FUNCTIONAL ANALYSIS

JUNAID AFTAB

Abstract. These are assorted notes on various topics in functional analysis. There may be errors or typos; please send any corrections to junaid.aftab1994@gmail.com.

Contents

Let $\mathbb{K}$ be a field. We usually consider the case where $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Part 1. **Banach Spaces**

1. Definitions

We first define normed spaces.

Definition 1.1. A normed space is a pair $(X, \|\cdot\|)$, where X is a $\mathbb{K}$-vector space and

$$\|\cdot\| : X \to [0, \infty)$$

is a norm function with the following properties:

(1) (**Non-negative**) $\|x\| = 0$ implies $x = 0$;
(2) (**Scalar Homogeneity**) $\|cx\| = |c|\|x\|$ for all $c \in \mathbb{K}$ and $x \in X$;
(3) (**Triangle Inequality**) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The triangle inequality implies that every normed space is a metric space with distance function

$$d(x, y) := \|x - y\|$$

This observation allows us to introduce notions from analysis in our study of infinite dimensional vector spaces. For instance, we can talk about open sets, closed sets, compact sets, limits, sequences, convergent sequences, continuity etc. In particular, we say that a sequence $(x_n)_{n\geq 1}$ in X is said to converge if there exists an element $x \in X$ such that

$$\lim_{n\to\infty} \|x_n - x\| = 0.$$

It is easy to check that this element, if it exists, is unique and is called the limit of the sequence $(x_n)_{n\geq 1}$. We then write $\lim_{n\to\infty} x_n = x$ or simply $x_n \to x$ as $n \to \infty$. Moreover, given a sequence $(x_n)_{n\geq 1}$ in a normed space X, the sum $\sum_{n=1}^{\infty} x_n$ is said to be convergent if there exists $x \in X$ such that

$$\lim_{N\to\infty} \left\| x - \sum_{n=1}^{N} x_n \right\| = 0.$$

The sum $\sum_{n=1}^{\infty} x_n$ is said to be absolutely convergent if $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

Remark 1.2. *We use the notation*

$$B(x_0, r) := \{x \in X : \|x - x_0\| < r\}$$
$$\overline{B}(x_0, r) := \{x \in X : \|x - x_0\| \leq r\}.$$

Here $B(x_0, r)$ the open ball centered at $x_0 \in X$ with radius $r > 0$. We shall see in Proposition 1.3 that $\overline{B}(x_0, r)$ is the closure of $B(x_0, r)$.

The geometry of a normed space can be very different from that of the usual Euclidean geometry. For instance, each $B(x_0, r)$ need to be "round" anymore. Nevertheless, some important important properties still hold.

Proposition 1.3. *Let X be a normed space. We have*

(1) $\overline{B(0,1)} = \overline{B}(0, 1)$
(2) *Each $B(x_0, r)$ and $\overline{B}(x_0, r)$ is a convex set.*

We discuss some properties of the norm function.

Proof. The proof is given below:

(1) The inclusion $\overline{B(0,1)} \subseteq \overline{B}(0,1)$ is trivial, because $\overline{B}(0,1)$ is a closed set that contains $B(0;1)$ and $\overline{B(0,1)}$ is the smallest closed set that contains $B(0,1)$. Let $x \in \overline{B}(0,1)$ and defines, for each $n \in \mathbb{N}$,

$$x_n := \left(1 - \frac{1}{n}\right) x$$

The sequence $(x_n)_{n\geq 1}$ converges to x in the norm $\|\cdot\|$ since

$$\|x_n - x\| = \left\| \left(1 - \frac{1}{n}\right) x - x \right\| = \frac{1}{n}\|x\| \leq \frac{1}{n} \longrightarrow 0,$$

It is clear that for all $n \in \mathbb{N}$, $x_n \in B(0,1)$. Hence, $\overline{B}(0,1) \subseteq \overline{B(0,1)}$.

(2) We prove the convexity of $B(x_0,r)$. Choose arbitrary $x, y \in B(x_0,r)$, $\lambda \in [0,1]$. We have,

$$\|\lambda x + (1-\lambda)y\| \leq \lambda\|x\| + (1-\lambda)\|y\| < \lambda r + (1-\lambda)r = r.$$

It follows that $\lambda x + (1-\lambda)y \in B(x_0,r)$ as required. Similarly, $\overline{B}(x_0,r)$ is a convex set.

This completes the proof. □

The fact that a normed space is a Banach space has a number of consequences. We state some immediate ones below:

Proposition 1.4. *Let $(X, \|\cdot\|)$ be a normed space. Then $(X, \|\cdot\|)$ has the following properties:*

(1) *For each $x, x' \in X$*

$$\big|\|x\| - \|x'\|\big| \leq \|x - x'\|$$

(2) *The function $x \mapsto \|x\|$ is a continuous and convex function.*

(3) *The addition and scalar multiplication vector space operations are continuous functions.*

Proof. The proof is given below:

(1) Triangle inequality implies that

$$\|x\| - \|x'\| \leq \|x - x'\|$$
$$\|x'\| - \|x\| \leq \|x' - x\|$$

Therefore,

$$\big|\|x\| - \|x'\|\big| \leq \|x - x'\|$$

(2) Continuity is a straightforward consequence of (1). Convexity of the norm follows from the norm axioms. Indeed, for every $x, y \in X$ and $\lambda \in [0,1]$ we have

$$\|\lambda x + (1-\lambda)y\| \leq \lambda\|x\| + (1-\lambda)\|y\|.$$

(3) Let $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} x'_n = x'$ in X, and $k \in \mathbb{K}$. Then

$$\lim_{n\to\infty} \|cx_n - cx\| = |c| \lim_{n\to\infty} \|x_n - x\| = 0$$

Similarly,

$$\lim_{n\to\infty} \|(x_n + x'_n) - (x - x')\| \leq \lim_{n\to\infty} \|x_n - x\| + \lim_{n\to\infty} \|x'_n - x'\| = 0$$

This completes the proof. □

Remark 1.5. *An important observation is that if the sub-level set*

$$\{x \in X : \|x\| \leq 1\}$$

is convex, then $\|\cdot\|$ *is ea norm on* E*. Indeed, let* $x, y \in X$*. We want to show that* $\|x+y\| \leq \|x\| + \|y\|$*. This is equivalent to*

$$\left\| \frac{x}{\|x\|+\|y\|} + \frac{y}{\|x\|+\|y\|} \right\| \leq 1.$$

The above inequality can be written as

$$\left\| \frac{\|x\|}{\|x\|+\|y\|} \frac{x}{\|x\|} + \frac{\|y\|}{\|x\|+\|y\|} \frac{y}{\|y\|} \right\| \leq 1.$$

This is indeed true by assumption.

We can now define Banach spaces.

Definition 1.6. A Banach space is a complete normed space.

The following proposition gives a necessary and sufficient condition for a normed space to be a Banach space.

Proposition 1.7. *A normed space* X *is a Banach space if and only if every absolutely convergent sum in* X *converges in* X*.*

Proof. Suppose that X is complete and let $\sum_{n\geq 1} x_n$ be absolutely convergent. If $n > m$ the triangle inequality implies

$$\left\| \sum_{j=1}^{n} x_j - \sum_{j=1}^{m} x_j \right\| = \left\| \sum_{j=m+1}^{n} x_j \right\| \leq \sum_{j=m+1}^{n} \|x_j\| \to 0$$

Hence $\left(\sum_{j=1}^{n} x_j\right)_{n\geq 1}$ is a Cauchy sequence and it converges by completeness. Conversely, let $(x_n)_{n\geq 1}$ be a Cauchy sequence. Choose indices $n_1 < n_2 < \dots$ in such a way that

$$\|x_i - x_j\| < \frac{1}{2}$$

for all $i, j > n_k$, $k = 1, 2, \dots$. The sum $x_{n_1} + \sum_{k\geq 1}(x_{n_{k+1}} - x_{n_k})$ is absolutely convergent since

$$\sum_{k\geq 1} \|x_{n_{k+1}} - x_{n_k}\| \leq \sum_{k\geq 1} \frac{1}{2^k} < \infty.$$

By assumption it converges to some $x \in X$. Then, by cancellation,

$$x = \lim_{m\to\infty} \left(x_{n_1} + \sum_{k=1}^{m} (x_{n_{k+1}} - x_{n_k}) \right) = \lim_{m\to\infty} x_{n_m+1}.$$

Therefore, the subsequence $(x_{n_m})_{m>1}$ is convergent. It is a standard fact from analysis that a Cauchy sequence with a convergent subsequence is itself convergent. □

Remark 1.8. *It turns out that every normed space can be completed to a Banach space. More precisely, if* X *is a normed space then there exists a unique Banach space* $\overline{X}$ *containing* X *isometrically as a dense subspace.*

2. Constructing New Banach Spaces

Several abstract constructions enable us to create new Banach spaces from given ones. We take a brief look at some basic constructions.

2.1. **Subspaces.** A subspace Y of a normed space X is a normed space with respect to the norm inherited from X.

Proposition 2.1. *A subspace Y of a Banach space X is a Banach pace with respect to the norm inherited from X if and only if Y is closed in X.*

Proof. Assume that $Y \subseteq X$ is a closed subspace of X. Suppose $(y_n)_{n\geq 1}$ is a Cauchy sequence in Y. Then it has a limit in X, by the completeness of X, and this limit belongs to Y, since Y is closed. Conversely, assume that Y is a Banach space. Let $x \in X$ such that there is a sequence $(y_n)_{n\geq 1}$ in Y such that $y_n \to x$. Since Y is complete, we have that $y_n \to y$ for some $y \in Y$. It follows from uniqueness of limits that $x = y \in Y$. Hence, $x \in Y$, and this shows that Y is closed. □

2.2. **Quotient Spaces.** If Y is a closed subspace of a Banach space X, the quotient space X/Y can be endowed with a norm by

$$\|[x]\| := \|x + Y\| := \inf_{y\in Y} \|x - y\|,$$

This is indeed a norm. If $\|[x]\| = 0$, there is a sequence $(y_n)_{n\geq 1}$ in Y such that $\|x - y_n\| < \frac{1}{n}$ for all $n \geq 1$. Then

$$\|y_n - y_m\| \leq \|y_n - x\| + \|x - y_m\| < \frac{1}{n} + \frac{1}{m},$$

So $(y_n)_{n\geq 1}$ is a Cauchy sequence in X. It has a limit $y \in X$ since X is complete, and we have $y \in Y$ since Y is closed. Then

$$\|x - y\| = \lim_{n\to\infty} \|x - y_n\| = 0$$

so $x = y$. This implies that $[x] = [y] = [0]$, the zero element of X/Y. The triangle inequality and scalar homogenity are trivially verified. In fact, X/Y is a Banach space.

Proposition 2.2. *Let X be a Banach space and let Y be a closed subspace of a Banach space X. The quotient space X/Y is a Banach space with the norm*

$$\|[x]\| := \|x + Y\| := \inf_{y\in Y} \|x - y\|,$$

Proof. Suppose that $\sum_{n\geq 1} \|[x_n]\| < \infty$. Choose $y_n \in Y$ are such that $\|x_n - y_n\| \leq \|[x_n]\| + \frac{1}{n^2}$. Since X is a Banach space, Proposition 1.7 implies that $\sum_{n\geq 1}(y_n - x_n)$ converges in X, say to x. Then, for all $n \geq 1$,

$$\left\| [x] - \sum_{n=1}^{N} [x_n] \right\| = \left\| \left[x - \sum_{n=1}^{N} x_n \right] \right\| \leq \left\| x - \sum_{n=1}^{N} x_n + \sum_{n=1}^{N} y_n \right\| = \left\| x - \sum_{n=1}^{N} (x_n - y_n) \right\|.$$

As $N \to \infty$, the right-hand side tends to 0 and therefore

$$\lim_{N\to\infty} \sum_{n=1}^{N} [x_n] = [x]$$

in X/Y. By Proposition 1.7, X/Y is a Banach space. □

3. Examples

This section is devoted to looking at a number of examples of Banach spaces.

3.1. **Spaces of Continuous Functions.** We discuss the following function spaces:

(1) $C(X)$, the space of continuous functions defined on a compact topological space X.
(2) $C_b(X)$, the space of continuous functions defined on a locally compact Hausdorff space X.
(3) $C_0(X)$, the space of continuous functions that vanish at infinity defined on locally compact Hausdorff space X[1].

Remark 3.1. *We restrict to the case of a locally compact Hausdorff topological space since then Urysohn's lemma applies. Hence, there exists an abundance of continuous functions on X.*

Proposition 3.2. *Let X be a topological space.*

(1) *If X is a compact topological space, $C(X)$ is a Banach space with respect to the supremum norm*
$$\|f\|_\infty := \sup_{x\in X} |f(x)|$$
(2) *If X is a locally compact Hausdorff topological space, $C_b(X)$ is a Banach space with respect to the supremum norm*
$$\|f\|_\infty := \sup_{x\in X} |f(x)|$$
(3) *If X is a locally compact Hausdorff topological space, $C_0(X)$ is a Banach space with respect to the supremum norm*
$$\|f\|_\infty := \sup_{x\in X} |f(x)|$$

Proof. Clearly, $\|\cdot\|_\infty$ is a norm in all three cases. We check completeness below:

(1) Suppose that $(f_n)_{n\geq 1}$ is a Cauchy sequence in $C(X)$. Then for each $x \in X$, $(f_n(x))_{n\geq 1}$ is a Cauchy sequence in $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and therefore convergent to some limit in $\mathbb{K}$ which we denote by $f(x)$. Hence the pointwise limit $f(x) = \lim_{n\to\infty} f_n(x)$ is a $\mathbb{K}$-valued function. Fix $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that
$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty < \varepsilon$$
for all $m, n \geq N$ and $x \in X$. Passing to the limit $m \to \infty$ while keeping n fixed we obtain
$$|f_n(x) - f(x)| \leq \varepsilon$$
for each $n \geq N$. Fix $x_0 \in X$ arbitrarily and let $U \subseteq X$ be an open set containing x_0 such that $|f_N(x) - f_N(x_0)| < \varepsilon$ whenever $x \in U$. Then, for $x \in U$,
$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon, \end{aligned}$$
This proves the continuity of f at the point x_0. Since x_0 is arbitrary, $f \in C(X)$. Moreover,
$$|f_n(x) - f(x)| \leq \|f_n - f\|_\infty < \varepsilon$$
Hence, $(f_n)_{n\geq 1}$ converges to f in $C(X)$.

[1]We say $f \in C_0(X)$ if and only if for every $\varepsilon > 0$ there is a compact set K_ε such that $|f(x)| < \varepsilon$ for $x \in K_\varepsilon^c$.

(2) The proof is similar to that of (1). Suppose that $(f_n)_{n\geq 1}$ is a Cauchy sequence in $C_b(X)$. For fixed $x \in X$, $(f_n(x))_{n\geq 1}$ is a Cauchy sequence in $\mathbb{K}$ and therefore convergent to some limit $f(x)$ in $\mathbb{K}$. Hence the pointwise limit $f(x) = \lim_{n\to\infty} f_n(x)$ is a $\mathbb{K}$-valued function. Let N be such that for $n, m \geq N$ we have $\|f_n - f_m\|_\infty < 1$. Then for all $x \in X$

$$\begin{aligned} |f(x)| &\leq |f(x) - f_N(x)| + |f_N(x)| \\ &= |\lim_{n\to\infty} f_n(x) - f_N(x)| + |f_N(x)| \\ &= \lim_{n\to\infty} |f_n(x) - f_N(x)| + |f_N(x)| \\ &\leq 1 + \|f_N\|_\infty \end{aligned}$$

Hence, $\|f\|_\infty \leq 1 + \|f_N\|_\infty$ implying that f is bounded. Let $\varepsilon > 0$. Let N be such that for $n, m \geq N$ we have $\|f_n - f_m\|_\infty < \varepsilon$. Then for all $n \geq N$

$$|f(x) - f_n(x)| = \lim_{m\to\infty} |f_m(x) - f_n(x)| \leq \varepsilon$$

for all $x \in X$ and hence $\|f - f_n\|_\infty \leq \varepsilon$. Hence, $(f_n)_{n\geq 1}$ converges to f in $C_b(X)$ is similar to the proof in (1).

(3) We claim that $C_0(X)$ is a closed subsapce of $C_b(X)$. Checking that it is a subspace is a straightforward computation. Let $(f_n)_{n\geq 1}$ be a sequence in $C_0(X)$ such that $(f_n)_{n\geq 1}$ converges to some $f \in C_b(X)$. Fix $\varepsilon > 0$. There is an $N \in \mathbb{N}$ such that $\|f - f_N\|_\infty < \epsilon/2$. Moreocer, there is a compact $K_N \subseteq X$ such that $|f_N(x)| < \epsilon/2$ for $x \in X \setminus K_N$. It follows that

$$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| < \epsilon$$

for $x \in X \setminus K_N$. Since ϵ was arbitrary, it follows that $f \in C_0(X)$. By Proposition 2.1, $C_0(X)$ is a Banach space.

This completes the proof. □

Remark 3.3. *If X is a locally compact Hausdorff space, we have the following inclusions of Banach spaces:*

$$C_0(X) \subset C_b(X) \subset C(X)$$

Let $C_c(X)$ be the space of continuous functions with compact support. We have,

$$C_c(X) \subset C_0(X) \subset C_b(X) \subset C(X)$$

However, $C_c(X)$ is not a Banach space. We can show that it is not a Banach space by showing that $C_c(X)$ is not closed in $C_0(X)$. In fact, one can show that $C_c(X)$ is dense in $C_0(X)$[2]. If X is a compact Hausdorff space, then

$$C_0(X) = C_b(X) = C(X)$$

Remark 3.4. *If $X = \mathbb{N}$, then $C_b(\mathbb{N})$ can be identified with $l^\infty(\mathbb{N})$, the space of all bounded sequences. Hence, we see that $l^\infty(\mathbb{N})$ is a Banach space. Moreover, $C_0(\mathbb{N})$ can be identified with c_0, the space of all seqences converging to zero. Hence, c_0 is also a Banach space.*

[2]This is most easily done by invoking properties about Lebesgue spaces over locally compact Hausdorff space. We don't discuss these in these notes so we skip details.

3.2. **Lebesgue Spaces.** Lebesgue spaces are function spaces that measure the intergrability of a function.

Definition 3.5. Let $(X, \mathscr{M}, \mu)$ be a measure space and $1 \leq p < \infty$. If $f : X \to \mathbb{K}$ is a measurable function, then we define

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p}$$

For $1 \leq p < \infty$, the space $\mathcal{L}^p(X, \mathscr{M}, \mu)$ is the set

$$\mathcal{L}^p(X, \mathscr{M}, \mu) = \{f : X \to \mathbb{R} \,|\, \|f\|_p < \infty\}.$$

We write $\mathcal{L}^p(X, \mathscr{M}, \mu)$ as $\mathcal{L}^p(X)$.

Proposition 3.6. *Let $(X, \mathscr{M}, \mu)$ be a measure space. Let $1 \leq p \leq \infty$. Then $\mathcal{L}^p(X)$ is a $\mathbb{K}$-vector space.*

Proof. Let $\alpha \in \mathbb{K}$ and $f, g \in \mathcal{L}^p(X)$. Note that:

$$\|\alpha f\|_p := |\alpha| \left(\int_X |f|^p \, d\mu \right)^{1/p} < \infty$$

Moreover, the elementary estimates

$$\begin{aligned} |f(x) + g(x)|^p &\leq (2 \max(|f(x)|, |g(x)|))^p \\ &= 2^p \max(|f(x)|^p, |g(x)|^p) \\ &\leq 2^p(|f(x)|^p + |g(x)|^p). \end{aligned}$$

implies that

$$\|f + g\|_p^p \leq 2 \left(\int_X |f(x)|^p d\mu + \int_X |g(x)|^p d\mu \right) < \infty$$

As a result $\|f + g\|_{L^p(X)} < \infty$. This shows that $\mathcal{L}^p(X)$ is a $\mathbb{K}$-vector space. □

We wish to show that show that $\mathcal{L}^p(X)$ is a Banach space. This poses a problem: for $1 \leq p < \infty$, $\| \cdot \|_p$ is not even a norm on $\mathcal{L}^p(X)$, because $\|f\|_p = 0$ only implies that $f = 0$ μ-almost everywhere. In spirit of Lebesgue's philosophy of ignoring whatever is going on on a set of measure zero, we define an equivalence relation $\sim$ on $\mathcal{L}^p(X)$ by

$$f \sim g \iff f = g \ \mu\text{-almost everywhere.}$$

The equivalence class of a function f modulo $\sim$ is denoted by $[f]$. On the quotient space

$$L^p(X) := \mathcal{L}^p(X)/ \sim,$$

we define scalar multiplication and addition in the natural way:

$$\begin{aligned} c[f] &:= [cf], \\ [f] + [g] &:= [f + g]. \end{aligned}$$

It is easy to check that both operations are well-defined. Following common practice, we make no distinction between functions in $\mathcal{L}^p(X)$ and their equivalence classes in $L^p(X)$, and call the latter 'functions' as well.

Proposition 3.7. *Let $(X, \mathscr{M}, \mu)$ be a measure space. Let $1 \leq p < \infty$. Then $L^p(X)$ is a normed vector space with the norm:*

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p}$$

Additionally, $L^p(X)$ is a Banach space.

Proof. It is clear that $\|\cdot\|_p$ is non-negative and scalar homogeneous. We prove the triangle inequality which is referred to as Minkowski's Inequality. Based on Remark 1.5, it suffices to check that the sub-level set,

$$\{f \in L^p(X) \mid \|f\|_p \leq 1\},$$

is a convex set. Let $f, g \in L^p(X)$ such that $\|f\|_p, \|g\|_p \leq 1$ and $\lambda \in [0, 1]$. Since the function $x \mapsto |x|^p$ is convex on $\mathbb{R}$ for $p \geq 1$, we have a pointwise inequality

$$|\lambda f(x) + (1-\lambda)g(x)|^p \leq \lambda |f(x)|^p + (1-\lambda)|g(x)|^p.$$

Integrating both sides of this inequality implies

$$\int_X |\lambda f + (1-\lambda)g|^p \, d\mu \leq \lambda \int_X |f|^p \, d\mu + (1-\lambda) \int_X |g|^p \, d\mu \leq 1$$

This shows that the triangle inequality holds in $L^p(X)$. Hence, $L^p(X)$ is a normed vector space.

We now show that $L^p(X)$ is a Banach space. Suppose $(f_n)_{n\geq 1}$ is a sequence $L^p(X)$, and $\sum_{n=1}^{\infty} \|f_n\|_p = B < \infty$. Let

$$G_k(x) = \sum_{n=1}^{k} |f_n(x)|, \qquad G = \sum_{n=1}^{\infty} |f_n(x)|$$

Minkowski's inequality implies that $\|G_k\|_p \leq B$ for all n, so by the monotone convergence theorem,

$$\int_X G^p d\mu = \lim_{n\to\infty} \int G_n^p \leq B^p$$

Hence $G \in L^p(X)$. In particular, $G < \infty$ almost everywhere which implies that the series

$$F(x) := \sum_{n=1}^{\infty} f_n(x)$$

converges almost everywhere. Hence $F \in L^p(X)$. Moreover, $|F - \sum_{n=1}^{k} f_n|^p \leq (2G)^p \in L^1(X)$. By the dominated convergence theorem,

$$\left\| F - \sum_{n=1}^{k} f_n \right\|_p^p = \int_X \left| F - \sum_{n=1}^{k} f_n \right|^p d\mu \to 0$$

Thus the series $\sum_{n=1}^{\infty} f_n$ converges in the $L^p(X)$ norm. $\square$

Remark 3.8. *If $X = \mathbb{N}$ and $\mu = \#$ is the counting measures on X, then*

$$\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \#) = \ell^p(\mathbb{N}) = \left\{ (a_n)_{n\geq 1} \in \mathbb{K}^\infty : \left(\sum_{n\geq 1} |a_n|^p \right)^{1/p} < \infty \right\}$$

for $1 \leq p < \infty$. Here $\ell^p(\mathbb{N})$ is the space of p-summable sequences. Hence, $\ell^p(\mathbb{N})$ is a Banach space.

In fact, the Lebesgue spaces make sense for $p = \infty$.

Definition 3.9. Let $(X, \mathscr{M}, \mu)$ be a measure space. If $f : X \to \mathbb{K}$ is a measurable function, then we define

$$\|f\|_\infty \equiv \inf\{C \geq 0 : |f(x)| \leq C \text{ for almost every } x\}.$$

The space $\mathcal{L}^\infty(X, \mathscr{M}, \mu)$ is the set

$$\mathcal{L}^\infty(X) = \{f : X \to \mathbb{R} \mid \|f\|_\infty < \infty\}.$$

Once again, we define

$$L^\infty(X) := \mathcal{L}^\infty(X)/\sim,$$

where equivalence relation $\sim$ on $\mathcal{L}^\infty(X)$ by

$$f \sim g \iff f = g \ \mu\text{-almost everywhere.}$$

Addition and scalar multiplication on $L^\infty(X)$ is defined as before. Once again, we make no distinction between functions in $\mathcal{L}^\infty(X)$ and their equivalence classes in $L^p(X)$, and call the latter 'functions' as well.

Proposition 3.10. *Let $(X, \mathscr{M}, \mu)$ be a measure space. Then $L^\infty(X)$ is a Banach space.*

Proof. It is clear that $\|\cdot\|_\infty$ is a norm. We first claim that $\|f_n - f\|_\infty \to 0$ if and only if there exists $E \subseteq X$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E. Suppose $\|f_n - f\|_\infty \to 0$. Given $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for $n \geq N$

$$\|f_n - f\|_\infty < \epsilon$$

Thus for any $n \geq N$, we have

$$|f_n(x) - f(x)| \leq \|f_n - f\|_\infty < \epsilon \tag{1}$$

a.e. on X. Let $M_n = \|f_n - f\|_\infty$ for all $n \geq N$ and set

$$A_n = \{x \in X : |f_n(x) - f(x)| > M_n\}$$

Then we have $\mu(A_n) = 0$. Now, let $A = \bigcup_{n \geq N} A_n$ then $\mu(A) = 0$. Let $E = A^c$ Then for $x \in E$ we have from (1)

$$|f_n(x) - f(x)| \leq \|f_n - f\|_\infty < \epsilon$$

for all $n \geq N$. Thus,

$$|f_n(x) - f(x)| < \epsilon$$

for all $n \geq N$. So $f_n \to f$ uniformly on E and clearly $\mu(E^c) = 0$. Conversely, suppose $E \subseteq M$ and $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E. Then for each $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \epsilon$$

for all $n \geq N$ and $x \in E$. Hence we also have

$$|f_n(x) - f(x)| < \epsilon \ \text{ a.e. on } \ X$$

Thus by definition of the$\|\cdot\|_\infty$ we have

$$\|f_n - f\|_\infty < \epsilon$$

We now show that $L^\infty(X)$ is a completed normed vector space. Let $\{f_n\}_{n \in \mathbb{N}}$ be a Cauchy sequence in $L^\infty(X)$. Thus given $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\|f_m - f_n\|_\infty < \epsilon$$

for all $m, n \geq N$. For each $m, n \in \mathbb{N}$, set

$$F_{m,n} = \{x \in X : |f_m(x) - f_n(x)| > \|f_m - f_n\|_\infty\}$$

Then clearly $\mu(F_{m,n}) = 0$ for all $m, n \in \mathbb{N}$. Set $F = \bigcup_{m,n \in \mathbb{N}} F_{m,n}$ and $E = F^c$. Note that $\mu(E^c) = \mu(F) = 0$. Moreover,

$$\begin{aligned} E &= \bigcap_{m,n \in \mathbb{N}} \{x \in X : |f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty\} \\ &= \{x \in X : |f_m(x) - f_n(x) \leq \|f_m - f_n\|_\infty \ \text{ for all } \ m, n \geq N\} \end{aligned}$$

Let $\epsilon > 0$. Then for $x \in E$ and for all $m, n \geq N$ we have

$$|f_m(x) - f_n(x)| \leq \|f_n - f_m\| < \epsilon \tag{2}$$

This shows that for every $x \in E$, $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{K}$. Since $\mathbb{K}$ is complete, there exists a limit

$$f(x) = \lim_{n\to\infty} f_n(x)$$

Note $f(x)$ is defined in E (i.e. outside of F). Thus for $x \in F$, $f(x) = 0$. Note that $f = \lim_{n\to\infty} f_n(x)\chi_E$ is measurable. We have

$$|f_n(x) - f(x)| \leq \epsilon$$

for $x \in E$. Thus for $n \geq N$

$$\|f_n - f\|_\infty \leq \epsilon$$

This shows that $f_n \to f$ in L^∞ norm. Finally, we note that $f \in L^\infty$ from the triangle inequality:

$$\|f\|_\infty \leq \|f_N\|_\infty + \|f_N - f\|_\infty \leq \|f_N\|_\infty + \epsilon < \infty$$

Hence, $L^\infty(X)$ is a Banach space. □

Remark 3.11. *If $X = \mathbb{N}$ and $\mu = \#$ is the counting measures on X, then*

$$\ell^\infty(\mathbb{N}) := \mathcal{L}^\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \#) = \{(a_n)_{n\geq 1} \in \mathbb{K}^\infty : \sup_{n\in\mathbb{N}} |a_n| < \infty < \infty\}$$

Here $\ell^\infty(\mathbb{N})$ is the space of bounded sequences. Hence, $\ell^p(\mathbb{N})$ is a Banach space.

4. Operators on Banach Spaces

Let X, Y be normed spaces. Linear operators respect the underlying vector space structure of X and Y. Since X and Y are normed spaces, continuous linear operators form the right class of operators to study between X and Y. We first define bounded operators between X and Y.

Definition 4.1. Let X an Y be normed spaces A linear operator $T : X \to Y$ is bounded if there exists a finite constant $C > 0$ such that

$$\|Tx\|_Y \leq C\|x\|_X$$

for all $x \in X$. The operator norm, $\|T\|$, is defined as

$$\|T\| = \inf\{C : \|Tx\|_Y \leq C\|x\|_X \text{ for all } x \in X\}$$

Remark 4.2. *In what follows, we will write a norm $\|\cdot\|_X, \|\cdot\|_Y$ as simply $\|\cdot\|$.*

Surprisingly, it turns out that continuous operators between X and Y and bounded operators (to be defined below) between X and Y define the same class of operators.

We now provide alternative characterizations of the operator norm of a bounded linear operator.

Proposition 4.3. *Let X and Y be normed spaces and let $T : X \to Y$ be a linear operator. The following are equivalent:*

(1) *T is bounded*
(2) *T is continuous*
(3) *T is continuous at some point $x_0 \in X$*

Proof. The implication (1) ⇒ (2) follows from the observation that

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\|\|x - x_0\|$$

and the implication (2) ⇒ (3) is trivial. To prove implication (3) ⇒ (1), suppose that T is continuous at x_0. Then there exists a $\delta > 0$ such that

$$\|x_0 - y\| < \delta \quad \Rightarrow \quad \|Tx_0 - Ty\| < 1$$

Since every $x \in X$ with $\|x\| < \delta$ is of the form $x = x_0 - y$ with $\|x_0 - y\| < \delta$ (take $y = x_0 - x$) and T is linear, it follows that $\|x\| < \delta$ implies $\|Tx\| < 1$. By scalar homogeneity and the linearity of T, we may scale both sides with a factor δ, and obtain that $\|x\| < 1$ implies $\|Tx\| < 1/\delta$. Hence, T is bounded and $\|T\| \leq 1/\delta$. $\square$

Proposition 4.4. *Let X an Y be normed spaces and let $T : X \to Y$ be a bounded operator. Then:*

$$\|T\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|\leq 1} \|Tx\| = \sup_{\|x\|=1} \|Tx\|| = \sup_{\|x\|<1} \|Tx\|$$

Proof. For each $x \in X$

$$\|Tx\| \leq \sup_{\|x\|\neq 0} \frac{\|Tx\|}{\|x\|}\|x\|$$

Therefore

$$\|T\| \leq \sup_{\|x\|\neq 0} \frac{\|Tx\|}{\|x\|}$$

Let $C > 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$. Then $\|Tx\| \leq C$ for each $x \in X$ such that $\|x\| \leq 1$. Hence, $\sup_{\|x\|\leq 1} \|Tx\| \leq C$. Moreover if $C > 0$ in the definition of $\|T\|$, then

$$\frac{\|Tx\|}{\|x\|} \leq \frac{C\|x\|}{\|x\|} = C$$

for each $x \neq 0$. Hence

$$\sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} \leq \|T\|$$

Clearly,

$$\sup_{\|x\|=1} \|Tx\| \leq \sup_{\|x\|\leq 1} \|Tx\| \leq \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} \leq \sup_{\|x\|=1} \|Tx\|$$

The last inequality follows from the observation that $\frac{\|Tx\|}{\|x\|} = T\Big(\frac{x}{\|x\|}\Big)$ for each $x \neq 0$. This proves that

$$\|T\| = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|\leq 1} \|Tx\| = \sup_{\|x\|=1} \|Tx\|$$

Note that

$$\sup_{\|x\|<1} \|Tx\| \leq \sup_{\|x\|\leq 1} \|Tx\|$$

If $\|x\| = 1$, then there is a sequence $(x_n)_{n\geq 1}$ such that $\|x_n\| < 1$ and $x_n \to x$. Since T is continuous (see Proposition 4.3), we have $Tx_n \to Tx$. This implies that

$$\sup_{\|x\|\leq 1} \|Tx\| \leq \sup_{\|x\|<1} \|Tx\|$$

Hence,

$$\|T\| = \sup_{\|x\|<1} \|Tx\|$$

This completes the proof. $\square$

Remark 4.5. *Here is a cute observation:*

$$\sup\{\|Tx\| : \|x\| \leq r\} = r\|T\|$$

Indeed, let x^* such that $\|x^*\| \leq 1$ and $\|T\| = \|T(x^*)\|$. It is clear that

$$\sup\{\|Tx\| : \|x\| \leq r\} = \|T(rx^*)\| = r\|T(x^*)\| = r\|T\|$$

The set of all bounded operators from X to Y is a vector space in a natural way with respect to pointwise scalar multiplication and addition. This vector space will be denoted by $\mathscr{B}(X, Y)$. In fact, it is a Banach space.

Proposition 4.6. *Let X and Y be normed spaces. If Y is a Banach space, then $\mathscr{B}(X,Y)$ is a Banach space.*

Proof. For all $T, T' \in \mathscr{B}(X,Y)$. We show that

$$\|T + T'\| \leq \|T\| + \|T'\|$$

For all $x \in X$, the triangle inequality gives

$$\|(T+T')x\| \leq \|Tx\| + \|T'x\| \leq (\|T\| + \|T'\|)\|x\|,$$

and the result follows by taking the supremum over all $x \in X$ with $\|x\| \leq 1$. Similarly, $\|cT\| = |c|\|T\|$ for $c \in \mathbb{K}$. Noting that $\|T\| = 0$ implies $T = 0$, it follows that $T \mapsto \|T\|$ is a norm on $\mathscr{B}(X,Y)$ and $\mathscr{B}(X,Y)$ is a normed space. We now show that $\mathscr{B}(X,Y)$ is a Banach space. Let $(T_n)_{n\geq 1}$ be a Cauchy sequence in $\mathscr{B}(X,Y)$. From

$$\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|$$

we see that $(T_n x)_{n\geq 1}$ is a Cauchy sequence in Y for every $x \in X$. Let Tx denote its limit. The linearity of each of the operators T_n implies that the mapping $T : x \mapsto Tx$ is linear and we have

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| \leq M\|x\|,$$

where $M := \limsup_{n\to\infty} \|T_n\|$ is finite since Cauchy sequences in normed spaces are bounded. This shows that the linear operator T is bounded, so it is an element of $\mathscr{B}(X,Y)$. Fix $\varepsilon > 0$ and let $N \in \mathbb{N}$ such that $\|T_n - T_m\| < \varepsilon$ for all $m, n > N$. Then, for $m, n \geq N$, from $\|T_n x - T_m x\| \leq \varepsilon\|x$ it follows, upon letting $m \to \infty$, that

$$\|T_n x - Tx\| \leq \varepsilon\|x\|$$

This being true for all $x \in X$ and $n > N$, it follows that $\|T_n - T\| \leq \varepsilon$ for all $n \geq N$. $\square$

Example 4.7. (Evaluation Operator) Let X be a compact topological space. For each $x_0 \in X$, we define the point evaluation map.

$$E_{x_0} : C(X) \to \mathbb{K}$$
$$f \mapsto f(x_0)$$

Clearly, E_{x_0} is a linear map. Moreover, it is a bounded linear map with norm $\|E_{x_0}\| = 1$. Boundedness with norm $\|E_{x_0}\| \leq 1$ follows from

$$|E_{x_0} f| = |f(x_0)| \leq \sup_{x\in X} |f(x)| = \|f\|_\infty.$$

By considering $f = 1$, the constant-one function on K, it is seen that $\|E_{x_0}\| = 1$. As an application of the use of the evaluation map, we claim that

$$A = \{f \in C(X) : f(x) \geq 0 \text{ for all } x \in X\}$$

is a closed set. Indeed,

$$A = \bigcap_{x\in X} E_x^{-1}[0,\infty)$$

Hence, A is closed.

Remark 4.8. *The set*

$$B = \{f \in C(X) : f(x) > 0 \textit{ for all } x \in X\}$$

can be shown to be an open set using the definition of the supremum norm.

Example 4.9. (Integration) Let $(X, \mathscr{M}, \mu)$ be a measure space. We can define the integration map.

$$\begin{aligned} I_\mu : L^1(X) &\to \mathbb{X} \\ f &\mapsto \int_X f \, d\mu \end{aligned}$$

Clearly, I_μ is a linear map. Moreover, it is a bounded linear map with norm $\|I_\mu\| = 1$. Boundedness with norm $\|I_\mu\| \leq 1$ follows from

$$|I_\mu f| = \left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu = \|f\|_1.$$

By considering non-negative functions it is seen that $\|I_\mu\| = 1$.

Example 4.10. We can similarly define an integration operator on $C(X)$ where X is a topological space. As an application of the integration operator, we show that a bounded operator need not attain their norm[3]. Consider

$$X = \{f \in C[0,1] : f(0) = 0\}$$

A simple argument shows that X is a closed subspace of $C([0,1])$. Hence, X is a Banach space. Consider the integration map on X:

$$\begin{aligned} T : X &\to \mathbb{K} \\ f &\mapsto \int_0^1 f(t) \, dt \end{aligned}$$

It is easy to check that that T is bounded with norm $\|T\| = 1$. If $f \in X$ such that $\|f\|_\infty \leq 1$, then a simple geometric argument[4] shows that the graph of $|f|$ is strictly contained in $[0,1] \times [0,1]$. Hence, $|Tf| < 1$ for each such f.

Example 4.11. (Integral Operators) Let (X, μ) be a compact metric space with a finite Borel measure. Then $X \times X$ is a compact metric space with the product metric. Let $k(s,t) \in C(X \times X)$ and define, for $f \in C(X)$, the function

$$\begin{aligned} T : C(X) &\to C(X) \\ f &\mapsto \left(s \mapsto \int_X k(s,t) f(t) \, d\mu(t) \right) \end{aligned}$$

It is easy to see that $Tf \in C(X)$ for each $f \in C(X)$. Indeed, given $\varepsilon > 0$, choose $\delta > 0$ so small that $d((s,t),(s',t')) < \delta$ implies $|k(s,t) - k(s',t')| < \varepsilon$. Then $d(s,s') < \delta$ implies

$$|Tf(s) - Tf(s')| \leq \varepsilon \int_X |f(t)| \, d\mu(t) \leq \varepsilon \mu(X) \|f\|_\infty.$$

Hence, T is is a linear operator on $C(X)$. To prove boundedness, we estimate

$$|Tf(s)| \leq \int_X |k(s,t)||f(t)| \, d\mu(t) \leq \mu(X) \|k\|_\infty \|f\|_\infty.$$

Taking the supremum over $s \in X$, this results in

$$\|Tf\|_\infty \leq \mu(X) \|k\|_\infty \|f\|_\infty.$$

It follows that T is bounded and $\|T\| \leq \mu(X) \|k\|_\infty$.

[3]This is because in general the unit ball is not a compact set in a Banach space.

[4]Which can be made rigorous

Example 4.12. If $k \in L^2(K \times K, \mu \times \mu)$, then the same formula for T yields a bounded operator $T : L^2(K, \mu) \to L^2(K, \mu)$ satisfying $\|T\| \le \|k\|_2$. Indeed, by the Cauchy–Schwarz inequality (to be proven for Hilbert spaces later on) in and Fubini's theorem we obtain

$$\begin{aligned}\int_K \left| \int_K k(s,t) f(t)\, d\mu(t) \right|^2 d\mu(s) &\le \int_K \left(\int_K |k(s,t)|^2 d\mu(t) \right) \left(\int_K |f(t)|^2 d\mu(t) \right) d\mu(s) \\ &= \|k\|_2^2 \|f\|_2^2\end{aligned}$$

We consider an example of an integral operator called the Volterra integral operator operator. Consider the integral operator $T : L^2[0,1] \to L^2[0,1]$

$$Tf(s) := \int_0^1 k(s,t) f(t)\, dt = \int_0^1 \mathbb{1}_{(0,s)}(t) f(t)\, dt = \int_0^s f(t)\, dt, \quad s \in [0,1],$$

For all $f \in L^2(0,1)$, the Cauchy–Schwarz inequality implies that the indefinite integral is well defined and that

$$|Tf(s) - Tf(s')| \le |s - s'|^{1/2} \|f\|_2 \quad \text{for all } s, s' \in [0,1].$$

From this, we infer that $Tf \in C[0,1]$. Hence, the Volterra integral operator is actually an integral operator from $L^2[0,1])$ to $C[0,1]$. A bound on the norm of the Volterra integral operator is obtained by applying the bound of the preceding example:

$$\|T\| \le \|k\|_2 = \frac{1}{\sqrt{2}} \approx 0.7071\ldots$$

5. Finite-Dimensional Spaces

Linear algebra studies finite dimensional vector spaces. We now discuss some basic properties of finite dimensional vector spaces. In particular, we show that every finite dimensional vector space is a Banach space. Therefore, linear algebra can be thought of as the study of finite dimensional Banach spaces.

Remark 5.1. *It is well-known that every finite dimensional vector space admits an inner product. Therefore, linear algebra can be thought of as the study of finite dimensional Hilbert spaces. Hilbert spaces will be discussed later on.*

Example 5.2. It is well-know that any finite dimensional vector space is isometrically isomorphic to $\mathbb{K}^n$ for some $n \ge 1$. Moreover, it is a standard fact that any two norms on $\mathbb{K}^n$ are equivalent. Therefore, we can think of $\mathbb{K}^n$ as being endowed with the norm

$$\|a\|_2 := \left(\sum_{k=1}^n |a_k|^2 \right)^{1/2}$$

With this norm, $\mathbb{K}^n$ is a Banach space. This follows from the observation that if $S_n = \{1, \cdots, n\}$ and $\mu = \#$ is the counting measures on X, then

$$\mathcal{L}^2(S_n, \mathcal{P}(S_n), \#) = \mathbb{K}^n$$

Finite-dimensional Banach spaces have properties that are not automatically carried over to infinite-dimensional Banach spaces. We discuss two such properties below:

Proposition 5.3. *Every linear operator from a finite-dimensional normed space X into a normed space Y is bounded.*

Proof. Let $(x_j)_{j=1}^d$ be a basis for X. If $T : X \to Y$ is linear, for $x = \sum_{j=1}^d c_j x_j$ we obtain, by the Cauchy–Schwarz inequality for $\mathbb{K}^n$

$$\begin{aligned}\|Tx\| &= \left\|\sum_{j=1}^d c_j T x_j\right\| \\ &\leq \sum_{j=1}^d |c_j| \|Tx_j\| \leq M d^{1/2} \|x\|_2,\end{aligned}$$

where $\|x\|_2 := \left(\sum_{j=1}^d |c_j|^2\right)^{1/2}$ and $M := \max_{1 \leq n \leq d} \|Tx_n\|$. Since all norms are equivalent on X, there exists a constant $K > 0$ such that $\|x\|_2 \leq K\|x\|$ for all $x \in X$. Combining this with the preceding estimate we obtain

$$\|Tx\| \leq M d^{1/2} \|x\|_2 \leq K M d^{1/2} \|x\|.$$

This shows that that T is bounded with norm at most $KMd^{1/2}$. □

Remark 5.4. *If X is an infinite-dimensional Banach space, discotinuous linear functionals can be produced on X using Zorn's lemma and the Hahn-Banah theorem.*

Proposition 5.5. *Let X be a normed space.*

(1) *(**Riesz's Lemma**) If Y is a proper closed subspace of a normed space X, then for every $\epsilon > 0$ there exists a norm one vector $x \in X$ with $d(x, Y) \geq 1 - \epsilon$.*
(2) *The unit ball of a normed space X is relatively compact if and only if X is finite-dimensional.*

Proof. The proof is given below:

(1) Fix any $x_0 \in X \setminus Y$; such x_0 exists since Y is a proper subspace of X. Fix $\epsilon > 0$ and choose $y_0 \in Y$ such that
$$\|x_0 - y_0\| \leq (1 + \epsilon) d(x_0, Y)$$
The vector $(x_0 - y_0)/\|x_0 - y_0\|$ has norm one, and for all $y \in Y$ we have
$$\begin{aligned}\left\|\frac{x_0 - y_0}{\|x_0 - y_0\|} - y\right\| &= \frac{\|x_0 - y_0 - y\|}{\|x_0 - y_0\|} \\ &\geq \frac{d(x_0, Y)}{(1 + \epsilon) d(x_0, Y)} \\ &= \frac{1}{1 + \epsilon}.\end{aligned}$$
It follows that
$$d\left(\frac{x_0 - y_0}{\|x_0 - y_0\|}, Y\right) \geq \frac{1}{1 + \epsilon}.$$
Since $(1 + \epsilon)^{-1} \to 1$ as $\epsilon \downarrow 0$, this completes the proof.
(2) Clearly, every bounded subset of a finite-dimensional normed space X is relatively compact. Conversely, suppose that X is infinite-dimensional and pick an arbitrary norm one vector $x_1 \in X$. Proceeding by induction, suppose that norm one vectors $x_1, \ldots, x_n \in X$ have been chosen such that $\|x_k - x_j\| \geq \frac{1}{2}$ for all $1 \leq j \neq k \leq n$. Choose a norm one vector $x_{n+1} \in X$ by applying (1) to the proper closed subspace[5]
$$Y_n = \operatorname{span}\{x_1, \ldots, x_n\}$$

[5] A finite-dimensional subspace of a Banach space is closed.

and $\epsilon = \frac{1}{2}$ Then $\|x_{n+1} - x_j\| \geq \frac{1}{2}$ for all $1 \leq j \leq n$. The resulting sequence $(x_n)_{n\geq 1}$ is contained in the closed unit ball of X and satisfies

$$\|x_j - x_k\| \geq \frac{1}{2}$$

for all $j \neq k \geq 1$, so $(x_n)_{n\geq 1}$ has no convergent subsequence. It follows that the closed unit ball of X is not compact.

This completes the proof. $\square$

Part 2. Hilbert Spaces

Arguably the most significant examples of Banach spaces are the Hilbert spaces. These spaces provide the fundamental framework for many areas of mathematics and physics, particularly in quantum mechanics and functional analysis.

6. Definitions & Examples

We define Hilbert spaces and discuss several important examples. Along the way, we highlight key features of Hilbert spaces that are absent in general Banach spaces.

Definition 6.1. Let X be a $\mathbb{K}$-vector space. X is an inner product space is there is a map
$$\langle , \rangle : H \times H \to \mathbb{K}$$
having the following properties:

(1) **(Bilinearity)**: For $x, x', y \in X$ and $\lambda \in \mathbb{K}$
$$\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$$
$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$$
(2) **(Skew-Symmetry)**: $\langle y, x\rangle = \overline{\langle x, y\rangle}$
(3) **(Positivity)**: $\langle x, x\rangle > 0$ for $x \neq 0$ and $\langle x, x\rangle = 0$ if and only if $x = 0$.

Remark 6.2. *A simple observation shows that we must have*
$$\langle x, \lambda(y + y')\rangle = \overline{\lambda}\langle x, y\rangle + \overline{\lambda}\langle x, y'\rangle$$

To equip inner product spaces with the structure of normed vector spaces, the following inequality is essential.

Theorem 6.3. ***(Cauchy-Schwarz)*** *If X is an inner product space with inner product $\langle\cdot, \cdot\rangle$, then*
$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$$
for all x and y in X. Moreover, equality occurs if and only if x and y are linearly dependent.

Proof. If $\alpha \in \mathbb{K}$ and $x, y \in X$, then
$$0 \leq \langle x - \alpha y, x - \alpha y\rangle = \langle x, x\rangle - \alpha\langle y, x\rangle - \bar{\alpha}\langle x, y\rangle + |\alpha|^2\langle y, y\rangle.$$
Suppose $\langle y, x\rangle = be^{i\theta}$, $b \neq 0$, and let $\alpha = te^{-i\theta}$, $t \in \mathbb{R}$. The above inequality becomes
$$0 \leq \langle x, x\rangle - 2bt + t^2\langle y, y\rangle = c - 2bt + at^2 \equiv q(t),$$
where $c = \langle x, x\rangle$ and $a = \langle y, y\rangle$. Thus, $q(t)$ is a quadratic polynomial in the real variable t, and $q(t) \geq 0$ for all t. This implies that the equation $q(t) = 0$ has at most one real solution t. From the quadratic formula, we find that the discriminant is non-positive, i.e., $4b^2 - 4ac \leq 0$. Hence,
$$b^2 - ac = |\langle x, y\rangle|^2 - \langle x, x\rangle\langle y, y\rangle \leq 0$$
proving the inequality. It is clear that the equality holds if x and y are linearly dependent. If x and y are not linearly dependent, then there must be a vector $z \perp y$ and a non-zero scalar a such that $x = ay + z$[6], in which case
$$\langle x, x\rangle\langle y, y\rangle = (a^2\langle y, y\rangle + \langle z, z\rangle) \cdot \langle y, y\rangle = a^2\langle y, y\rangle^2 + \langle z, z\rangle\langle y, y\rangle$$
whereas
$$|\langle x, y\rangle|^2 = a^2\langle y, y\rangle^2$$
so strict inequality holds. □

[6]By Gram-Schmidt for finite-dimensional inner product spaces.

Corollary 6.4. *If X is an inner product space with an inner product $\langle \cdot, \cdot \rangle$ and if we define*

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

for all $x \in X$, then:

(1) $\|x + y\| \leq \|x\| + \|y\|$ *for all $x, y \in X$,*
(2) $\|\alpha x\| = |\alpha| \|x\|$ *for $\alpha \in k$ and $x \in X$.*
(c) $\|x\| = 0$ *implies $x = 0$.*

Hence, X is in particular a normed space with norm $\| \cdot \|$.

Proof. (2) and (3) are clear. To see (1), note that for $x, y \in X$,

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

The inequality now follows by taking the square root. □

Inner product spaces possess properties that are not shared by every normed vector space, or even by all Banach spaces. It is these special characteristics that distinguish Hilbert spaces and form the basis of their rich structure. In particular, we will discuss the parallelogram law and the notion of strict convexity, which highlight some of these unique features.

Proposition 6.5. *Let X be an inner product space.*

(1) ***(Parallelogram Law)*** *If x and $y \in X$, then*
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$
(2) ***(Strict Convexity)*** *X is strictly convex. That is, For $x, y \in X$ such that $\|x\| = \|y\| = 1$ with $x \neq y$ and $0 < \lambda < 1$ we have*
$$\|(1 - \lambda)x + \lambda y\| < 1$$

Proof. The proof is given below:

(1) For any x and y in X, we have:
$$\begin{aligned}
\|x + y\|^2 &= \|x\|^2 + 2\operatorname{Re}\langle x, y \rangle + \|y\|^2, \\
\|x - y\|^2 &= \|x\|^2 - 2\operatorname{Re}\langle x, y \rangle + \|y\|^2.
\end{aligned}$$
Adding the two equations yields the desired result.
(2) Note that
$$\begin{aligned}
&\lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \|\lambda x + (1 - \lambda)y\|^2 \\
=&\lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \big(\lambda^2\|x\|^2 + (1 - \lambda)^2\|y\|^2 + 2\lambda(1 - \lambda)\operatorname{Re}\langle x, y \rangle\big) \\
=&(\lambda - \lambda^2)\|x\|^2 + \underbrace{\big((1 - \lambda) - (1 - \lambda)^2\big)}_{=\lambda - \lambda^2}\|y\|^2 - 2\lambda(1 - \lambda)\operatorname{Re}\langle x, y \rangle \\
=&\lambda(1 - \lambda)\big(\|x\|^2 + \|y\|^2 - 2\operatorname{Re}\langle x, y \rangle\big) \\
=&\lambda(1 - \lambda)\underbrace{\|x - y\|^2}_{\neq 0} > 0
\end{aligned}$$
That is
$$\|\lambda x + (1 - \lambda)y\|^2 < \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 = 1,$$

Hence, X is strictly convex.

This completes the proof. $\square$

Remark 6.6. *If* $(X, \|\cdot\|)$ *is a Banach space, then the associated norm* $\|\cdot\|$ *need not satisfy the parallelogram law*[7]*. Consider* $X = C([0,1])$ *Consider the functions* $f(x) = 1 - x$ *and* $g(x) = x$*. We have*

$$\|f - g\|_\infty^2 + \|f + g\|_\infty^2 = \|1 - 2x\|_\infty^2 + \|1\|_\infty^2 = 1 + 1 = 2$$

but

$$2(\|f\|_\infty^2 + \|g\|_\infty^2) = 2\,\|1 - x\|_\infty^2 + 2\,\|x\|_\infty^2 = 2 + 2 = 4$$

Similarly, a Banach space need not be strictly convex. Consider $X = l^1(\mathbb{R})$ *and consider*

$$x = (1, 0, 0, \cdots,) \qquad y = (0, 1, 0, \cdots,)$$

Then $\|x\|_1 = \|y\|_1 = 1$ *but for each* $0 < \lambda < 1$*, we have*

$$\|(1 - \lambda)x + \lambda y\|_1 = \|(1 - \lambda, \lambda, 0, \cdots,)\|_1 = 1$$

We are now in a position to define Hilbert spaces.

Definition 6.7. A Hilbert space, H, is a Banach space together with an inner product $\langle \cdot, \cdot \rangle$.

Thus, Hilbert spaces are special Banach spaces endowed with an inner product that induces their norm. We now examine examples of Hilbert spaces, many of which arise from Banach spaces considered previously.

Example 6.8. The following is a list of examples of Hilbert spaces.

(1) $\mathbb{K}^n$ is a Hilbert space. Indeed, we can define an inner product on $\mathbb{K}^n$ as:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$

It quite easy to verify that properties in Definition 6.1 are satisfied. It is a standard fact that $\mathbb{K}^n$ is a complete normed space under the norm induced by $\langle , \rangle$ is complete.

(2) Consider

$$\ell^2(\mathbb{N}) = \{(x_n)_{n \geq 1} \in \mathbb{K}^\infty : \|x\|_2 < \infty\}$$

Then, $\ell^2(\mathbb{N})$ is a Hilbert space. Indeed, we can define an inner product on $\ell^2(\mathbb{N})$ as:

$$\langle x, y \rangle = \sum_{n \geq 1} x_n \overline{y_n}$$

We claim the expression above is finite. Since $\mathbb{K}^n$ is an inner product space, Theorem 6.3 implies that:

$$\begin{aligned} \sum_{n=1}^{M} x_n y_n &\leq \left(\sum_{n=1}^{M} |x_n|^2 \right)^{1/2} \left(\sum_{n=1}^{M} |y_n|^2 \right)^{1/2} \\ &\leq \left(\sum_{n=0}^{\infty} |x_n|^2 \right) \left(\sum_{n=0}^{\infty} |y_n|^2 \right) < \infty \end{aligned}$$

Letting $M \to +\infty$, we get:

$$\sum_{n \geq 1} x_n \overline{y_n} \leq \left(\sum_{n=0}^{\infty} |x_n|^2 \right) \left(\sum_{n=0}^{\infty} |y_n|^2 \right) < \infty$$

[7] In fact, it can be proved that a normed vector space is an inner product space if and only if the norm satisfies the parallelogram law.

It quite easy to verify that properties in Definition 6.1 are satisfied.

(3) Let $(X, \mathscr{M}, \mu)$ be a measurable space. Then

$$L^2(X) = \left\{ f : X \to \mathbb{K} : f \text{ is measurable and } \|f\|_2 < \infty \right\}$$

is a Hilbert space. Indeed, we can define an inner product on $L^2(X)$ as:

$$\langle f, g \rangle = \int_X f\overline{g} d\mu < \infty$$

The expression above is finite. Indeed, since $f, \overline{g}$ and $f + \overline{g}$ belong to $L^2(X)$, we have that $|f|^2, |\overline{g}|^2$ and

$$(f + \overline{g})^2 = f^2 + 2f\overline{g} + \overline{g}^2$$

Hence, $f\overline{g}$ is integrable, as required. It quite easy to verify that properties in Definition 6.1 are satisfied. The norm induced by $\langle , \rangle$ is the one discussed above. Hence, $\ell^2(\mathbb{N})$ is a Hilbert space. The norm induced by $\langle , \rangle$ is the one discussed above. Hence, $L^2(X)$ is a Hilbert space.

(4) We can generalize Example 6.8(2). Let A be a possibly uncountable set. Let $\mathscr{M} = \mathcal{P}(A)$ and $\mu = \#$ be the counting measure. In this case, recall that

$$\|f\|_2^2 = \int_X |f|^2 d\mu = \sum_{x \in A} |f(x)|^2$$

where the summation over a possibly uncountable set can be thought of as[8]

$$\sum_{x \in A} |f(x)|^2 = \sup \left\{ \sum_{i \in F} |x_i|^2 : F \subseteq A, \ F \text{ finite} \right\}$$

It can be checked that

$$\begin{aligned} L^2(A) &= \{ f : A \to \mathbb{K} : \|f\|_2 < \infty \} \\ &= \left\{ (x_i)_{i \in A} : \sup \left\{ \sum_{i \in F} |x_i|^2 : F \subseteq A, \ F \text{ finite} \right\} < \infty \right\} \end{aligned}$$

is a Hilbert space. We label this Hilbert space as $l^2(A)$.

Remark 6.9. *The Cauchy-Schwartz inequality for $L^2(X)$ states that:*

$$|\langle f, g \rangle| = \left| \int_X f\overline{g} d\mu \right| \leq \Big(\int |f|^2 \, d\mu \Big)^{1/2} \Big(\int |g|^2 \, d\mu \Big)^{1/2} = \|f\|_2 \|g\|_2$$

In particular, applying the Cauchy-Schwartz inequality to $|f|$ and $|g|$, we have,

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2$$

This is a special case Holder's inequality (proof omitted) which states that

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

if $f \in L^p(X), g \in L^q(X)$ and $1/p + 1/q = 1$.

[8] One can also use the the concept of nets from general topology to make sense of a sum over a possibly uncountable set. We shall not delve in these details.

Definition 6.10. Let H and K be Hilbert space. A linear operator $T : H \to K$ is bounded if it is bounded as a linear operator on the underlying Banach spaces. Moreover, T is an isomorphism if T is a bijective linear map and that

$$\langle x, y\rangle_H = \langle Tx, Ty\rangle_K$$

for each $x, y \in H$[9].

As in the case of Banach spaces, one can make analytic arguments in infinite-dimensional Hilbert spaces since it is a complete metric space. Here is a sample analytic argument we can make:

Theorem 6.11. *Let H be a Hilbert space with an inner product $\langle , \rangle$ Then $\langle x, y\rangle$ is jointly continuous as a function of x and y.*

Proof. It suffices to show that if $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$. We have

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &\le |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &= |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\le \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|. \end{aligned}$$

Since convergent sequences are bounded, the number $M := \sup_{n>1} \|x_n\|$ is finite, and we find

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| \le M\|y_n - y\| + \|x_n - x\|\|y\|.$$

Both terms on the right-hand side tend to 0 as $n \to \infty$. □

7. Orthogonality, Best Approximation & Projections

We explore the geometric structure of Hilbert spaces through the notions of orthogonality and projection. These concepts play a fundamental role in understanding best approximation problems and the decomposition of elements in Hilbert spaces.

7.1. **Orthogonality.**

The greatest advantage of a Hilbert space is its underlying concept of orthogonality which is induced by its underlying inner product.

Definition 7.1. If H is a Hilbert space and $x, y \in H$, then x and y are orthogonal if $\langle x, y\rangle = 0$. We write $x \perp y$.

Here are some consequences of the notion of orthogonality which are similar to notions in classical Euclidean geometry.

Proposition 7.2. ***(Pythagorean Theorem)***[10] *Let H be a Hilbert space. If $x_1, \cdots, x_n$ are pairwise orthogonal vectors in H, then*

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

Proof. If $x_1, x_2 \in \mathcal{H}$, then

$$\|x_1 + x_2\|^2 = \langle x_1 + x_2, x_1 + x_2\rangle = \|x_1\|^2 + 2\,\mathrm{Re}\langle x_1, x_2\rangle + \|x_2\|^2$$

Since $x_1 \perp x_2$, this implies the result for $n = 2$. The general case follows by an easy inductive argument. □

---

[9]Note that the condition $\langle x, y\rangle_H = \langle Tx, Ty\rangle_K$ actually implies that T is injective. Moreover, it turns out that T is an isomorphism if and only if T is $\|Tx\|_K = \|x\|_H$ for each $x \in H$. Moreover, this definition will be sufficient for our purposes by now because we later on we shall show that T is bijective if and only if T^{-1} is a bijective, bounded linear map. This will be a consequence of the Open Mapping Theorem.

[10]The proof doesn't use the completeness of the underlying Hilbert space. It only uses the properties of the inner product.

Definition 7.3. The orthogonal complement of a subset A of H is the set

$$A^{\perp} := \{x \in H : x \perp a \text{ for all } a \in A\}$$

The orthogonal complement $A^{\perp}$ of a subset A is a closed subspace of H. Indeed, it is trivially checked that $A^{\perp}$ is a vector space. To prove its closedness, let $x_n \to x$ in H with $x_n \in A^{\perp}$. By Theorem 6.11, we obtain

$$\langle x, a\rangle = \lim_{n\to\infty} \langle x_n, a\rangle = 0.$$

for all $a \in A$.

Example 7.4. The following are some computations of orthogonal complements:

(1) Let

$$Y := \{f \in L^2(0,1) : f(t) = 0 \text{ for almost all } t \in (0, 1/2)\}$$

We compute $Y^{\perp}$. If $g \in Y^{\perp}$, then

$$\int_{\frac{1}{2}}^{1} f(t)g(t)\mathrm{d}t = 0$$

for each $f \in Y$. In particular, if $f(t) = \chi_A(t)$ for each Borel set in $[1/2, 1)$, then

$$\int_A g(t)\mathrm{d}t = 0$$

for each Borel set in $[1/2, 1)$. Hence $g(t) = 0$ for almost all $t \in [1/2, 1)$. Conversely, every such function is in $Y^{\perp}$. Hence, $Y^{\perp}$ consists of all functions g such that $g(t) = 0$ for almost all $t \in [1/2, 1)\}$.

(2) Let

$$Y := \{f \in L^2(0,1) : \int_0^1 f(t)\, dt = 0\}$$

Take any $g \in Y^{\perp}$. Set $\overline{g} = \int_0^1 g(t)dt$, and take

$$f(x) = g(x) - \overline{g}$$

We have $f \in Y$ and so

$$0 = \langle f, g\rangle = \langle f, g\rangle - \langle f, \overline{g}\rangle = \langle f, g - \overline{g}\rangle = \langle g - \overline{g}, g - \overline{g}\rangle = \|g - \overline{g}\|_2^2.$$

Thus $g = \overline{g}$, and so g is constant. Thus $Y^{\perp}$ contains all constant functions. The argument seems a bit artificial, but the conclusion follows naturally once we know that complex exponentials form an orthonormal basis for $L^2(0,1)$.

7.2. **Best Approximation.** The most important result on orthogonality is certainly the fact that every closed subspace Y of a Hilbert space is orthogonally complemented by $Y^{\perp}$. For its proof, we need the the approximation theorem for convex closed sets in Hilbert space, which is of independent interest.

Proposition 7.5. ***(Best Approximation)*** *Let C be a non-empty convex closed subset of H. Then, for all $x \in H$, there exists a unique $c \in C$ that minimizes the distance from x to the points of C:*

$$\|x - c\| = \min_{y\in C} \|x - y\|.$$

Proof. By considering $C - x = \{c - x : c \in K\}$ instead of C, it suffices to assume that $x = 0$. We how there is a unique vector c in C such that

$$\|c\| = \inf\{\|c\| : c \in C\} := d.$$

By definition, there is a sequence $\{c_n\}$ in C such that $\|c_n\| \to d$. Proposition 6.5 implies:

$$\left\|\frac{c_n - c_m}{2}\right\|^2 = \frac{\|c_n\|^2 + \|c_m\|^2}{2} - \left\|\frac{c_n + c_m}{2}\right\|^2.$$

Since C is convex, $\frac{1}{2}(c_n + c_m) \in K$. Hence,

$$\left\|\frac{1}{2}(c_n + c_m)\right\|^2 \leq d^2.$$

If $\varepsilon > 0$, choose N such that for $n \geq N$,

$$\|c_n\|^2 < d^2 + \frac{\varepsilon^2}{4}.$$

If $n, m \geq N$, then

$$\left\|\frac{c_n - c_m}{2}\right\|^2 < \frac{1}{2}\left(2d^2 + \frac{\varepsilon^2}{2}\right) - d^2 = \frac{\varepsilon^2}{4}.$$

Thus, $\|c_n - c_m\| < \varepsilon$ for $n, m \geq N$, and $\{c_n\}$ is a Cauchy sequence. Since H is complete and C is closed, there is a $c \in C$ such that $\|c_n - c\| \to 0$. Also, for all c_n,

$$d \leq \|c\| = \|c - c_n + c_n\| \leq \|c - c_n\| + \|c_n\| \to d.$$

Thus, $\|c\| = d$. To prove that c is unique, suppose $c' \in K$ such that $\|c'\| = d$. By convexity, $\frac{1}{2}(c + c') \in K$. Hence,

$$d \leq \left\|\frac{1}{2}(c' + c)\right\| \leq \frac{1}{2}\|c'\| + \frac{1}{2}\|c\| = d.$$

So, $\|\frac{1}{2}(c' + c)\| = d$. Proposition 6.5 implies that:

$$\left\|\frac{c - c'}{2}\right\|^2 = \frac{\|c\|^2 + \|c'\|^2}{2} - \left\|\frac{c + c'}{2}\right\|^2 = d^2 - d^2 = 0.$$

Hence,

$$\|c - c'\| = 0,$$

implying that $c' = c$. $\square$

Remark 7.6. *Proposition 7.5 can fail for Banach spaces. Consider $X = l^\infty(\mathbb{R})$ and consider the convex set.*

$$C = \{(x, 1, 0, \cdots,) : x \in [-1, 1]\}$$

Let $x = (0, 0, \cdots)$ Each element in C has norm 1. Hence,

$$\{c \in C : \|c\| = \min_{y \in C} \|y\|\| = C$$

If the closed, convex set in Proposition 7.5 is in fact a closed linear subspace of H, more can be said. For $x \in H$, let $x_0 \in C$ such that

$$\|x - x_0\| = \min_{y \in C} \|x - y\|$$

We claim that $x - x_0 \in C^\perp$. Fix a nonzero $c \in C$. For any $\lambda \in k$ we have,

$$\begin{aligned}
\|x - x_0\|^2 &\leq \|x - (x_0 - \lambda c)\|^2 \\
&= \|\lambda c - (x - x_0)\|^2 \\
&= |\lambda|^2\|c\|^2 + 2\operatorname{Re}(\lambda c, x - x_0) + \|x - x_0\|^2.
\end{aligned}$$

Taking $\lambda = -\frac{\overline{\langle c, x - x_0 \rangle}}{\|c\|^2}$, this gives

$$0 \leq \frac{|(c, x - x_0)|^2}{\|y\|^2} - 2\frac{|(c, x - x_0)|^2}{\|y\|^2}$$

which is only possible if $\langle x - x_0, c\rangle$. This shows that $x - x_0 \in C^{\perp}$. Conversely, suppose $c \in C$ such that $x - c \in C^{\perp}$. If $c' \in C$, then $x - c \perp x - c'$ so that

$$\|x - c'\|^2 = \|(x - c) + (c - c')\|^2 = \|x - c\|^2 + \|c - c'\|^2 \geq \|x - c\|^2.$$

Thus

$$\|x - c\| = \min_{c' \in C} \|x - y\|$$

7.3. **Projections.** The discussion above allows us to define a projection operator on a closed subspace, M, in a Hilbert space. This projection operator minimizes the distance from $x \in H$ to M.

Proposition 7.7. *Let H be a Hilbert space and let M be a closed linear subspace of M. Define a map*

$$\pi_M : H \to M$$

such that $\pi_M(x)$ is the unique point in M such that $x - \pi_M(x) \perp M$. Then

(1) *π_M is a linear transformation on H,*
(2) *$\|\pi_M(x)\| \leq \|x\|$ for every $x \in H$,*
(3) *$\pi_M^2 = \pi_M$,*
(4) *$\ker \pi_M = M^{\perp}$ and $\operatorname{ran} \pi_M = M$.*

π_M is the projection map onto M.

Proof. The proof is given below:

(1) Let $x_1, x_2 \in H$ and $\alpha_1, \alpha_2 \in \mathbb{K}$. If $f \in M$, then

$$\begin{aligned}\langle(\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 \pi_M(x_1) + \alpha_2 \pi_M(x_2)), f\rangle &= \alpha_1(x_1 - \pi_M(x_1), f) + \alpha_2(x_2 - \pi_M(x_2), f) \\ &= 0.\end{aligned}$$

By uniqueness, we have that

$$\pi_M(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \pi_M(x_1) + \alpha_2 \pi_M(x_2)$$

(2) If $x \in X$, then $x = (x - \pi_M(x)) + \pi_M(x)$, $\pi_M(x) \in M$, and $h - \pi_M(x) \in M^{\perp}$. Thus

$$\|x\|^2 = \|x - \pi_M(x)\|^2 + \|\pi_M(x)\|^2 \geq \|\pi_M(x)\|^2.$$

(3) If $x \in M$, then $\pi_M(x) = x$. Hence $\pi_M^2 = \pi_M$.
(4) If $\pi_M(x) = 0$, then $x = x - \pi_M(x) \in M^{\perp}$. Conversely, if $x \in M^{\perp}$, then 0 is the unique vector in M such that $x - 0 = x \in M^{\perp}$. Therefore $\pi_M(x) = 0$. Clearly, $\operatorname{ran} \pi_M = M$.

This completes the proof. □

Corollary 7.8. *Let H be a Hilbert space and let M be a closed subspace of H. Then*

(1) *$H = M \oplus M^{\perp}$*
(2) *$(M^{\perp})^{\perp} = M$*
(3) *More generally, if M is any subspace of H, then $(M^{\perp})^{\perp} = \overline{M}$.*
(4) *M is dense in H if and only if $M^{\perp} = \{0\}$*

Proof. The proof is given below:

(1) Let $x \in H$ and write x as $x = \pi_M(x) + (x - \pi_M(x))$. We know that $\pi_M(x) \in M$. Moreover,

$$\pi_M(x - \pi_M(x)) = \pi_M(x) - \pi_M^2(x) = \pi_M(x) - \pi_M(x) = 0$$

Hence $x - \pi_M(x) \in M^{\perp}$. It is clear that $M \cap M^{\perp} = \{0\}$. Hence, $H = M \oplus M^{\perp}$.
(2) Note that $I - \pi_M$ is an orthogonal projection onto $M^{\perp}$. By part (d) of the preceding theorem, $(M^{\perp})^{\perp} = \ker(I - \pi_M)$. But $0 = (I - \pi_M)x$ if and only if $x = \pi_M x$. Thus $(M^{\perp})^{\perp} = \ker(I - \pi_M) = \operatorname{ran} \pi_M = M$.

(3) Note that $(M^\perp)^\perp$ is a closed subspace containing M. Hence, $\overline{M} \subseteq (M^\perp)^\perp$. If C is any closed subspace of H containing M, we have

$$M \subseteq C \iff C^\perp \subseteq M^\perp \iff (M^\perp)^\perp \subseteq (C^\perp)^\perp = C$$

Hence, $M^\perp)^\perp$ is the smallest closed subspace of H containing M. Hence, $(M^\perp)^\perp = \overline{M}$.

(4) Simply note that

$$\overline{M} = H \iff \{0\} = H^\perp = (\overline{M})^\perp = ((M^\perp)^\perp)^\perp = M^\perp$$

This completes the proof. □

8. Orthornormal Systems

Just as in the finite-dimensional setting of Euclidean space, every Hilbert space admits a notion of coordinatization. This is achieved through the use of orthonormal sets, which serve as a foundation for representing elements in terms of scalar components. Orthonormal systems not only facilitate the construction of orthogonal projections but also enable the expansion of vectors in terms of basis elements in infinite-dimensional settings.

Definition 8.1. Let I be a non-empty set. A family $(h_i)_{i\in I}$ in H is called an orthonormal system if

$$\langle h_i, h_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

An orthonormal system is a maximal orthonormal system if M is the linear span of $(h_i)_{i\in I}$, then $\overline{M} = H$. That is, every $x \in H$ can be represented as a convergent series

$$x = \sum_{i\in I} c_i h_i$$

for suitable coefficients $c_i \in k$.

Remark 8.2. *Recall from Corollary 7.8 that $\overline{M} = H$ if and only if $M^\perp = \{0\}$. Therefore, an orthonormal system is a maximal orthonormal system if and only if*

$$\langle h, h_i \rangle = 0 \textit{ for all } i \in I \textit{ implies that } h = 0\ .$$

Proposition 8.3. *Let H be a non-zero Hilbert space. Then H has an maximal orthonormal system.*

Proof. (**Sketch**) This follows by Zorn's Lemma. Partially order the set of all orthonormal systems in the nonzero Hilbert space H by set inclusion. By Zorn's lemma, this set has a maximal element, say $(h_i)_{i\in I}$, where I is some index set. It is clear that this set is an orthonormal set. If there were a nonzero $h \in H$ such that $\langle h, h_i \rangle = 0$, then after normalizing h to unit length, we obtain a new orthonormal system $(h_i)_{i\in I} \cup \{h\}$ properly containing $(h_i)_{i\in I}$, contradicting the maximality of $(h_i)_{i\in I}$. Therefore, the orthonormal system is a maximal orthonormal system. □

A countable maximal orthonormal set is also called an orthonormal basis. Intuitively, Hilbert spaces that admit an orthonormal basis are of the form '$\mathbb{K}^\infty$.' We shall make this intuition precise in Corollary 8.6. Orthonormal sets are quite tractable since they have a number of simple properties. All of these properties fundamentally arise from the ability to define projection operators onto subspaces determined by finite subsets of an orthonormal system. By employing such projections, one can systematically construct elements of the Hilbert space H as limits of these finite approximations.

Proposition 8.4. *Let H be a Hilbert space and let $(h_n)_{n\geq 1}$ be an orthonormal sequence in H. Let $(c_n)_{n\geq 1}$ be a sequence of scalars in $\mathbb{K}$. Then:*

(1) *$\sum_{n\geq 1} c_n h_n$ converges in H if and only if $\sum_{n\geq 1} |c_n|^2 < \infty$.*

(2) *(**Bessel's Inequality**) For $x \in H$,*

$$\|x\|^2 \geq \sum_{n\geq 1} |\langle x, h_n\rangle|$$

(3) *Let M be the closed linear span of $(h_n)_{n\geq 1}$. Then for each $x \in H$*

$$\pi_M(x) = \sum_{n\geq 1} \langle x, h_n\rangle h_n$$

(4) *(**Parseval's Identity**) If $(h_n)_{n\geq 1}$ is an orthonormal basis, then*

$$\|x\|^2 = \sum_{n\geq 1} |\langle x, h_n\rangle|$$

Proof. The proof is given below:

(1) If $\sum_{n\geq 1} c_n h_n$ converges in H, say to x, then $x = \lim_{N\to\infty} \sum_{n=1}^N c_n h_n$ in H and therefore

$$\infty > \|x\|^2 = \lim_{N\to\infty} \sum_{n=1}^N \|h_n\|^2 = \lim_{N\to\infty} \sum_{n=1}^N \|c_n\|^2 = \sum_{n\geq 1} |c_n|$$

Conversely, suppose that $\sum_{n\geq 1} |c_n| < \infty$. Then

$$\lim_{\substack{M,N\to\infty \\ N>M}} \frac{1}{2} \left\| \sum_{n=1}^N c_n h_n - \sum_{n=1}^M c_n h_n \right\|^2 = \lim_{\substack{M,N\to\infty \\ N>M}} \frac{1}{2} \left\| \sum_{n=M+1}^N c_n h_n \right\|^2 = \lim_{\substack{M,N\to\infty \\ N>M}} \frac{1}{2} \sum_{n=M+1}^N |c_n|^2 = 0.$$

It follows that $\sum_{n\geq 1} c_n h_n$ is Cauchy, and hence convergent.

(2) Let $x_n = x - \sum_{k=1}^n (x, h_k) e_k$. Then $h_n \perp e_k$ for $1 \leq k \leq n$. By the Pythagorean Theorem (Proposition 7.2),

$$\infty > \|x\|^2 = \|x_n\|^2 + \sum_{k=1}^n |\langle x, h_k\rangle|^2 \leq \sum_{k=1}^n |\langle x, h_k\rangle|^2 \leq \sum_{n\geq 1}^\infty |\langle x, h_n\rangle|^2$$

(3) By (1) and (2),

$$\sum_{n\geq 1} \langle x, h_n\rangle h_n$$

is finite and well-defined. Call this expression $\pi_M(x)$. We have

$$\begin{aligned}\langle x - \pi_M(x), \pi_M(x)\rangle &= \lim_{N,M\to\infty} \left\langle x - \sum_{n=1}^N \langle x, h_n\rangle h_n, \sum_{m=1}^M \langle x, h_m\rangle h_n \right\rangle \\ &= \lim_{N,M\to\infty} \sum_{m=1}^M \overline{\langle x, h_m\rangle} \left\langle x - \sum_{n=1}^N \langle x, h_n\rangle h_n, h_m \right\rangle = \lim_{N,M\to\infty} 0 = 0\end{aligned}$$

because

$$\left\langle x - \sum_{n=1}^N \langle x, h_n\rangle h_n, h_m \right\rangle = \langle x, h_m\rangle - \left\langle \sum_{n=1}^N \langle x, h_n\rangle h_n, h_m \right\rangle = \langle x, h_m\rangle - \langle x, h_m\rangle = 0$$

Hence,

$$\pi_M(x) = \sum_{n\geq 1} \langle x, h_n\rangle h_n$$

(4) Assume $(h_n)_{n\geq 1}$ is an orthonormal basis of H. Let M is the linear span of $(h_n)_{n\geq 1}$. Then $\overline{M} = H$. Corollary 7.8 implies that $M^\perp = \{0\}$. Therefore,

$$x - \pi_M(x) = 0$$

Hence,

$$x = \pi_M(x) = \sum_{n\geq 1} \langle x, h_n\rangle h_n$$

The claim now follows from this observation.

This completes the proof. □

Using Proposition 8.4, we can now establish two key results. First, a Hilbert space is separable if and only if it admits a countable orthonormal basis. Second, this allows us to conclude that every separable Hilbert space is isometrically isomorphic to $\ell^2(\mathbb{N})$.

Proposition 8.5. *A Hilbert space, H, has an orthonormal basis if and only if it is separable.*

Proof. Assume H is seperable and let $(h_i)_{i\in I}$ be a maximal orthonormal set for some index set I. If h_i, h_j are elements in the maximal orthonormal set, then

$$\|h_i - h_j\|_2^2 = \|h_i\|_2^2 + \|h_j\|_2^2 = 2$$

Hence $\mathscr{C} = \{B(h_i; 1/\sqrt{2})\}_{i\in I}$ is a collection of pairwise disjoint open balls in H. Since H is separable, the collection $\mathscr{C}$ must be a countable collection. Hence, the maximal orthonormal set is in fact an orthonormal basis. Conversely, if H has an orthonormal basis then its linear span is dense in H and is generated by countably many elements. This is sufficient to conclude that H is seperable. □

Corollary 8.6. *Every non-zero separable Hilbert space, H, is isomorphic to $\ell^2(\mathbb{N})$. In this case, we say that* $\dim = |N| = \infty$.[11].

Proof. Let H be a non-zero separable Hilbert space with orthonormal basis $(h_n)_{n\geq 1}$. Let $T : H \to \ell^2(\mathbb{N})$ be defined for each $x \in H$ by:

$$T(x) = (\langle h_n, x\rangle)_{n\geq 1}$$

By Bessel's inequality in Proposition 8.4, we have:

$$\|Tx\|_2^2 = \|(\langle x_n, h\rangle)\|_2^2 = \sum_{n\geq 1} |\langle h_n, x\rangle^2| \leq \|x\|^2 < \infty$$

Hence, T is well-defined. It is clear that T is a linear operator. Moreover, we have that

$$\begin{aligned}
\langle x, y\rangle_H &= \left\langle \sum_{n\geq 1}\langle h_n, x\rangle h_n, \sum_{m\geq 1}\langle h_m, y\rangle h_m \right\rangle \\
&= \sum_{n\geq 1}\sum_{m\geq 1} \langle h_n, x\rangle \overline{\langle h_m, y\rangle} \langle e_n, e_m\rangle \\
&= \sum_{n\geq 1} \langle h_n, x\rangle \overline{\langle h_n, y\rangle} \\
&= \langle Tx, Ty\rangle_{\ell^2(\mathbb{N})}
\end{aligned}$$

[11]This statement justifies the intuition that every Hilbert spaces that admit an orthonormal basis is of the form "k^∞."

for each $x, y \in H$. Hence, T is inner-product preserving. Lastly, we show that T is surjective. Let $(x_n)_{n\geq 1}$ and consider

$$x = \sum_{n\geq 1} x_n h_n$$

Note that x converges by Proposition 8.4(a) and we have

$$\|x\|_2^2 = \sum_{n\geq 1} |x_n|^2 < \infty$$

We see that:

$$T(x) = (\langle e_n, x\rangle)_{n\geq 1} = \left(\left\langle e_n, \sum_{n\geq 1} x_n e_n \right\rangle\right) = (x_n\|e_n\|_2)_{n\geq 1} = (x_n)_{n\geq 1}$$

So T is surjective. Hence, T is an isomorphism. □

Note that every non-zero Hilbert space is not a separable space. Consider the Hilbert space:

$$l^2(\mathbb{R}) = \left\{(x_i)_{i\in\mathbb{R}} : \sup\left\{\sum_{i\in F} |x_i|^2 : F \subset \mathbb{R},\ F \text{ finite}\right\} < \infty\right\}$$

The functions f_y defined by

$$f_y(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{else} \end{cases}$$

are an uncountable set of elements with distance $\sqrt{2}$, hence $l^2(\mathbb{R})$ is not separable. More generally, we have the following generalization of Corollary 8.6.

Proposition 8.7. *Let H be a non-zero Hilbert space.*

(1) *H is isomorphic to $l^2(A)$ for some possibly uncountable set A. We say that* $\dim = |A|$.
(2) *$l^2(A)$ and $l^2(B)$ are isomorphic if and only if $|A| = |B|$. Hence, two Hilbert spaces are isomorphic if and only if they have the same dimension.*

Proof. The proof is given below:

(1) (Sketch) By Proposition 8.3, there is a maximal orthonormal system $\{h_\alpha\}_{\alpha\in A}$. Choose A to be the index set. Fix $x \in H$. By Remark 8.2, x can be represented as

$$x = \sum_{a\in A} c_\alpha h_\alpha$$

for $c_\alpha \in \mathbb{K}$. In fact, $c_\alpha = \langle x, h_\alpha\rangle$. We claim that the number of terms in the summation above are at most countable. Consider the set

$$\mathscr{E} = \{\alpha \in A : \langle x, h_\alpha\rangle \neq 0\} = \bigcup_{n\geq 1} \{\alpha \in A : |\langle x, h_\alpha\rangle| \geq 1/n\} := \bigcup_{n\geq 1} \mathscr{E}_n$$

Assume $\mathscr{E}_n$ is an infinite set. Pick a countably infinite sequence $(\beta_i)_{i\geq 1}$ in $\mathscr{E}_n$. Fix any $N \in \mathbb{N}$. Then:

$$x = \left(x - \sum_{n=1}^{N} (x, h_{\beta_i}) h_{\beta_i}\right) + \sum_{n=1}^{N} (x, h_{\beta_i}) h_{\beta_i}$$

is an orthogonal decomposition. Therefore,

$$\begin{aligned}\|x\|^2 &= \left\| x - \sum_{i=1}^{N} (x, h_{\beta_i}) h_{\beta_i} \right\|^2 + \left\| \sum_{i=1}^{N} (x, h_{\beta_i}) h_{\beta_i} \right\|^2 \\ &\geq \left\| \sum_{n=1}^{N} (x, h_{\beta_i}) h_{\beta_i} \right\|^2 \\ &= \sum_{i=1}^{N} |(x, h_{\beta_i})|^2 \geq \sum_{i=1}^{N} \frac{1}{n}\end{aligned}$$

Letting $N \to \infty$ yields a contradiction since $\|x\| < \infty$. Hence, each $\mathscr{E}_n$ is at most finite implying that $\mathscr{E}$ is at most countable. Let $\{\alpha_n\}$ be an enumeration of the $\alpha \in A$ for which $(x, e_n) \neq 0$. We have

$$\|x\|^2 = \left\| x - \sum_{n=1}^{N} (x, h_{\alpha_n}) h_{\alpha_n} \right\|^2 + \sum_{n=1}^{N} \|(x, h_{\alpha_n})\|^2.$$

Based on Remark 8.2, we have

$$\|x\|^2 = \lim_{N \to \infty} \left\| x - \sum_{n=1}^{N} (x, h_{\alpha_n}) h_{\alpha_n} \right\|^2 + \lim_{N \to \infty} \sum_{n=1}^{N} \|(x, h_{\alpha_n})\|^2 = \sum_{n \geq 1} \|(x, h_{\alpha_n})\|^2$$

Let $T : H \to l^2(A)$ be defined for each $x \in H$ by:

$$T(x) = (\langle h_\alpha, x \rangle)_{\alpha \in A}$$

Our discussion above implies that T is well-define since $(\langle h_\alpha, x \rangle)_{\alpha \in A} \in l^2(A)$ for each fixed $x \in H$. As in Corollary 8.6, one can check that T is a linear map that preserves the inner product and T is surjective[12].

(2) This is a straightforward consequence of (1).

This completes the proof. □

9. Hilbert Dual Space

We investigate the dual space of a Hilbert space. Unlike the general setting of Banach spaces, the presence of an inner product allows for a more concrete and elegant characterization of the dual space of a Hilbert space. This culminates in the Riesz Representation Theorem, which establishes a natural isomorphism between a Hilbert space and its dual.

Definition 9.1. Let H be a Hilbert space over the field $\mathbb{K}$. The (Hilbert) dual space of H, denoted by H^*, is the set of all continuous/bounded linear functionals $\ell : H \to \mathbb{K}$. That is,

$$H^* = \{\ell : H \to \mathbb{K} \mid \ell \text{ is linear and bounded}\}.$$

The dual space H^*, equipped with the operator norm, is a Banach space. Moreover, the inner product structure on H enables a more explicit description and a full characterization of H^*, as established by the Riesz Representation Theorem below.

Proposition 9.2. ***(Riesz Representation Theorem)*** *Let H be a Hilbert space. Then for every continuous linear functional $\ell \in H^*$, there exists a unique element $h \in H$ such that*

$$\ell(x) = \langle h, x \rangle \quad \textit{for all } x \in H.$$

[12] We will have to repeatedly use the observation that the norm of each element in $l^2(A)$ is determined by a sum over countably many indices. This argument is similar to what we have given above.

Moreover, the mapping

$$\begin{aligned} H &\to H^*, \\ h &\mapsto \langle h, \cdot \rangle \end{aligned}$$

is an isometric isomorphism.

Proof. Let $\ell \in H^*$, where we may assume WLOG that $\ell \neq 0$. It is clear that $\ker(\ell)$ is a proper closed subspace of H: Corollary 7.8 implies that $H = \ker(\ell) \oplus \ker(\ell)^\perp$. Let z be a non-zero element in $\ker(\ell)^\perp$. Consider the element

$$h := \frac{\ell(z) z}{\|z\|^2} \in H$$

We now show that $\ell = \langle h, \cdot \rangle$. Fix any $x \in H$, and define

$$w := x - \frac{\ell(x)}{\ell(z)} z.$$

Note that $\ell(w) = 0$. Hence, $w \in \ker(\ell)$ and $\langle w, z \rangle = 0$ implying that $\langle w, h \rangle = 0$. We have

$$\begin{aligned} \langle x, h \rangle &= \left\langle w + \frac{\ell(x)}{\ell(z)} z, h \right\rangle \\ &= \langle w, h \rangle + \left\langle \frac{\ell(x)}{\ell(z)} z, h \right\rangle \\ &= \frac{\ell(x)}{\ell(z)} \langle z, h \rangle = \frac{\ell(x)}{\ell(z)} \left\langle \frac{\ell(z) z}{\|z\|^2}, z \right\rangle = \ell(x). \end{aligned}$$

We now show uniqueness. Suppose $\ell = \langle h, \cdot \rangle = \langle h', \cdot \rangle$ for some $h, h' \in H$. We then have

$$\langle h, h - h' \rangle = \langle h', h - h' \rangle \implies \langle h - h', h - h' \rangle = 0 \implies \|h - h'\| = 0$$

Hence, $h = h'$. This shows that the map

$$\begin{aligned} H &\to H^*, \\ h &\mapsto \langle h, \cdot \rangle \end{aligned}$$

is a bijection. It is clear that the map is linear. We claim that the map is an isometry as well. By Cauchy-Schwartz, $\|\langle h, \cdot \rangle\| \leq \|h\|$. But the bound is attained with the input $h/\|h\|$. □

Remark 9.3. *It is important to observe that this correspondence is linear if $K = \mathbb{R}$, but conjugate-linear if $K = \mathbb{C}$. This is a consequence of the conjugate-linearity of inner products with respect to their second variable.*

9.1. **Hilbert Space Adjoint.** The Riesz Representation Theorem (**??**) allows us to define the adjoint of a bounded linear operator on a Hilbert space in a natural and elegant way. This concept, known as the Hilbert space adjoint, has important applications in spectral theory and quantum mechanics.

Proposition 9.4. *Let H and K be Hilbert spaces, and let $T : H \to K$ be a bounded linear operator. Then there exists a unique bounded linear operator $T^* : K \to H$, called the adjoint of T, such that*

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

for all $x \in H$ and $y \in K$.

Proof. Fix $y \in K$ and consider the map:

$$\begin{aligned} L_y : H &\to \mathbb{K} \\ x &\mapsto \langle Tx, y\rangle \end{aligned}$$

The map L_y defines a bounded linear functional on H. By the Riesz Representation Theorem (**??**) there exists a unique $h \in H$ such that

$$\langle Tx, y\rangle = L_y(x) = \langle h, x\rangle.$$

Define $T^*y = h$. It is clear that T^* is a linear map. To see that it is bounded, observe that

$$\begin{aligned} \|T^*y\|_H &= \|h\|_H \\ &= \sup_{\|x\|=1} |\langle h, x\rangle_H| \\ &= \sup_{\|x\|=1} |\langle Tx, y\rangle_K| \\ &\leq \sup_{\|x\|=1} \|Tx\|_K \cdot \|y\|_K \\ &\leq \sup_{\|x\|=1} \|T\| \cdot \|x\| \cdot \|y\| = \|T\| \cdot \|y\|. \end{aligned}$$

We conclude that T^* is bounded, and that $\|T^*\| \leq \|T\|$. We now show that T^* is unique. Suppose that $S \in \mathscr{B}(K, H)$ also satisfies

$$\langle Tx, y\rangle_K = \langle x, Sy\rangle_H.$$

for all $x \in H$ and $y \in K$ Then, for each fixed $y \in K$, we have

$$\langle x, Sy - T^*y\rangle_H = 0$$

for all $x \in H$ and $y \in K$. This implies $Sy - T^*y = 0$ for all $y \in K$. Hence, $S = T^*$, proving uniqueness. $\square$

We now discuss various properties of the adjoint operator.

Proposition 9.5. *Let H, K, L be Hilbert space.*

(1) *If $T \in \mathscr{B}(H, K)$, then $(T^*)^* = T$.*

(2) *If $A, B \in \mathscr{B}(H, K)$ and $\alpha, \beta \in \mathbb{K}$, then*

$$(\alpha T + \beta S)^* = \overline{\alpha}T^* + \overline{\beta}S^*.$$

(3) *If $A \in \mathscr{B}(H, K)$ and $B \in \mathscr{B}(K, L)$, then*

$$(S \circ T)^* = T^* \circ S^*.$$

(4) *If $A \in \mathscr{B}(H)$ is invertible in $\mathscr{B}(H)$, then A^* is invertible in $\mathscr{B}(H)$ and*

$$(A^{-1})^* = (A^*)^{-1}.$$

Proof. The proof is given below:

(1) Let $x \in H, y \in K$. We have that

$$\langle Tx, y\rangle_K = \langle x, T^*y\rangle_H = \overline{\langle T^*y, x\rangle_H} = \overline{\langle y, T^{**}x\rangle_K} = \langle T^{**}x, y\rangle_K$$

This shows that $T = T^{**}$.

(2) Let $x \in H, y \in K$. We have that

$$\langle(\alpha T + \beta S)x, y\rangle_H = \langle x, (\alpha T + \beta S)^*y\rangle_K = \langle x, (\overline{\alpha}T^* + \overline{\beta}S^*)y\rangle_K$$

This shows that $(\alpha T + \beta S)^* = \overline{\alpha}T^* + \overline{\beta}S^*$.

(3) Let $x \in H, z \in L$.

$$\langle x, (S \circ T)^* z \rangle_H = \langle (S \circ T)x, z \rangle_L = \langle S(Tx), z \rangle_L = \langle Tx, S^* z \rangle_K = \langle x, T^* S^* z \rangle_H$$

This shows that $(S \circ T)^* = T^* \circ S^*$.

(4) Skipped.

This completes the proof. $\square$