k \alpha_k e_k$ is a finite sum or a series which converges to an $x \in H$, and $\alpha_k = \langle x, e_k \rangle$ by the same theorem. We thus have $\widetilde{x} = Tx$ by definition of T.

Subsection 7

Representation of Functionals on Hilbert Spaces

Functionals on Hilbert Spaces

Riesz's Theorem (Functionals on Hilbert Spaces)

Every bounded linear functional f on a Hilbert space H can be represented in terms of the inner product, namely,

$$f(x) = \langle x, z \rangle,$$

where z depends on f, is uniquely determined by f and has norm $\|z\| = \|f\|$.

- We prove that:
  (a) f has the required representation;
  (b) z in the representation is unique;
  (c) $\|z\| = \|f\|$.

(a) If $f = 0$, then we take $z = 0$. Let $f \neq 0$. We look at properties z must have if a representation exists:
  - First, $z \neq 0$, since otherwise $f = 0$.
  - Second, $\langle x, z \rangle = 0$, for all x for which $f(x) = 0$, i.e., for all x in the null space $\mathscr{N}(f)$ of f. Hence, $z \perp \mathscr{N}(f)$. This suggests that we consider $\mathscr{N}(f)$ and its orthogonal complement $\mathscr{N}(f)^{\perp}$.

Riesz's Theorem: Part (a)

- $\mathcal{N}(f)$ is a vector space and is closed. Furthermore, $f \neq 0$ implies $\mathcal{N}(f) \neq H$, so that $\mathcal{N}(f)^\perp \neq \{0\}$ by the projection theorem. Hence $\mathcal{N}(f)^\perp$ contains a $z_0 \neq 0$. For an arbitrary $x \in H$, set

$$v = f(x)z_0 - f(z_0)x.$$

Then $f(v) = f(x)f(z_0) - f(z_0)f(x) = 0$. This shows that $v \in \mathcal{N}(f)$. Since $z_0 \perp \mathcal{N}(f)$, we have

$$0 = \langle v, z_0 \rangle = \langle f(x)z_0 - f(z_0)x, z_0 \rangle = f(x)\langle z_0, z_0 \rangle - f(z_0)\langle x, z_0 \rangle.$$

Noting that $\langle z_0, z_0 \rangle = \|z_0\|^2 \neq 0$, we can solve for $f(x)$. The result is $f(x) = \frac{f(z_0)}{\langle z_0, z_0 \rangle}\langle x, z_0 \rangle = \left\langle x, \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0 \right\rangle$. So, we take $z = \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0$.

Riesz's Theorem: Parts (b) and (c)

(b) We prove that z is unique: Suppose that for all $x \in H$,

$$f(x) = \langle x, z_1 \rangle = \langle x, z_2 \rangle.$$

Then $\langle x, z_1 - z_2 \rangle = 0$, for all x. Choosing $x = z_1 - z_2$, we have

$$\langle x, z_1 - z_2 \rangle = \langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2 = 0.$$

Hence $z_1 - z_2 = 0$, so that $z_1 = z_2$.

(c) We finally prove $\|z\| = \|f\|$. If $f = 0$, then $z = 0$ and the property holds. Let $f \neq 0$. Then $z \neq 0$.

- We get $\|z\|^2 = \langle z, z \rangle = f(z) \leq \|f\| \|z\|$. Division by $\|z\| \neq 0$ yields $\|z\| \leq \|f\|$.
- It remains to show that $\|f\| \leq \|z\|$. By the Schwarz inequality, we see that $|f(x)| = |\langle x, z \rangle| \leq \|x\| \|z\|$. This implies $\|f\| = \sup_{\|x\|=1} |\langle x, z \rangle| \leq \|z\|$.

The Equality Lemma

Lemma (Equality)

If $\langle v_1, w\rangle = \langle v_2, w\rangle$, for all w in an inner product space X, then $v_1 = v_2$. In particular, $\langle v_1, w\rangle = 0$, for all $w \in X$, implies $v_1 = 0$.

- By assumption, for all w,

$$\langle v_1 - v_2, w\rangle = \langle v_1, w\rangle - \langle v_2, w\rangle = 0.$$

For $w = v_1 - v_2$, we get $\|v_1 - v_2\|^2 = 0$. Hence $v_1 - v_2 = 0$. So $v_1 = v_2$. In particular, $\langle v_1, w\rangle = 0$, with $w = v_1$, gives $\|v_1\|^2 = 0$. So $v_1 = 0$.

Sesquilinear Forms over Vector Spaces

Definition (Sesquilinear Form)

Let X and Y be vector spaces over the same field K ($=\mathbb{R}$ or $\mathbb{C}$). Then a **sesquilinear form** (or **sesquilinear functional**) h on $X \times Y$ is a mapping $h : X \times Y \to K$, such that, for all $x, x_1, x_2 \in X$ and $y, y_1, y_2 \in Y$ and all scalars α, β,

(a) $h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y)$;

(b) $h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2)$;

(c) $h(\alpha x, y) = \alpha h(x, y)$;

(d) $h(x, \beta y) = \overline{\beta} h(x, y)$.

Hence h is **linear** in the first argument and **conjugate linear** in the second one.

If X and Y are real ($K = \mathbb{R}$), then (d) is simply $h(x, \beta y) = \beta h(x, y)$ and h is called **bilinear** since it is linear in both arguments.

Bounded Sesquilinear Forms over Normed Spaces

Definition (Bounded Sesquilinear Form)

If X and Y are normed spaces and if there is a real number c, such that for all x, y, $|h(x,y)| \leq c\|x\|\|y\|$, then h is said to be **bounded**, and the number

$$\|h\| = \sup_{\substack{x \in X - \{0\} \\ y \in Y - \{0\}}} \frac{|h(x,y)|}{\|x\|\|y\|} = \sup_{\substack{\|x\|=1 \\ \|y\|=1}} |h(x,y)|$$

is called the **norm** of h.

Example: The inner product is sesquilinear and bounded.

By the conditions in the definition, we have

$$|h(x,y)| \leq \|h\|\|x\|\|y\|.$$

Riesz Representation

Theorem (Riesz Representation)

Let H_1, H_2 be Hilbert spaces and $h : H_1 \times H_2 \to K$ a bounded sesquilinear form. Then h has a representation

$$h(x,y) = \langle Sx, y \rangle,$$

where $S : H_1 \to H_2$ is a bounded linear operator. S is uniquely determined by h and has norm $\|S\| = \|h\|$.

- We consider $\overline{h(x,y)}$. This is linear in y, because of the bar. To exploit the preceding representation, we keep x fixed. Then, we get a representation in which y is variable, say, $\overline{h(x,y)} = \langle y, z \rangle$. Hence, $h(x,y) = \langle z, y \rangle$. Here $z \in H_2$ is unique but, of course, depends on our fixed $x \in H_1$. It follows that this representation with variable x defines an operator given by $z = Sx$. Substituting $z = Sx$, $h(x,y) = \langle Sx, y \rangle$.

Linearity of S

- S is linear:

  In fact, its domain is the vector space H_1. From sesquilinearity, we obtain, for all y in H_2,

$$\begin{array}{rcl} \langle S(\alpha x_1 + \beta x_2), y\rangle & = & h(\alpha x_1 + \beta x_2, y) \\ & = & \alpha h(x_1, y) + \beta h(x_2, y) \\ & = & a\langle Sx_1, y\rangle + \beta\langle Sx_2, y\rangle \\ & = & \langle \alpha Sx_1 + \beta Sx_2, y\rangle. \end{array}$$

  So by the Equality Lemma,

$$S(\alpha x_1 + \beta x_2) = \alpha Sx_1 + \beta Sx_2.$$

Boundedness and Uniqueness of S

- S is bounded: Leaving aside the trivial case $S = 0$, we have

$$\|h\| = \sup_{\substack{x\neq 0\\ y\neq 0}} \frac{|\langle Sx, y\rangle|}{\|x\|\|y\|} \geq \sup_{\substack{x\neq 0\\ Sx\neq 0}} \frac{|\langle Sx, Sx\rangle|}{\|x\|\|Sx\|} = \sup_{x\neq 0} \frac{\|Sx\|}{\|x\|} = \|S\|.$$

  This proves boundedness. Moreover, $\|h\| \geq \|S\|$. But, note that

$$\|h\| = \sup_{\substack{x\neq 0\\ y\neq 0}} \frac{\|\langle Sx, y\rangle\|}{\|x\|\|y\|} \leq \sup_{x\neq 0} \frac{\|Sx\|\|y\|}{\|x\|\|y\|} = \|S\|.$$

  Now, we obtain $\|S\| = \|h\|$.
- S is unique: Assume that there is a linear operator $T : H_1 \to H_2$, such that, for all $x \in H_1$ and $y \in H_2$, we have

$$h(x, y) = (Sx, y) = (Tx, y).$$

  Then $Sx = Tx$, for all $x \in H_1$, by the Equality Lemma. Hence $S = T$, by definition.

Subsection 8

Hilbert-Adjoint Operator

Hilbert-Adjoint Operators

Definition (Hilbert-Adjoint Operator T^*)

Let $T : H_1 \to H_2$ be a bounded linear operator, where H_1 and H_2 are Hilbert spaces. Then the **Hilbert-adjoint operator** T^* of T is the operator

$$T^* : H_2 \to H_1$$

such that, for all $x \in H_1$ and $y \in H_2$,

$$\langle Tx, y \rangle = \langle x, T^* y \rangle.$$

Existence of Hilbert-Adjoint Operators

Theorem (Existence)

The Hilbert-adjoint operator T^* of a bounded linear operator T exists, is unique and is a bounded linear operator with norm $\|T^*\| = \|T\|$.

- The formula $h(y,x) = \langle y, Tx\rangle$ defines a sesquilinear form on $H_2 \times H_1$ because the inner product is sesquilinear and T is linear. In fact, conjugate linearity of the form is seen from

$$\begin{aligned} h(y, \alpha x_1 + \beta x_2) &= \langle y, T(\alpha x_1 + \beta x_2)\rangle \\ &= \langle y, \alpha Tx_1 + \beta Tx_2\rangle \\ &= \overline{\alpha}\langle y, Tx_1\rangle + \overline{\beta}\langle y, Tx_2\rangle \\ &= \overline{\alpha} h(y, x_1) + \overline{\beta} h(y, x_2). \end{aligned}$$

.

Existence of Hilbert-Adjoint Operators (Cont'd)

- h is bounded: Indeed, by the Schwarz inequality,

$$|h(y,x)| = |\langle y, Tx\rangle| \le \|y\|\|Tx\| \le \|T\|\|x\|\|y\|.$$

This also implies $\|h\| \le \|T\|$. Moreover

$$\|h\| = \sup_{\substack{x\neq 0\\ y\neq 0}} \frac{|\langle y, Tx\rangle|}{\|y\|\|x\|} \ge \sup_{\substack{x\neq 0\\ Tx\neq 0}} \frac{|\langle Tx, Tx\rangle|}{\|Tx\|\|x\|} = \|T\|,$$

whence $\|h\| \ge \|T\|$. Together, $\|h\| = \|T\|$.

We obtain a Riesz representation $h(y,x) = \langle T^* y, x\rangle$. We know that $T^* : H_2 \to H_1$ is a uniquely determined bounded linear operator with norm $\|T^*\| = \|h\| = \|T\|$.

Also $\langle y, Tx\rangle = \langle T^* y, x\rangle$. By taking conjugates, $\langle Tx, y\rangle = \langle x, T^* y\rangle$.

The Zero Operator Lemma

Lemma (Zero Operator)

Let X and Y be inner product spaces and $Q : X \to Y$ a bounded linear operator. Then:

(a) $Q = 0$ if and only if $\langle Qx, y \rangle = 0$, for all $x \in X$ and $y \in Y$.

(b) If $Q : X \to X$, where X is complex, and $\langle Qx, x \rangle = 0$, for all $x \in X$, then $Q = 0$.

(a) $Q = 0$ means $Qx = 0$, for all x and implies $\langle Qx, y \rangle = \langle 0, y \rangle = 0\langle w, y \rangle = 0$. Conversely, $\langle Qx, y \rangle = 0$, for all x and y implies $Qx = 0$, for all x, by Equality, so that $Q = 0$, by definition.

(b) By assumption, $\langle Qv, v \rangle = 0$, for every $v = \alpha x + y \in X$, that is, $0 = \langle Q(\alpha x + y), \alpha x + y \rangle = |\alpha|^2 \langle Qx, x \rangle + \langle Qy, y \rangle + \alpha \langle Qx, y \rangle + \overline{\alpha} \langle Qy, x \rangle$. The first two terms on the right are zero by assumption. $\alpha = 1$ gives $\langle Qx, y \rangle + \langle Qy, x \rangle = 0$. $\alpha = i$ gives $\overline{\alpha} = -i$ and $\langle Qx, y \rangle - \langle Qy, x \rangle = 0$. By addition, $\langle Qx, y \rangle = 0$, and $Q = 0$ follows from (a).

Remark on the Zero Operator Lemma

- In part (b), it is essential that X be complex: Indeed, the conclusion may not hold if X is real.

  A counterexample is a rotation Q of the plane $\mathbb{R}^2$ through a right angle:
  - Q is linear, and $Qx \perp x$, hence $\langle Qx, x \rangle = 0$, for all $x \in \mathbb{R}^2$;
  - $Q \neq 0$.

Properties of Hilbert-Adjoint Operators

Theorem (Properties of Hilbert-Adjoint Operators)

Let H_1, H_2 be Hilbert spaces, $S : H_1 \to H_2$ and $T : H_1 \to H_2$ bounded linear operators and α any scalar. Then we have:

(a) $\langle T^* y, x\rangle = \langle y, Tx\rangle$, $x \in H_1, y \in H_2$;

(b) $(S + T)^* = S^* + T^*$;

(c) $(\alpha T)^* = \overline{\alpha} T^*$;

(d) $(T^*)^* = T$;

(e) $\| T^* T\| = \| TT^*\| = \| T\|^2$;

(f) $T^* T = 0$ iff $T = 0$;

(g) $(ST)^* = T^* S^*$ (Assuming $H_2 = H_1$).

(a) We have

$$\langle T^* y, x\rangle = \overline{\langle x, T^* y\rangle} = \overline{\langle Tx, y\rangle} = \langle y, Tx\rangle.$$

Property (b)

(b) For all x and y,

$$\begin{aligned} \langle x,(S+T)^* y\rangle &= \langle (S+T)x,y\rangle \\ &= \langle Sx,y\rangle + \langle Tx,y\rangle \\ &= \langle x,S^* y\rangle + \langle x,T^* y\rangle \\ &= \langle x,(S^*+T^*)y\rangle. \end{aligned}$$

Hence, for all y,

$$(S+T)^* y = (S^*+T^*)y.$$

Properties (c)-(d)

(c) Calculate

$$\begin{array}{rcl}\langle(\alpha T)^*y,x\rangle & = & \langle y,(\alpha T)x\rangle=\langle y,\alpha(Tx)\rangle \\ & = & \overline{\alpha}\langle y,Tx\rangle=\overline{\alpha}\langle T^*y,x\rangle \\ & = & \langle\overline{\alpha}T^*y,x\rangle.\end{array}$$

Now apply the Zero Operator Lemma with $Q=(\alpha T)^*-\overline{\alpha}T^*$.

(d) $(T^*)^*$ is written T^{**} and equals T, since, for all $x\in H_1$ and $y\in H_2$, we have

$$\langle(T^*)^*x,y\rangle=\langle x,T^*y\rangle=\langle Tx,y\rangle.$$

So (d) follows from the Zero Operator Lemma with $Q=(T^*)^*-T$.

Properties (e)-(g)

(e) We see that $T^*T : H_1 \to H_1$ but $TT^* : H_2 \to H_2$. By the Schwarz inequality,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \|T^*Tx\|\|x\| \leq \|T^*T\|\|x\|^2.$$

Taking the supremum over all x of norm 1, we obtain $\|T\|^2 \leq \|T^*T\|$. So $\|T\|^2 \leq \|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$, giving $\|T^*T\| = \|T\|^2$. Replacing T by T^*, we get $\|TT^*\| = \|T^{**}T^*\| = \|T^*\|^2 = \|T\|^2$.

(f) From (e), we immediately obtain (f).

(g) We have

$$\langle x, (ST)^*y\rangle = \langle (ST)x, y\rangle = \langle Tx, S^*y\rangle = \langle x, T^*S^*y\rangle.$$

Hence $(ST)^*y = T^*S^*y$, by Equality.

Subsection 9

Self-Adjoint, Unitary and Normal Operators

Self-adjoint, Unitary and Normal Operators

Definition (Self-adjoint, Unitary and Normal Operators)

A bounded linear operator $T: H \to H$ on a Hilbert space H is said to be:

- **self-adjoint** or **Hermitian** if $T^* = T$;
- **unitary** if T is bijective and $T^* = T^{-1}$;
- **normal** if $TT^* = T^*T$.

- The *Hilbert-adjoint* T^* of T was defined by $\langle Tx, y \rangle = \langle x, T^*y \rangle$.

  If T is self-adjoint, the formula becomes $\langle Tx, y \rangle = \langle x, Ty \rangle$.
- If T is self-adjoint or unitary, T is normal.
- A normal operator need not be self-adjoint or unitary:

  Example: If $I: H \to H$ is the identity operator, then:
  - $T = 2iI$ is normal since $T^* = -2iI$, so that $TT^* = T^*T = 4I$;
  - $T^* \neq T$ and $T^* \neq T^{-1} = -\frac{1}{2}iI$.

Matrices of Hilbert-Adjoint Operators

- Consider $\mathbb{C}^n$ with the inner product $\langle x, y\rangle = x^\top \overline{y}$, where x and y are written as column vectors, and $^\top$ means transposition.

  Thus, $x^\top = (\xi_1, \ldots, \xi_n)$, and we use the ordinary matrix multiplication.

  Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator (which is bounded). A basis for $\mathbb{C}^n$ being given, we can represent T and its Hilbert-adjoint operator T^* by two n-rowed square matrices, say, A and B, respectively.

  Using the inner prodct and the familiar rule $(Bx)^\top = x^\top B^\top$, we obtain

  $$\langle Tx, y\rangle = (Ax)^\top \overline{y} = x^\top A^\top \overline{y} \quad \text{and} \quad \langle x, T^* y\rangle = x^\top \overline{B}\overline{y}.$$

  By definition, the left-hand sides are equal for all $x, y \in \mathbb{C}^n$. Hence we must have $A^\top = \overline{B}$. Consequently, $B = \overline{A}^\top$.

  If a basis for $\mathbb{C}^n$ is given and a linear operator on $\mathbb{C}^n$ is represented by a certain matrix, then its *Hilbert-adjoint operator is represented by the complex conjugate transpose of that matrix*.

Types of Matrices

Definition

A square matrix $A = (\alpha_{jk})$ is said to be:

- **Hermitian** if $\overline{A}^\top = A$ ($\overline{\alpha}_{kj} = \alpha_{jk}$);
- **skew-Hermitian** if $\overline{A}^\top = -A$ ($\overline{\alpha}_{kj} = -\alpha_{jk}$);
- **unitary** if $\overline{A}^\top = A^{-1}$;
- **normal** if $A\overline{A}^\top = \overline{A}^\top A$.

A real square matrix $A = (\alpha_{jk})$ is said to be:

- **(real) symmetric** if $A^\top = A$ ($\alpha_{kj} = \alpha_{jk}$);
- **(real) skew-symmetric** if $A^\top = -A$ ($\alpha_{kj} = -\alpha_{jk}$);
- **orthogonal** if $A^\top = A^{-1}$.

- Hence:
  - a real Hermitian matrix is a (real) symmetric matrix;
  - a real skew-Hermitian matrix is a (real) skew-symmetric matrix;
  - a real unitary matrix is an orthogonal matrix.

Properties of Operators and Representing Matrices

- Representing matrices are:
  - **Hermitian** if T is self-adjoint (Hermitian);
  - **unitary** if T is unitary;
  - **normal** if T is normal.
- For a linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$, representing matrices are:
  - **Real symmetric** if T is self-adjoint;
  - **orthogonal** if T is unitary.

Criterion for Self-Adjointness

Theorem (Self-Adjointness)

Let $T : H \to H$ be a bounded linear operator on a Hilbert space H. Then:

(a) If T is self-adjoint, $\langle Tx, x\rangle$ is real for all $x \in H$.

(b) If H is complex and $\langle Tx, x\rangle$ is real for all $x \in H$, the operator T is self-adjoint.

(a) If T is self-adjoint, then for all x, $\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle Tx, x\rangle$. Hence, $\langle Tx, x\rangle$ is equal to its complex conjugate, so that it is real.

(b) If $\langle Tx, x\rangle$ is real for all x, then $\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \overline{\langle x, T^* x\rangle} = \langle T^* x, x\rangle$. Hence, $0 = \langle Tx, x\rangle - \langle T^* x, x\rangle = \langle (T - T^*)x, x\rangle$ and $T - T^* = 0$ by a preceding lemma, since H is complex.

- In Part (b) it is essential that H be complex: For a real H the inner product is real-valued, which makes $\langle Tx, x\rangle$ real without any further assumptions about T.

Self-Adjointness of Product

Theorem (Self-Adjointness of Product)

The product of two bounded self-adjoint linear operators S and T on a Hilbert space H is self-adjoint if and only if the operators commute, i.e., if and only if $ST = TS$.

- We have

$$\begin{aligned}(ST)^* &= T^*S^* \\ &= TS.\end{aligned}$$

Hence $ST = (ST)^*$ iff $ST = TS$.

Sequences of Self-Adjoint Operators

Theorem (Sequences of Self-Adjoint Operators)

Let (T_n) be a sequence of bounded self-adjoint linear operators $T_n : H \to H$ on a Hilbert space H. Suppose that (T_n) converges, say, $T_n \to T$, i.e., $\|T_n - T\| \to 0$, where $\|\cdot\|$ is the norm on the space $B(H, H)$. Then the limit operator T is a bounded self-adjoint linear operator on H.

- We must show that $T^* = T$. This follows from $\|T - T^*\| = 0$. To prove the latter, we use $\|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\|$. We obtain by the triangle inequality in $B(H, H)$

$$\begin{array}{rcl} \|T - T^*\| & \leq & \|T - T_n\| + \|T_n - T_n^*\| + \|T_n^* - T^*\| \\ & = & \|T - T_n\| + 0 + \|T_n - T\| \\ & = & 2\|T_n - T\| \stackrel{n\to\infty}{\longrightarrow} 0. \end{array}$$

Hence $\|T - T^*\| = 0$ and $T^* = T$.

Unitary Operators

Theorem (Unitary Operator)

Let the operators $U: H \to H$ and $V: H \to H$ be unitary, where H is a Hilbert space. Then:

(a) U is isometric; thus $\|Ux\| = \|x\|$, for all $x \in H$;

(b) $\|U\| = 1$, provided $H \neq \{0\}$;

(c) U^{-1} $(= U^*)$ is unitary;

(d) UV is unitary;

(e) U is normal.

Furthermore:

(f) A bounded linear operator T on a complex Hilbert space H is unitary if and only if T is isometric and surjective.

(a) $\|Ux\|^2 = \langle Ux, Ux \rangle = \langle x, U^* Ux \rangle = \langle x, Ix \rangle = \|x\|^2$.

(b) follows immediately from (a).

Unitary Operators (Cont'd)

(c) Since U is bijective, so is U^{-1}, and $(U^{-1})^* = U^{**} = U = (U^{-1})^{-1}$.

(d) UV is bijective, and $(UV)^* = V^*U^* = V^{-1}U^{-1} = (UV)^{-1}$.

(e) follows from $U^{-1} = U^*$ and $UU^{-1} = U^{-1}U = I$.

(f) Suppose that T is isometric and surjective. Isometry implies injectivity. So T is bijective. We show that $T^* = T^{-1}$. By the isometry, $\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \langle x, x\rangle = \langle Ix, x\rangle$. Hence, $\langle (T^*T - I)x, x\rangle = 0$ and $T^*T - I = 0$. So $T^*T = I$. From this, $TT^* = TT^*(TT^{-1}) = T(T^*T)T^{-1} = TIT^{-1} = I$. Together, $T^*T = TT^* = I$. Hence $T^* = T^{-1}$ so that T is unitary.

The converse is clear since T is isometric by (a) and surjective by definition.

- An isometric operator need not be unitary (it may fail to be surjective):
  Example: The **right shift operator** $T : \ell^2 \to \ell^2$, given by $(\xi_1, \xi_2, \xi_3, \ldots) \mapsto (0, \xi_1, \xi_2, \xi_3, \ldots)$, where $x = (\xi_j) \in \ell^2$, is isometric but not unitary.