Metric and Banach Spaces

Alexandre Daoud
King's College London
alex.daoud@mac.com

April 28, 2016

Chapter 1

Sequence Spaces

1.1 Finite Dimensional Case

Definition 1.1. Let K be a field and V a vector space over K. A function $||\cdot|| : V \to \mathbb{R}$ is called a **norm** if, given $x, y \in V$ and $\lambda \in K$, we have

1. Positivity - $||x|| \geq 0$ with equality if and only if $x = 0$
2. Homogeneity - $||\lambda x|| = |\lambda| x$
3. Triangle inequality - $||x + y|| \leq ||x|| + y||$

Definition 1.2. Let $p \in [1, \infty]$ and n a natural number. We define the **p-norm** on $\mathbb{C}^n$ to be

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \qquad \text{for } p < \infty$$
$$||x||_\infty = \max_{i=1,\ldots,n} |x_i|, \qquad \text{for } p = \infty$$

The positivity and homogeneity norm axioms are easy to check. The triangle inequality is more complicated. We first require the following two results:

Lemma 1.3 (Young's Inequality)**.** *Let $p, q \in \mathbb{R}$ be such that $p, q > 0$ and $1/p + 1/q = 1$. Then for all $a, b \in \mathbb{R}$ such that $a, b \geq 0$ we have*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

with equality if and only if $a = b^{-1}$.

Proof. We have

$$ab = \exp[\ln(ab)] = \exp[\ln a + \ln b] = \exp\left[\frac{1}{p} p \ln a + \frac{1}{q} q \ln b\right] = \exp\left[\frac{1}{p} \ln a^p + \frac{1}{q} \ln b^q\right]$$

Now the exponential function is strictly increasing and strictly convex so, combined with the hypothesis $1/p + 1/q = 1$, we have

$$ab \leq \frac{1}{p} \exp[\ln a^p] + \frac{1}{q} \exp[\ln b^q] = \frac{a^p}{p} + \frac{b^q}{q}$$

$\square$

Lemma 1.4 (Hölder's Inequality)**.** *Let $p, q \in \mathbb{R}$ be such that $p > 1$ and $q > 1$ and $1/p+1/q = 1$. Then, given any $x, y \in \mathbb{C}^n$, we have*

$$||xy||_1 \leq ||x||_p \, ||y||_q$$

Proof. Let $u = x/||x||_p$ and $v = y/||y||_q$. Then, clearly, $||u||_p = ||v||_q = 1$. By Young's Inequality we have

$$|u_i v_i| \leq \frac{|u_i|^p}{p} + \frac{|v_i|^q}{q}$$

Passing to the sum on both sides we see that

$$||uv||_1 \leq \frac{1}{p}||u||_p + \frac{1}{q}||v||_q = 1$$

Hence

$$||xy||_1 = ||x||_p \, ||y||_q \, ||uv||_1 \leq ||x||_p \, ||y||_q$$

□

Remark. In the case that $p = q = 2$, Hölder's inequality reduces to the Cauchy-Schwarz inequality.

Proposition 1.5 (Minkowski's inequality)**.** *Given $x, y \in \mathbb{C}^n$, we have that*

$$||x + y||_p \leq ||x||_p + ||y||_p$$

Proof. We shall first prove this for the case that $p = 1$. We have that

$$\begin{aligned} ||x + y||_1 &= \sum_{i=1}^{n} |x_i + y_i| \\ &\leq \sum_{i=1}^{n} |x_i| + |y_i| \\ &= \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| \\ &= ||x||_1 + ||y||_1 \end{aligned}$$

We now prove the proposition for the case where $p = \infty$. We have that

$$\begin{aligned} ||x + y||_\infty &= \max_{i=1,\ldots,n} |x_i + y_i| \\ &\leq \max_{i=1,\ldots,n} (|x_i| + |y_i|) \\ &\leq \max_{i=1,\ldots,n} |x_i| + \max_{i=1,\ldots,n} |y_i| \\ &= ||x||_\infty + ||y||_\infty \end{aligned}$$

Now fix $p \in (1, \infty)$ and choose $q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. We first consider the triangle inequality for the modulus of complex numbers. Let $x, y \in \mathbb{C}^n$. Then

$$|x_i + y_i| \leq |x_i| + |y_i|$$

multiplying through by $|x_i + y_i|^{p-1}$ we get

$$|x_i + y_i|^p \leq |x_i||x_i + y_i|^{p-1} + |y_i||x_i + y_i|^{p-1}$$

We now sum over i on both sides of the inequality:

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i||x_i + y_i|^{p-1}$$

Now applying Hölder's inequality to both terms of the right hand side of the above inequality, it follows that

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}}$$

Now, $q = (1 - \frac{1}{p})^{-1}$ whence $(p-1)q = p$ so this becomes

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |x_i + y_i|^{p}\right)^{\frac{1}{q}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |x_i + y_i|^{p}\right)^{\frac{1}{q}}$$

Dividing through by the common factor in the terms of the right hand side yields

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1-\frac{1}{q}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$$

We now note that $1 - 1/q = 1/p$ and Minkowski's Inequality is proven. $\square$

Definition 1.6. Let K be a field and V be a vector space over K. Suppose that $||\cdot||_1$ and $||\cdot||_2$ are two norms on V. We say that $||\cdot||_1$ and $||\cdot||_2$ are **equivalent** if there exist $c, C \in \mathbb{R}$ such that for all $x \in X$ we have

$$c||x||_1 \leq ||x||_2 \leq C||x||_1$$

Proposition 1.7. *Let $p, q \in [1, \infty]$. Then $||\cdot||_p$ and $||\cdot||_q$ are equivalent as norms on $\mathbb{C}^n$.*

Proof. Fix $x \in \mathbb{C}^n$. It suffices to prove that $||x||_\infty \leq ||x||_p \leq n^{\frac{1}{p}}||x||_\infty$. We have that

$$\max_{i=1,\ldots,n} |x_i|^p \leq \sum_{i=1}^{n} |x_i|^p \leq n \max_{i=1\ldots n} |x_i|^p$$

Now taking the p^{th} root across these inequalities yields the desired result. $\square$

1.2 ℓ^p spaces

Proposition 1.8. *Let $p \in [1, \infty]$. Consider the set*

$$\ell^p = \left\{ (x_1, x_2, \ldots) \in \mathbb{C}^\infty \;\middle|\; \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} < \infty \right\}$$

when $p < \infty$ *and*

$$\ell^\infty = \left\{ (x_1, x_2, \ldots) \in \mathbb{C}^\infty \;\middle|\; \sup_{i \in \mathbb{N}} |x_i| < \infty \right\}$$

when $p = \infty$. *Then* ℓ^p *is a normed vector space over* $\mathbb{C}$ *with norm given by*

$$||x||_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}}$$

when $p < \infty$ *and*

$$||x||_\infty = \sup_{i \in \mathbb{N}} |x_i|$$

when $p = \infty$.

Proof. We shall prove the case where $p \in [1, \infty)$. It is clear that ℓ^p contains an additive identity, contains additive inverses, and contains scalar multiples. All additive and scalar multiplicative properties of vector spaces are also satisfied trivially. We must check that if $\lambda \in \mathbb{C}$ and $x, y \in \ell^p$ then $x + \alpha y \in \ell^p$. In the finite dimensional case, Minkowski's Inequality implies that

$$\left(\sum_{i=1}^{n} |x_i + \alpha y_i|^p \right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i| \right)^{1/p} + |\alpha| \left(\sum_{i=1}^{n} |y_i| \right)^{1/p}$$

Now, passing to the limit $n \to \infty$, the right hand side is bounded by some constant and thus $x + \alpha y \in \ell^p$. Hence ℓ^p is a (infinite dimensional) vector space over $\mathbb{C}$.

The positivity and homogeneity properties of $|| \cdot ||_p$ follow immediately. The triangle inequality also follows by a similar argument to the above where we pass to the limit $n \to \infty$ in the finite dimensional Minkowski Inequality. □

Recall that a vector space is **complete** if every Cauchy sequence in the vector space converges to a limit in the vector space.

Definition 1.9. Let K be a field and V a vector space over K. We say that V is a **Banach space** if V is a complete normed linear space.

Theorem 1.10. *ℓ^p is a Banach space for any* $1 \leq p \leq \infty$.

Proof. By Proposition 1.8, ℓ^p is a normed linear space. We must show that ℓ^p is complete. Let $\{ x^{(k)} \}_{k \in \mathbb{N}}$ be a Cauchy sequence in ℓ^p and fix $i \in \mathbb{N}$. It is clear that $|x_i^{(k)}| \leq ||x^{(k)}||_p$ for all $k \in \mathbb{N}$ whence $\{ x_i^{(k)} \}_{k \in \mathbb{N}}$ is a Cauchy sequence of complex numbers. Since $\mathbb{C}$ is complete, we have that $x_i^{(k)} \to x_i$ as $k \to \infty$.

We shall prove the theorem in the case $1 \leq p < \infty$. Since $x^{(k)}$ is Cauchy, given any $\varepsilon > 0$, there exists $N_\varepsilon \in \mathbb{N}$ such that

$$\left(\sum_{i=1}^{n} |x_i^{(k)} - x_i^{(m)}|^p \right)^{1/p} \leq \varepsilon$$

for all $k, m \geq N_\varepsilon$ and for all $n \in \mathbb{N}$. Letting $m \to \infty$, we have

$$\left(\sum_{i=1}^{n} |x_i^{(k)} - x_i|^p \right)^{1/p} \leq \varepsilon$$

for all $k \geq N_\varepsilon$ and for all $n \in \mathbb{N}$. Since n is an arbitrary natural number, it follows that

$$\left(\sum_{i=1}^{\infty} |x_i^{(k)} - x_i|^p\right)^{1/p} \leq \varepsilon \tag{1.1}$$

for all $k \geq N_\varepsilon$. By the definition of ℓ^p, it follows that $x^{(k)} - x$ is an element of ℓ^p for any k. Since $x^{(k)}$ is also an element of ℓ^p, by linearity, we must have that $x \in \ell^p$. We can now rewrite (1.1) as follows:

$$||x^{(k)} - x||_p \leq \varepsilon$$

for all $k \in N_\varepsilon$. Since ε is arbitrary, we then have that $||x^{(k)} - x||_p \to 0$ as $k \to \infty$ whence ℓ^p is complete. $\square$

Proposition 1.11. *Let $p, q \in [1, \infty]$. Then ℓ^p and ℓ^q are not equivalent as normed spaces.*

Proof. Suppose, without loss of generality, that $p > q$. Consider the linear subspace f of ℓ^p and ℓ^q consisting of sequences with only finitely many non-zero terms. Let $\{ g^{(n)} \}_{n \in \mathbb{N}}$ be the sequence in f whose i^{th} term is the sequence in f whose first i terms are 1 and the rest zero. In other words,

$$\begin{aligned} g^{(1)} &= (1, 0, 0, 0, \dots) \\ g^{(2)} &= (1, 1, 0, 0, \dots) \\ g^{(3)} &= (1, 1, 1, 0, \dots) \\ &\vdots \end{aligned}$$

Define the sequence

$$f^{(n)} = n^{-1/p} g^{(n)}$$

Then

$$||f^{(n)}||_p = n^{-1/p} ||g^{(n)}||_p = n^{-1/p} \left(\sum_{i=1}^{\infty} |g_i^{(n)}|^p\right)^{1/p} = n^{-1/p} n^{1/p} = 1$$

So $f^{(n)}$ does not converge to 0 in ℓ^p. Now,

$$||f^{(n)}||_q = n^{-1/p} ||g^{(n)}||_q = n^{-1/p} n^{1/q} = n^{1/q - 1/p}$$

But $p > q$ whence $1/q > 1/p$ so $f^{(n)} \to 0$ as $n \to \infty$ and is thus Cauchy. Hence ℓ^p and ℓ^q cannot be equivalent as normed spaces. $\square$

Definition 1.12. Let K be a field and X, Y vector spaces over K. Suppose that X is equipped with the norm $|| \cdot ||_X$ and Y is equipped with the norm $|| \cdot ||_Y$. If X is a subspace of Y and there exists a constant $c \in \mathbb{R}$ such that $||x||_Y \leq c||x||_X$ for all $x \in X$ then we say that X **embeds** in Y and we denote it as $X \hookrightarrow Y$.

Proposition 1.13. *Let $p, r \in [1, \infty]$ such that $p < r$. Then $\ell^p \hookrightarrow \ell^r$.*

Proof. We first prove that $\ell^p \hookrightarrow \ell^\infty$. To see that ℓ^p is a subspace of ℓ^∞, let $x \in \ell^p$. Then there exists some constant $\varepsilon \in \mathbb{R}$ such that

$$\left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} = \varepsilon$$

Taking the p^{th} power of both sides yields the following comparison:

$$||x||_\infty = \sup_{i \in \mathbb{N}} |x_i| \leq \sup_{i \in \mathbb{N}} |x_i|^p \leq \sum_{i=1}^{\infty} |x_i|^p = \varepsilon^p < \infty$$

and so $x \in \ell^\infty$. Furthermore, for any $i \in \mathbb{N}$ we have

$$|x_i|^p \leq \sum_{i=1}^{\infty} |x_i|^p$$

and so $\ell^p \hookrightarrow \ell^\infty$. We now prove the general case. We may write $r = p + s$ so that for all $i \in \mathbb{N}$

$$||x||_r^r = |x_i|^r = |x_i|^p |x_i|^s \leq |x_i|^p \sup_{i \in \mathbb{N}} |x_i|^s \leq |x_i|^p ||x||_\infty^s \leq ||x||_p^p ||x||_p^s = ||x||_p^r$$

Taking the $1/r^{th}$ power across this inequality yields $||x||_r \leq ||x||_p$ and so $\ell^p \hookrightarrow \ell^r$. □

1.3 Separability of ℓ^p spaces

Definition 1.14. Let A and B be two sets. We say that A and B have the same **cardinality** if there exists a bijection $f : A \to B$. We say that A is **finite** if there exists some $n \in \mathbb{N}$ such that A has the same cardinality as $\{ 1, \ldots, n \}$. We say that A is **countable** if it is either finite or has the same cardinality as $\mathbb{N}$.

Theorem 1.15. *Let A be a set and $X \subseteq A$. If A is countable then X is countable.*

Proof. It suffices to show that any subset $A \subseteq \mathbb{N}$ is countable. Let

$$\begin{aligned} a_1 &= \min A, \quad A_1 = A \backslash \{ a_1 \} \\ a_2 &= \min A_1, \quad A_2 = A_1 \backslash \{ a_2 \} \\ &\vdots \end{aligned}$$

If at some stage, A_n is empty then A is finite. If not then the map $n \mapsto a_n$ is a bijection between $\mathbb{N}$ and A. □

Theorem 1.16. *Let A be a set. If there exsists a surjection $f : \mathbb{N} \to A$ then A is countable.*

Proof. Consider $a \in A$ and let $g(a) = \min f^{-1}(\{ a \})$ and set $B = \{ g(a) \mid a \in A \}$. Clearly $B \subseteq \mathbb{N}$ and g is a bijection between B and A. The theorem then follows from Theorem 1.15. □

Theorem 1.17. *A countable union of countable sets is countable.*

Proof. Let A_n be countable sets where $n \in N$ and N is itself countable. First assume that $N, A_1, \dots$ are all infinite. Then we may assume that $N = \mathbb{N}$. Suppose $A_n = \{ a_1^{(n)}, a_2^{(n)}, \dots \}$. We can enumerate the elements of $A = \bigcup_{n=1}^{\infty} A_n$ as follows (in a diagonal fashion):

$$a_1^{(1)}, a_2^{(1)}, a_1^{(2)}, a_3^{(1)}, a_2^{(2)}, a_1^{(3)}, a_4^{(1)}, \dots$$

This establishes a surjection $f : \mathbb{N} \to A$. Now by Theorem 1.16, A is countable. In the case that any of the sets $N, A_1, \dots$ are finite, we can add extra elements to these sets to make them infinite. ☐

Corollary 1.18. *If A and B are countable then the Cartesian product $A \times B$ is countable.*

Proof. Note that we can write

$$A \times B = \bigcup_{b \in B} A \times \{ b \}$$

This is a countable union of countable sets so by 1.17, it is countable. ☐

Example 1.19. $\mathbb{Z}$ and $\mathbb{Q}$ are countable.

Theorem 1.20. *Let X be non-empty. Then the power set 2^X of X cannot have the same cardinality as X.*

Proof. Assume there exists a bijection $f : X \to 2^X$. Consider the set

$$S = \{ x \in X \mid x \notin f(x) \} \subseteq X$$

Now let $s \in X$ be such that $f(s) = S$. First suppose that $s \in S$. then $s \notin f(s)$. But this contradicts the definition of s. Now suppose that $s \notin S$. We have that $s \in f(s)$. Again, this contradicts the definition of s. Hence there can exist no such bijection. ☐

Corollary 1.21. *$\mathbb{R}$ and $\mathbb{N}$ have different cardinalities.*

Proof. Consider the interval $[0, 1]$. We claim that $2^{\mathbb{N}}$ has the same cardinality as $[0, 1]$. Let $x \in [0, 1]$. Then x can be written uniquely in the form of a binary expansion $0.a_1a_2a_3\dots$ where $a_j \in \{ 0, 1 \}$. Such a binary expansion uniquely defines a subset of $\mathbb{N}$. This establishes a bijection between $[0, 1]$ and $2^{\mathbb{N}}$. ☐

Definition 1.22. Let X be a Banach space. We say that X is separable if it contains a countable dense subset.

Example 1.23. Since $\mathbb{Q}$ is countable and dense in $\mathbb{R}$, the space $\mathbb{R}$ is separable.

Theorem 1.24. *Let $p < \infty$. Then ℓ^p is separable. ℓ^∞ is not separable.*

Proof. First suppose $p < \infty$. Let $A_n \subseteq \ell^p$ be the set of all sequences $x \in \ell^p$ such that all coordinates of x are rational and $x_k = 0$ for all $k > n$. Furthermore, let $A = \cup_{n=1}^{\infty} A_n$. Clearly each A_n is countable and hence A is itself countable. We claim that A is dense in ℓ^p.

To this end, let $y \in \ell^p \backslash A$. We need to exhibit an $x \in A$ such that x is arbitrarily close to y. Let $n \in \mathbb{N}$ and denote $P_n : \ell^p \to \ell^p$ to be the mapping

$$P_n(y) = (y_1, y_2, \dots, y_n, 0, 0, \dots)$$

Given any $\varepsilon > 0$, we can clearly find an $n \in \mathbb{N}$ such that $||y - P_n(y)||_p \leq \frac{\varepsilon}{2}$.
Now since $\mathbb{Q}$ is dense in $\mathbb{R}$, it follows that the set of all complex rationals is dense in $\mathbb{C}$. It is thus clear that we can find an $x \in A_n$ such that $||P_n(y) - x||_p \leq \frac{\varepsilon}{2}$. Now using the triangle inequality, we have that there exists an $n \in \mathbb{N}$ such that $||y - x||_p < \varepsilon$. Since ε is arbitrary, we can always find an x arbitrarily close to y and we are done.

We now show that ℓ^∞ is not separable. Let $X \subseteq \mathbb{N}$ and define $e^X \in \ell^\infty$ as follows: $e^X_n = 1$ if $n \in X$ and $e^X_n = 0$ if $n \notin X$. Then the number of such elements of ℓ^p is uncountable and, for any $X \neq Y$, we have $||e^X - e^Y||_\infty = 1$.

Now suppose that $D \subseteq \ell^\infty$ is a countable dense subset. Consider the balls $B_{\frac{1}{2}}(e^X)$ over all $X \subseteq N$. Since $2^{\mathbb{N}}$ is uncountable, this collection of balls is uncountable. Now, these balls are disjoint and each contain an element of D since D is dense. Hence there exists a bijection between a subset of D (which is countable) and the set of all such balls (which is uncountable). This is a contradiction and hence there can exist no such set D. $\square$

Chapter 2

Lebesgue integration and L^p spaces

2.1 Riemann Integral

Definition 2.1. Fix an interval $[a, b] \in \mathbb{R}$. We define a **step function** to be a finite linear combination of characteristic functions of bounded intervals:

$$f(x) = \sum_n c_n \chi_{\delta_n}(x)$$

Definition 2.2. Let f be a step function. We define the **integral** of f as follows:

$$\int_a^b f(x)\, dx = \sum_n c_n \mu(\delta_n)$$

where $\mu(\delta)$ is understood to be the length of the interval δ.

There are multiple issues with the Reimann integral:

1. If $f_n \to f$ pointwise then we cannot conclude that $\int f_n \to \int f$.
2. The class of Riemann integrable functions $R[a, b]$ contains many functions which, intuitively, should be integrable
3. $C[a, b]$ with the norm

   $$||f||_1 = \int_a^b |f(x)|dx$$

   is not complete. It's completion is much larger than $R[a, b]$ so we would like to be able to better describe it.

2.2 Lebesgue measure

Definition 2.3. Let X be a set and Σ a collection of subsets of X such that

1. $\varnothing \in X$
2. Σ is closed under complements
3. Σ is closed under countable union

Then we say that Σ is a **σ-algebra** over X.

Definition 2.4. We define the **Borel** σ-algebra of $\mathbb{R}$, denoted $\mathcal{B}(\mathbb{R})$ to be the σ-algebra of $\mathbb{R}$ containing every open intervals. Any set in $\mathcal{B}(\mathbb{R})$ is called a **Borel set**.

Example 2.5. Let S be the set of all irrational numbers whose continued fractions are of the form

$$a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}$$

such that there exists a sequence $0 < i_1 < i_2 < \dots$ where each a_{i_k} divides $a_{i_{k+1}}$. Then $S \notin \mathcal{B}(\mathbb{R})$.

Definition 2.6. We define the **Lebesgue measure** to be the function $\mu : \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ given by

$$\mu(B) = \sum_i (b_i - a_i)$$

when B is a countable union of some open intervals (a_i, b_i). Otherwise it is defined as

$$\mu(B) = \inf_{B \subseteq O} \mu(O)$$

where the infimum is taken over open sets O containing B.

Theorem 2.7. *Let A_n be a countable collection of mutually disjoint Borel sets. Then*

$$\mu(\cup_n A_n) = \sum_n \mu(A_n)$$

In other words, the Lebesgue measure is countably additive.

Definition 2.8. A **measure space** is a triple $(\Omega, \mathcal{A}, \mu)$ where Ω is a set, $\mathcal{A}$ is a σ-algebra over Ω and μ is a countably additive function on A.

2.3 Borel functions

Definition 2.9. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We say that f is a **Borel** function if $f^{-1}[(a, b)]$ is a Borel set for any interval $(a, b) \subseteq \mathbb{R}$.

Remark. Clearly, if f is continuous then f is Borel.

Proposition 2.10. *A function f is Borel if and only if for all $B \in \mathcal{B}(\mathbb{R})$, the set $f^{-1}[B]$ is Borel.*

Proof. First suppose that f is Borel. Let X be the set

$$X = \{\, B \in \mathcal{B}(\mathbb{R}) \mid f^{-1}(B) \in \mathcal{B}(\mathbb{R}) \,\}$$

Clearly, $X \subseteq \mathcal{B}(\mathbb{R})$. We claim that X is a σ-algebra that contains the open intervals. By the minimality of $\mathcal{B}(\mathbb{R})$, it would then follow that $X = \mathcal{B}(\mathbb{R})$. Since f is Borel, X necessarily contains the open intervals. We now check the axioms of a σ-algebra one by one.

We have $\varnothing \in \mathcal{B}(\mathbb{R})$ and $f^{-1}(\varnothing) = \varnothing$ and so $\varnothing \in X$.

Let $B \in X$. Then $B \in \mathcal{B}(\mathbb{R})$ and $f^{-1}(B) \in \mathcal{B}(\mathbb{R})$. Now, $\mathcal{B}(\mathbb{R})$ is a σ-algebra and so $B^c \in \mathcal{B}(\mathbb{R})$. We have $f^{-1}(B^c) = f^{-1}(B)^c \in \mathcal{B}(\mathbb{R})$ whence $B^c \in X$.

Let $\{ A_n \}$ be a countable collection of elements of X. Then each $A_n \in \mathcal{B}$ and $f^{-1}(A_n) \in \mathcal{B}$. Since $\mathcal{B}(\mathbb{R})$ is a σ-algebra, we have $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$. Hence

$$f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} f^{-1}(A_i) \in \mathcal{B}(\mathbb{R})$$

Thus X is a σ-algebra. Furthermore, since f is Borel, X necessarily contains all open intervals whence $X = \mathcal{B}(\mathbb{R})$.

Conversely, suppose that for all $B \in \mathcal{B}(\mathbb{R})$ we have $f^{-1}(B) \in \mathcal{B}(\mathbb{R})$. Then clearly for all open intervals $(a, b) \subseteq \mathbb{R}$ we have $f^{-1}((a, b)) \in \mathcal{B}(\mathbb{R})$ and so f is Borel. □

Corollary 2.11. *Let f and g be Borel functions. Then the following functions are Borel:*

1. $f \circ g$
2. $|f|$
3. $f_{\pm} = \max\{\pm f, 0\}$

Proof. Let $(a, b) \subseteq \mathbb{R}$. Since f is Borel, we have $g^{-1}((a, b)) = B$ for some Borel set B. Since f is Borel, we have $f^{-1}(B) = B'$ for some Borel set B'. Hence $(f \circ g)^{-1}((a, b)) = B'$ whence $f \circ g$ is Borel.

Now, the absolute value function is Borel since it is continuous. Hence, by the first part of the corollary, $|f|$ is Borel.

To show that $f_{\pm}$ is Borel, consider the function g sending x to $\max\{\pm x, 0\}$. $\pm x$ is clearly a continuous function as is the zero function. Since the maximum of any two continuous functions is continuous, it follows that g is continuous. Therefore, g is Borel. Appealing to the first part of the corollary, we see that $f_{\pm} = f \circ g$ is Borel. □

Theorem 2.12. *Let f_n be a sequence of Borel functions. If $f_n(x) \to f(x)$ for all x as $n \to \infty$ then f is Borel.*

Remark. The previous theorem is not necessarily true for $C[a, b]$ and $R[a, b]$.

Definition 2.13. Let B be a set and $A \subseteq B$ a subset. We define the **characteristic** function of A to be:

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } \textit{otherwise} \end{cases}$$

Definition 2.14. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We say that f is a **simple** function if it is a linear combination of characteristic functions of Borel sets. The vector space of all simple functions over $\mathbb{R}$ is denoted $\mathrm{Simp}(\mathbb{R})$.

Proposition 2.15. *Let f be a simple function. Then f is Borel.*

Proof. Given $c \in \mathbb{R}$ and B be a Borel set, it suffices to show that

$$f(x) = c\chi_B(x)$$

is Borel. Let $(a, b) \subseteq \mathbb{R}$ be an interval. First suppose that $0, c \in B$. Then $f^{-1}((a, b)) = \mathbb{R} \in \mathcal{B}(\mathbb{R})$. Next suppose that $0 \in B$ and $c \notin (a, b)$. Then $f^{-1}((a, b)) = \mathbb{R}\backslash B \in \mathcal{B}(\mathbb{R})$. Now suppose that $0 \notin (a, b)$ but $c \in (a, b)$. Then $f^{-1}((a, b)) = B \in \mathcal{B}(\mathbb{R})$. Finally, suppose that neither $0 \in (a, b)$ nor $c \in (a, b)$. Then $f^{-1}((a, b)) = \varnothing \in \mathcal{B}(\mathbb{R})$. Hence in all cases, we see that $f^{-1}((a, b))$ is a Borel set whence f is Borel. □

Proposition 2.16. *Let f be a positive Borel function. Then there exists a sequence of simple functions f_n such that $f_{n+1} \geq f_n$ and $f_n \leq f$ for all n and $f_n(x) \to f(x)$ for all x as $n \to \infty$. Furthermore, if f is bounded then f_n can be chosen so as to converge uniformly to f.*

Corollary 2.17. *A linear combination of Borel functions is a Borel function.*

Proof. Let $af + bg$ be a linear combination of Borel functions for some $a, b \in \mathbb{R}$. Since b and g are Borel, Proposition 2.16 implies that there exists monotone sequences of simple functions $\{ f_n \}$ and $\{ g_n \}$ such that $f_n \leq f$ and $g_n \leq g$ for all n and $f_n(x) \to f(x)$, $g_n(x) \to g(x)$ for all x as $n \to \infty$. Then $af_n + bg_n$ is a sequence of simple functions and are thus also Borel. It clearly converges to $af + bg$. By Theorem 2.12, $af + bg$ is Borel. □

2.4 Lebesgue Integral

Definition 2.18. We define the **Lebesgue integral** to be the functional given by

$$\int_{\mathbb{R}} f(x)\, d\mu = \sum_{i=1}^{n} c_i \mu(B_i)$$

when $f = \sum_{i=1}^{n} c_i \chi_{B_i}$ is a simple function. If $f \geq 0$ is a Borel function then it is defined by

$$\int_{\mathbb{R}} f(x)\, d\mu = \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \leq f}} g(x)\, d\mu$$

Proposition 2.19. *Let f be a Borel function. Then $\int |f|\, d\mu$ is finite if and only if both $\int f_+\, d\mu$ and $\int f_-\, d\mu$ are finite.*

Proof. First suppose that $\int |f|\, d\mu$ is finite. Then it is clear that $\int f_\pm\, d\mu$ are finite.

Conversely, suppose that $\int f_\pm\, d\mu$ is finite. Then

$$\begin{aligned}
\int |f|\, d\mu &= \int f_+ - f_-\, d\mu \\
&= \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \leq f_+ - f_-}} \int g\, d\mu \leq \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \leq f_+}} \int g\, d\mu - \sup_{\substack{h \in \mathrm{Simp}(\mathbb{R}) \\ h \leq f_+}} \int h\, d\mu \leq \int f_+\, d\mu - \int f_-\, d\mu < \infty
\end{aligned}$$

□

Definition 2.20. Denote by $\mathcal{L}^1(a, b)$ the class of all Borel functions f on (a, b) such that the integrals $\int f_\pm\, dx$ are both finite. For $f \in \mathcal{L}^1$ we define the **Lebesgue integral** of f by

$$\int f\, d\mu = \int f_+\, d\mu - \int f_-\, d\mu$$

Theorem 2.21 (Montonone Convergence Theorem)**.** *Let $f_n \geq 0$ be a sequence of Borel functions on $\mathbb{R}$ such that $f_{n+1} \geq f_n$ for all n and $f_n(x) \to f(x)$ for all x as $n \to \infty$. If $\int f_n(x)\, dx \leq C$ for all n then $f \in \mathcal{L}^1(\mathbb{R})$ and*

$$\int f_n(x)\, d\mu \to \int f\, d\mu$$

Theorem 2.22 (Dominated Convergence Theorem)**.** *Let f_n be a sequence of Borel functions on $\mathbb{R}$ such that $f_n(x) \to f(x)$ for all x as $n \to \infty$ and $f_n(x) \le F(x)$ for all n and x and some $F \in \mathcal{L}^1(\mathbb{R})$ then*

$$\int |f_n(x) - f(x)|\ \mu \to 0$$

as $n \to \infty$ and thus

$$\int |f_n(x)|\ d\mu \to \int |f|\ d\mu$$

as $n \to \infty$.

Theorem 2.23. *Let $f, g \in \mathcal{L}^1(a, b)$. Then*

1. *$\int f \le \int g$ whenever $f \le g$*
2. *$|\int f| \le \int |f|$*
3. *$|\int f| \le (b-a)||f||_\infty$*
4. *$\int (af + bg) = a \int f + b \int g$*

where $||f||_\infty = \sup_{x \in (a,b)} |f(x)|$.

Proof.

Part 1: Suppose that $f \le g$. Then

$$\int f\ d\mu = \sup_{\substack{h_1 \in \mathrm{Simp}(\mathbb{R}) \\ h_1 \le f}} \int h_1\ d\mu \le \sup_{\substack{h_2 \in \mathrm{Simp}(\mathbb{R}) \\ h_2 \le g}} \int h_2\ d\mu = \int g\ d\mu$$

Part 2: We have that

$$\left|\int f\ d\mu\right| = \left|\sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \int g\ d\mu\right| \le \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \left|\int g\ d\mu\right| \le \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \int |g|\ d\mu = \int |f|\ d\mu$$

for simple functions $g \le f$.

Part 3: We have that

$$\left|\int f\ d\mu\right| \le \int |f|\ d\mu = \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \int |g|\ d\mu$$

Now if $g = \sum_n c_n^{(g)} \chi_{B_n^{(g)}}$ for some $c_n^{(g)} \in \mathbb{R}$ and Borel sets $B_n^{(g)}$, we have

$$\begin{aligned}
\left|\int f\ d\mu\right| \le \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \int |g|\ dx = \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \sum_n |c_n^{(g)} \mu(B_n^{(g)})| &\le \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \sum_n |c_n^{(g)} (b-a)| \\
&= (b-a) \sup_{\substack{g \in \mathrm{Simp}(\mathbb{R}) \\ g \le f}} \sum_n |c_n^{(g)}| \\
&\le (b-a) \sup_{x \in (a,b)} |f| \\
&= (b-a)||f||_\infty
\end{aligned}$$

Part 4: We first prove that the integral is a linear functional for simple functions. Let $\phi = \sum_{i=1}^{n} c_n \chi_{C_n}$ and $\psi = \sum_{j=1}^{m} d_n \chi_{D_n}$ (for collections of pairwise disjoint sets C_i and D_i) be simple functions and $a, b \in \mathbb{R}$. We have that

$$\begin{aligned}
\int (a\phi + b\psi)\ d\mu &= \sum_{i=1}^{n}\sum_{j=1}^{m}(ac_i + bd_j)\chi_{C_i \cap D_j} \\
&= a\sum_{i=1}^{n}\sum_{j=1}^{m} c_i \chi_{C_i\cap D_j} + b\sum_{i=1}^{n}\sum_{j=1}^{m} d_j \chi_{C_i \cap D_j} \\
&= a\sum_{i=1}^{n} c_i \sum_{j=1}^{m} \chi_{C_i \cap D_j} + b\sum_{j=1}^{m} d_j \sum_{i=1}^{n} \chi_{C_i\cap D_j} \\
&= a\sum_{i=1}^{n} c_i \chi_{C_i} + b\sum_{j=1}^{m} d_j \chi_{D_i} \\
&= a\int \phi\ d\mu + b\int \psi\ d\mu
\end{aligned}$$

Now let f and g be Borel functions. By Lemma 2.16 we can always find a sequence of increasing simple functions f_n and g_n such that $f_n \to f$ and $g_n \to g$. Then by the Monotone Convergence Theorem, we have that

$$\begin{aligned}
a\int f\ d\mu + b\int g\ d\mu &= \lim_{n\to\infty}\int af_n\ d\mu + \lim_{n\to\infty}\int bg_n\ d\mu \\
&= \lim_{n\to\infty}\left(\int af_n\ d\mu + \int bg_n\ d\mu\right)
\end{aligned}$$

Now f_n and g_n are all simple functions which are linear by the previous claim whence

$$\int af\ d\mu + \int bg\ d\mu = \lim_{n\to\infty}\left(\int af_n + bg_n\ d\mu\right)$$

Now again by the monotone convergence theorem, the limit on the right is equal to $\int af + bg\ d\mu$.

□

Theorem 2.24 (Fatou's Lemma)**.** *Let f_n be a sequence of non-negative Borel functions. Then*

$$\int \liminf f_n(x)\ d\mu = \liminf \int f_n(x)\ d\mu$$

Proof. Let $h_m(x) = \inf_{n\geq m}\{f_n(x)\}$. Clearly, $h_m(x) \leq f_m(x)$ and $h_m(x) \leq h_{m+1}(x)$ for all x and m. We first observe that

$$\int h_m(x)\ dx \leq \int f_n(x)$$

for all $n \geq m$. This is equivalent to

$$\int h_m(x)\ dx \leq \inf_{n\geq m}\int f_n(x)$$

Now, by the Monotone Convergence Theorem

$$\begin{aligned}\int \liminf f_n(x)\,dx &= \int \lim_{m\to\infty} h_m(x)\,dx \\ &= \lim_{m\to\infty}\int h_m(x)\,dx \\ &\leq \liminf \int f_n(x)\,dx\end{aligned}$$

□

Definition 2.25. Let $f : \mathbb{C} \to \mathbb{C}$ be a complex valued function. We define the **Lebesgue integral** of f to be

$$\int f\,d\mu = \int \mathrm{Re}(f)\,d\mu + i\int \mathrm{Im}(f)\,d\mu$$

2.5 The L^p spaces

Definition 2.26. Let $p > 1$ and $(\Omega, \mathcal{A}, \mu)$ a measure space. We denote by $\mathcal{L}^p(\Omega)$ the collection of measurable functions f which satisfy

$$||f||_p = \left(\int_\Omega |f(x)|^p\,d\mu(x)\right)^{\frac{1}{p}} < \infty$$

Example 2.27. $\mathcal{L}^p(\mathbb{R})$ where $\mathcal{A} = \mathcal{B}(\mathbb{R})$ and μ is the Lebesgue measure.

Proposition 2.28. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $f \in \mathcal{L}^p(\Omega)$ integrable. Then $||f||_p = 0$ if and only if $f = 0$ almost everywhere.*

Proof. First suppose that $||f||_p = 0$. Denote

$$A = \{\, x \in \Omega \mid f(x) \neq 0 \,\}$$

Let $A_n = \{\, x \in \Omega \mid |f(x)| > 1/n \,\}$. Then, clearly, $A = \bigcup_n A_n$. Now suppose, for a contradiction, that A has strictly positive measure. Then at least one of the A_n has strictly positive measure, say A_k. Then $|f(x)|$ dominates $1/k\chi_{A_k}$ on A_k. We then have that

$$0 = \int_{A_k} |f|^p\,d\mu > \int_{A_k} \frac{1}{k^p}\chi_{A_k}\,d\mu = \frac{1}{k^p}\mu(A_n) > 0$$

which is a contradiction.

Now suppose that $f = 0$ almost everywhere. Then for any simple function g satisfying $0 \leq g \leq |f|^p$, we must have that g is 0 almost everywhere. Let such a g have representative $\sum_n c_n^{(g)} B_n^{(g)}$ for some $c_n^{(g)} \in \mathbb{R}$ and Borel sets $B_n^{(g)}$. Then

$$\int |f|^p\,d\mu = \sup_{\substack{g\in\mathrm{Simp}(\mathbb{R})\\ g\leq|f|^p}} c_n^{(g)} B_n^{(g)} = 0$$

since, given any n, either $c_n^{(g)} = 0$ or $\mu(B_n) = 0$. □

Proposition 2.29. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Then $\mathcal{L}^p(\Omega)$ is a vector space over $\mathbb{R}$ (or $\mathbb{C}$).*

Proof. Let $f, g \in \mathcal{L}^p(\Omega)$ and $\alpha \in \mathbb{R}$. Then $f + \alpha g$ is a measurable function. Hence

$$\begin{aligned}\int |f+\alpha g|^p \, d\mu &\leq \int (|f| + |\alpha||g|)^p \, d\mu \\ &\leq \int (2\max\{|f|, |\alpha||g|\})^p \, d\mu \\ &= 2^p \int \max\{|f|^p, |\alpha||g|^p\} \, d\mu \\ &\leq 2^p \int |f|^p + |\alpha||g|^p \, d\mu \\ &= 2^p \left(\int |f|^p \, d\mu + \int |\alpha||g|^p \, d\mu \right) < \infty\end{aligned}$$

and so $f + \alpha g \in \mathcal{L}^p(\Omega)$. The rest of the vector space axioms are clear from the basic properties of functions and thus $\mathcal{L}^p(\Omega)$ is a vector space. □

Proposition 2.30 (Hölder's Inequality for L^p spaces)**.** *Let $p, q \in \mathbb{R}$ be such that $p, q > 1$ and $1/p + 1/q = 1$. If $(\Omega, \mathcal{A}, \mu)$ is a measure space and $f \in \mathcal{L}^p(\Omega), g \in \mathcal{L}^q(\Omega)$ then $fg \in \mathcal{L}^1(\Omega)$ and*

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

Proof. First assume that either $||f||_p = 0$ or $||g||_q = 0$. It then follows that fg is equal to 0 almost everywhere. Hölder's Inequality then follows immediately in this case.

Hence, we may assume that neither $||f||_p$ and $||g||_q$ are zero. Let $u = f/||f||_p$ and $v = g/||g||_q$. We claim first that $||uv||_1 \leq ||u||_p ||v||_q = 1$. By Young's Inequality we have, for all $x \in \mathbb{R}$,

$$|u(x)v(x)| \leq \frac{|u(x)|^p}{p} + \frac{|v(x)|^q}{q}$$

Passing to the Lebesgue integral we have

$$\int |uv| \, d\mu \leq \frac{1}{p} \int |u|^p \, d\mu + \frac{1}{q} \int |v|^q \, d\mu = 1$$

We then have that

$$\left|\left| \frac{fg}{||f||_p ||g||_q} \right|\right|_1 \leq 1$$

whence

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

□

Proposition 2.31 (Minkowski's Inequality for L^p spaces)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p > 1$. Suppose that $f, g \in \mathcal{L}^p(\Omega)$. Then*

$$||f+g||_p \leq ||f||_p + ||g||_p$$

Proof. If $p = 1$ then Minkowski's Inequality follows directly by the properties of the Lebesgue integral hence suppose $p > 1$. Let $q \in \mathbb{R}$ be such that $1/p + 1/q = 1$. We have that

$$|f + g| \leq |f| + |g|$$

Multiplying this inequality through by $|f + g|^{p-1}$ we have

$$|f + g|^p \leq |f||f + g|^{p-1} + |g||f + g|^{p-1}$$

Now, $(p - 1)q = p$ whence $(f + g)^{p-1} \in \mathcal{L}^p(\Omega)$ so we may pass to the Lebesgue integral to get

$$||f + g||_p^p \leq ||(|f||f + g|^{p-1})||_1 + ||(|g||f + g|^{p-1})||_1$$

Applying Hölder's Inequality yields

$$||f + g||_p^p \leq ||f||_p||(f + g)^{p-1}||_q + ||g||_p||(f + g)^{p-1}||_q$$

Now note that

$$||(f + g)^{p-1}||_q = \left(\int |f + g|^{q(p-1)}\, d\mu\right)^{1/q} = \left(\int |f + g|^p\, d\mu\right)^{1/q} = ||f + g||_p^{p/q}$$

and so

$$||f + g||_p^p \leq ||f||_p||f + g||_p^{p/q} + ||g||_p||f + g||_p^{p/q}$$

Since $p - p/q = 1/p$ and we may assume that $||f + g||_p > 0$ it follows that

$$||f + g||_p \leq ||f||_p + ||g||_q$$

□

Corollary 2.32. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space. Then $\mathcal{L}^p(\Omega)$ is a semi-normed*[1] *vector space over $\mathbb{R}$ (or $\mathbb{C}$).*

Definition 2.33. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p \geq 1$. Define an equivalence relation on $\mathcal{L}^p(\Omega)$ by $f \sim g$ if and only if $f = g$ almost everywhere. We define the $L^p(\Omega)$ space to be the collection of equivalence classes of $\sim$. By an abuse of notation, for any equivalence class $[f] \in L^p(\Omega)$ we shall write $f \in L^p(\Omega)$ to be one of its representatives (if a continuous represntative exists then we shall usually chose that one).

Proposition 2.34. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p \geq 1$. Then $L^p(\Omega)$ is a normed vector space over $\mathbb{R}$ (or $\mathbb{C}$).*

Proof. $L^p(\Omega)$ is clearly a semi-normed vector space over $\mathbb{R}$ - this follows directly from the results for $\mathcal{L}^p(\Omega)$. To see that $|| \cdot ||_p$ is a norm on $L^p(\Omega)$, we just need to prove that $||f|| = 0$ if and only if $f = 0$. Now, $f = 0$ means that f is the equivalence class containing 0. This equivalence class contains all functions that are 0 almost everywhere. We may thus choose such a representative, such as the 0 function, to see that $||f||_p = 0$. □

[1] recall that a semi-normed space is a set equipped with a so-called semi-norm which satisfies all properties of a norm except $||x|| = 0 \iff x = 0$.

Proposition 2.35. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p \geq 1$. Then $L^p(\Omega)$ is a Banach space.*

Proof. We need to show that $L^p(\Omega)$ is complete. Let $\{ f_n \}_{n\in\mathbb{N}}$ be a Cauchy sequnce in $L^p(\Omega)$. Let $\{ f_{n_k} \}_{k\in\mathbb{N}}$ be a subsequence of $\{ f_n \}$ such that

$$||f_{n_{k+1}} - f_{n_k}||_p \leq 2^{-k}$$

for all $k \geq 1$. Now consider the functions

$$f(x) = f_{n_1}(x) + \sum_{i=1}^{\infty}(f_{n_{k+1}}(x) - f_{n_k}(x))$$

$$g(x) = |f_{n_1}(x)| + \sum_{i=1}^{\infty}|f_{n_{k+1}}(x) - f_{n_k}(x)|$$

and denote their partial sums by $S_K(f), S_K(g)$. We have

$$||S_K(g)||_p \leq ||f_{n_1}(x)||_p + \sum_{i=1}^{K}||f_{n_{k+1}} - f_{n_k}||_p \leq ||f_{n_1}(x)||_p + 2^{-k}$$

Clearly, there exists a constant $C \in \mathbb{R}$ such that $||S_K(g)||_p < C$ for all $K \geq 1$. Furthermore, $S_{K+1}(g) \geq S_K(g)$ for all $K \geq 1$ and $S_K(g) \to g$ pointwise. Appealing to the Montonone Convergence Theorem, we have that $g \in L^p(\Omega)$. It then follows that $f \in L^p(\Omega)$.

We now claim that f is the limit of $\{ f_n \}$. Observe that $S_{K-1}(g) = f_{n_K}$ so $f_{n_K} \to f$ pointwise as $K \to \infty$. We also show that $f_{n_K} \to f$ in $L^p(\Omega)$. We have that

$$\begin{aligned}|f(x) - S_K(f)(x)|^p &\leq (2\max\{|f(x)|, |S_K(f)(x)|\})^p \\ &\leq 2^p|f(x)|^p + 2^p|S_K(f)(x)|^p \\ &\leq 2^{p+1}|g(x)|^p\end{aligned}$$

Appealing to the Dominated Convergence Theorem, we see that $||f - f_{n_K}||_p \to 0$ as $k \to \infty$.

Finally, since $\{ f_n \}$ is Cauchy, for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n, m > N$ we have $||f_n - f_m||_p \leq \varepsilon/2$. Now choose n_K such that $n_K > N$. Then $||f_{n_K} - f||_p \leq \varepsilon/2$. By the triangle inequality, we then have that

$$||f_n - f||_p \leq ||f_n - f_{n_K}||_p + ||f_{n_K} - f||_p \leq \varepsilon$$

and we are done. □

Proposition 2.36. *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and $p \geq 1$. Then the collection of all simple functions in $L^p(\Omega)$ is dense in $L^p(\Omega)$.*

Proof. Let $f \in L^p(\Omega)$. For all $\varepsilon > 0$, it suffices to exhibit a simple function g such that $||f - g||_p < \varepsilon$.

Recall that every measurable function f can be approximated pointwise by a sequence of simple functions $\{ f_n \}$ such that $f_{n+1} \geq f_n$. We clearly have that $f_n(x) \leq f(x)$ for all x and n. By the Dominated Convergence Theorem, we thus have that $||f_n - f||_p \to 0$ as $n \to \infty$. Hence for all $\varepsilon > 0$, there exists an n such that $||f_n - f||_p < \varepsilon$. Therefore, the simple functions are dense in $L^p(\Omega)$. □

Proposition 2.37. *The collection of all continuous functions in $L^p(a, b)$ is dense in $L^p(a, b)$ for all intervals $(a, b) \subseteq \mathbb{R}$ (including infinite intervals).*

Proof. By 2.36, it suffices to show that the characteristic function of any Borel set B can be approximated arbitrarily well by a continuous function. In other words, for all $\varepsilon > 0$, we need to show that there exists a continuous function $g \in L^p(\mathbb{R})$ such that $||\chi_B - g||_p < \varepsilon$.

To this end, fix $\varepsilon > 0$, choose an open set $G \supseteq B$ such that $\mu(G\backslash B) < \varepsilon$. Let $f(x)$ denote the function

$$f(x) = \frac{d(x, \overline{G})}{d(x, \overline{G}) + d(x, B)}$$

where $d(x, C) := \inf_{y \in C} |x - y|$. Since d is continuous, so is $f(x)$. Now if $x \in B$ then $d(x, B) = 0$ and, consequently, $f(x) = 1$. Now if $x \notin G$ then $f(x) = 0$ by definition of $d(x, C)$. Hence $f(x) - \chi_B(x) = 0$ for all $x \in B$ and $x \in \overline{G}$ and $|f(x) - \chi_B(x)| < 1$ for all $x \in G\backslash B$. Thus

$$\int |f(x) - \chi_B(x)|^p \, d\mu < \mu(G\backslash B) < \varepsilon$$

$\square$

Corollary 2.38. *$L^p(a, b)$ is separable for any (possibly unbounded) interval $(a, b) \subseteq \mathbb{R}$.*

Proof. We prove the corollary for $L^p(\mathbb{R})$. By 2.5, the continuous functions are dense in $L^p(\mathbb{R})$. By the Stone-Weierstrass Theorem, we can approximate continuous functions arbitrarily well with polynomial functions. Now, we can approximate any polynomial arbitrarily well with polynomials with rational coefficients. The latter collection is clearly countable and dense in $L^p(\mathbb{R})$ so $L^p(\mathbb{R})$ is separable. $\square$

Definition 2.39. Let f be a Borel function. We define the **essential supremum** of f to be

$$\operatorname{ess\,sup} f = \sup\{\, t \mid \mu(\{\, x \mid f(x) > t \,\}) > 0 \,\}$$

Remark. Let f be a Borel function and $L = \operatorname{ess\,sup} f$. Clearly, if $t < L$ then $\mu(\{\, x \mid f(x) > t \,\}) > 0$. If $t > L$ then $\mu(\{\, x \mid f(x) > t \,\}) = 0$.

Proposition 2.40. *Let f be a Borel function. Then $\mu(\{\, x \mid f(x) > \operatorname{ess\,sup} f \,\}) = 0$.*

Proof. Let $L = \operatorname{ess\,sup} f$. If $L = \infty$ then the proposition is clear so assume that L is finite. Note that $\{\, x \mid f(x) > \operatorname{ess\,sup} \,\} = f^{-1}((L, \infty])$. We have that

$$f^{-1}((L, \infty]) = \bigcup_{k=1}^{\infty} f^{-1}((L + 1/k, \infty])$$

By the definition of ess sup, each set in this union has measure zero whence $f^{-1}((L, \infty])$ has measure zero. $\square$

Definition 2.41. Denote by $\mathcal{L}^\infty(a, b)$ the collection of all Borel functions f on (a, b) such that there exists $M \in \mathbb{R}$ with $|f(x)| \leq M$ almost everywhere. Define the function $||\cdot||_\infty : L^\infty(a, b) \to \mathbb{R}$ by

$$||f||_\infty = \operatorname{ess\,sup} |f(x)|$$

Define an equivalence relation on $\mathcal{L}^\infty(a, b)$ where $f \sim g$ if and only if $||f - g||_\infty = 0$ almost everywhere. We denote by $L^\infty(a, b)$ the collection of all equivalence classes of $\sim$.

Proposition 2.42. *$L^\infty(a,b)$ is a normed vector space over $\mathbb{R}$ (or $\mathbb{C}$) for all (possibly unbounded) intervals $(a,b) \subseteq \mathbb{R}$.*

Proof. We first show that $L^\infty(\mathbb{R})$ is a vector space over $\mathbb{R}$. To this end, let $\alpha \in \mathbb{R}$ and $f, g \in L^\infty(\mathbb{R})$. We need to show that $f + \alpha g \in L^\infty(a,b)$. That is to say, we need to show that there exists $M \in \mathbb{R}$ such that

$$|f + \alpha g| \leq M$$

almost everywhere. By hypothesis, there exists $M_1, M_2 \in \mathbb{R}$ such that $|f| \leq M_1$ and $|g| < M_2$ almost everywhere. Hence

$$|f + \alpha g| \leq |f| + |\alpha||g| \leq M_1 + |\alpha| M_2$$

almost everywhere whence $f + \alpha g \in L^\infty(\mathbb{R})$. The rest of the vector space axioms follow directly from the basic properties of functions and so $L^\infty(\mathbb{R})$ is a vector space.

For the norm axioms, we first prove homogeneity. Let $c \in \mathbb{R}$ and $f \in L^\infty(\mathbb{R})$. We have that

$$||cf||_\infty = \operatorname{ess\,sup} |cf| = \sup \{ t \mid \mu(\{ x \mid |cf(x)| > t \} > 0) \} = |c| ||f||_\infty$$

We next prove the triangle inequality. Let $f, g \in L^\infty(\mathbb{R})$. Then there exist sets of measure zero $X, Y \subseteq \mathbb{R}$ such that $|f(x)| \leq ||f(x)||_\infty$ for all $x \in \mathbb{R} \backslash X$ and $|g(x)| \leq ||g(x)||_\infty$ for all $x \in \mathbb{R} \backslash Y$. Then $X \cup Y$ is again a set of measure zero. Then for all $x \in \mathbb{R} \backslash (X \cup Y)$ we have

$$||f + g||_\infty = \operatorname{ess\,sup} |f(x) + g(x)| \leq |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq ||f||_\infty + ||g||_\infty$$

Now if $f \in L^\infty(\mathbb{R})$ then it is clear that $||f||_\infty \geq 0$. Now suppose that $||f||_\infty = 0$. By Proposition 2.40, we have that $\{ x \mid |f(x)| > 0 \} = \{ x \mid f(x) \neq 0 \}$ has measure zero. Hence f is zero almost everywhere whence f is the equivalence class of the 0 function and we are done. ☐

Proposition 2.43. *$L^\infty(a,b)$ is a Banach space for any (possibly unbounded) interval $(a,b) \subseteq \mathbb{R}$.*

Proof. We shall prove the proposition for $L^\infty(\mathbb{R})$. It suffices to show that $L^\infty(a,b)$ is complete. To this end, let $\{ f_n \}$ be a Cauchy sequence in $L^\infty(\mathbb{R})$. That is to say, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n, m > N$ we have $||f_m - f_n||_\infty < \varepsilon$. This is equivalent to there existing sets of measure zero $Y_{m,n}$ such that for all $x \in \mathbb{R} \backslash Y_{m,n}$ we have $|f_m - f_n| < \varepsilon$. Denote $Y = \bigcup_{m,n} Y_{m,n}$. Then Y has measure zero and for all $x \in \mathbb{R} \backslash Y$ we have $|f_m - f_n| < \varepsilon$ for all $m, n > N$.

Hence for all $x \in \mathbb{R} \backslash Y$, $\{ f_n(x) \}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, this sequence converges pointwise to some limit $f(x)$. This limit f is defined outside of the measure zero set Y. For $x \in Y$, we may take $f(x) = 0$. In other words, $f = \lim_n \chi_{\mathbb{R} \backslash Y} f_n$. Note that this function is measurable. In the Cauchy sequence condition, we may let $n \to \infty$ so that for all $m \geq N$ we have

$$|f_m(x) - f| < \varepsilon$$

But then

$$||f_m(x) - f||_\infty = \operatorname{ess\,sup} |f_m(x) - f| \leq |f_m(x) - f| < \varepsilon$$

so that $f_m(x) \to f$ as $m \to \infty$ with respect to the L^∞ norm.

Finally, we show that $f \in L^\infty(\mathbb{R})$. We have that

$$||f||_\infty = ||f + f_m - f_m||_\infty \leq ||f - f_m||_\infty + ||f_m||_\infty \leq ||f_m||_\infty + \varepsilon < \infty$$

and so $f \in L^\infty(\mathbb{R})$. $\square$

Proposition 2.44. *$L^\infty(a, b)$ is not separable for any (possibly unbounded) interval $(a, b) \subseteq \mathbb{R}$.*

Proof. Consider the collection of functions $\chi_{(c,d)}$ over the intervals $(c, d) \subseteq (a, b)$. This collection is clearly uncountable. Furthermore, if $c \neq c'$ and $d \neq d'$ then $||\chi_{(c,d))} - \chi_{(c',d')}||_\infty = 1$.

Now suppose there exists a dense countable subset $D \subseteq L^\infty(a, b)$. Consider the balls $B_{\frac{1}{2}}(\chi_{(c,d)})$ around each interval. These balls are clearly disjoint and uncountable. Since D is dense, there must be an element of D in each such ball. But this is a contradiction as it implies the existence of a surjection from a countable set onto an uncountable set. $\square$

Proposition 2.45. *Let $p, q \geq 1$ be such that $p > q$. Then $L^p(a, b) \hookrightarrow L^q(a, b)$ for any bounded interval (a, b).*

Proof. Assume that $f \in L^p(a, b)$. Since $p > q$, we have $q/p < 1$. Observe that $q/p + (1 - q/p) = 1$. Let $u = p/q$ and $v = 1/(1 - q/p)$ so that $1/u + 1/v = 1$. By Hölder's Inequality, we have that

$$||f||_q = ||\,|f|^q\,||_1 = ||\,|f|^q \cdot 1\,||_1 \leq ||\,|f|^q\,||_u ||1||_v = \left(\int |f|^{q(p/q)}\, d\mu\right)^{q/p} (b - a)^{1-q/p}$$

Taking the q^{th} root across this inequality we have that

$$||f||_q \leq (b - a)^{1/q - 1/p} ||f||_p$$

as desired. $\square$

Proposition 2.46. *Let $r \geq 1$ and $f \in L^r(a, b)$ for some bounded interval (a, b). Then*

$$\lim_{p\to\infty} ||f||_p = ||f||_\infty$$

Proof. Let $t \in [0, ||f||_\infty)$. By definition,

$$A = \{\, x \mid |f(x)| \geq t \,\}$$

is a set of positive measure. We then have the following inequality:

$$\begin{aligned} ||f||_p &\geq \left(\int_A |f|^p\right)^{1/p} \\ &\geq (t^p \mu(A))^{1/p} \\ &= t\mu(A)^{1/p} \end{aligned}$$

If $\mu(A)$ is finite then $\mu(A)^{1/p} \to 1$ as $p \to \infty$. If $\mu(A)$ is infinite then $\mu(A)^{1/p}$ is infinite for all p. In either case, we have

$$\liminf_{p\to\infty} ||f||_p \geq t$$

Now t is arbitrary, we have

$$\liminf_{p\to\infty} ||f||_p \geq ||f||_\infty$$

For the reverse inequality, note that $|f(x)| \leq ||f||_\infty$ for almost all x. Then for all $p \geq r$ we have

$$||f||_p = \left(\int |f|^p \, d\mu\right)^{1/p} = \left(\int |f|^r |f|^{p-r} \, d\mu\right)^{1/p} \leq \left(\int |f|^r ||f||_\infty^{p-r} \, d\mu\right)^{1/p} = ||f||_\infty^{1-r/p} ||f||_r^{r/p}$$

Now, $||f||_r^{r/p} < \infty$ for all p and so

$$\limsup_{p\to\infty} ||f||_p \leq \limsup_{p\to\infty} ||f||_\infty^{1-r/p} ||f||_r^{r/p} = ||f||_\infty$$

$\square$

Chapter 3

Hilbert spaces

3.1 The geometry of Hilbert spaces

Definition 3.1. Let V be a vector space over $\mathbb{C}$. We say that V is an **inner product space** if there is a complex valued function $(\cdot,\cdot)$ on $V \times V$ such that, given any $x, y, z \in V$ and $\alpha \in \mathbb{C}$,:

1. $(x, x) \geq 0$ and $(x, x) = 0$ if and only if $x = 0$
2. $(x, y + z) = (x, y) + (x, z)$
3. $(x, \alpha y) = \alpha(x, y)$
4. $(x, y) = \overline{(y, x)}$

$(\cdot,\cdot)$ is referred to as an **inner product**.

Example 3.2. $\mathbb{C}^n$ is an inner product space. Given $x = \langle x_1, \ldots, x_n\rangle$ and $y = \langle y_1, \ldots, y_n\rangle$, we define their inner product as

$$(x, y) = \sum_{j=1}^{n} \overline{x}_j y_j$$

Example 3.3. Let $C[a, b]$ denote the complex-valued continuous functions on the interval $[a, b]$. Given $f(x), g(x) \in C[a, b]$, we define their inner product as

$$(f, g) = \int_a^b \overline{f(x)} g(x)\, dx$$

Definition 3.4. Let V be an inner product space and $x, y \in V$. We say that x and y are **orthogonal** if $(x, y) = 0$. A collection of elements $\{ x_i \} \subseteq V$ is said to be an **orthonormal** set if $(x_i, x_i) = 1$ for all i and $(x_i, x_j) = 0$ for all $i \neq j$.

Remark. Let V be an inner product space with inner product given by $(\cdot,\cdot)$. If $x \in V$ we write $||x|| = \sqrt{(x, x)}$.

Theorem 3.5 (Pythagorean theorem)**.** *Let V be an inner product space and $x_1, \ldots, x_N \in V$ be an orthonormal set. Then, given any $x \in V$, we have that*

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

Proof. We first claim that

$$\sum_{n=1}^{N}(x_n, x)x_n, \qquad x - \sum_{n=1}^{N}(x_n, x)x_n$$

are orthogonal. We have that

$$\left(\sum_{n=1}^{N}(x_n, x)x_n, x - \sum_{n=1}^{N}(x_n, x)x_n\right) = \underbrace{\left(\sum_{n=1}^{N}(x_n, x)x_n, x\right)}_{A} - \underbrace{\left(\sum_{n=1}^{N}(x_n, x)x_n, \sum_{n=1}^{N}(x_n, x)x_n\right)}_{B}$$

Firstly, we have

$$A = \sum_{n=1}^{N}\overline{(x_n, x)}(x_n, x)$$

Secondly, we have

$$\begin{aligned} B &= \sum_{i=1}^{N}\sum_{j=1}^{N}\overline{(x_i, x)}(x_j, x)(x_i, x_j) \\ &= \sum_{i=1}^{N}\overline{(x_i, x)}(x_i, x) \end{aligned}$$

where we have used the fact that $(x_i, x_j) = 0$ for all $i \neq j$. Hence $A - B = 0$ as desired. It then follows that

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

□

Corollary 3.6 (Bessel's inequality)**.** *Let V be an inner product space and $x_1, \dots, x_N \in V$ an orthonormal set. Then, given any $x \in V$, we have*

$$||x||^2 \geq \sum_{n=1}^{N}|(x, x_n)|^2$$

Corollary 3.7 (Cauchy-Schwarz inequality)**.** *Let V be an inner product space and $x, y \in V$. Then*

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

Proof. This is clearly true for $y = 0$ hence assume $y \neq 0$. Consider the vector $\frac{y}{||y||}$. This vector by itself clearly forms an orthonormal set. Hence we may apply Bessel's inequality to any $x \in V$ so that

$$\begin{aligned} ||x||^2 &\geq \left|\left(x, \frac{y}{||y||}\right)\right| \\ &= \frac{|(x, y)|^2}{||y||^2} \end{aligned}$$

whence the result follows by muptiplying through by $||y||^2$ and taking the square root across the inequality. □

Theorem 3.8. *Let V be an inner product space. Then V is a normed linear space with norm given by $||x|| = \sqrt{(x,x)}$.*

Proof. The first two axioms of a norm are satisfied directly from the first four properties of the inner product. It remains to show that the triangle inequality holds. We have that

$$\begin{aligned}||x+y||^2 &= (x,x) + (x,y) + (y,x) + (y,y)\\ &= (x,x) + 2\mathrm{Re}(x,y) + (y,y)\\ &\leq (x,x) + 2|(x,y)| + (y,y)\\ &\leq (x,x) + 2\sqrt{(x,x)}\sqrt{(y,y)} + (y,y)\\ &= (||x|| + ||y||)^2\end{aligned}$$

where in the second to last inequality we have used the Cauchy-Schwarz inequality. The result then follows by taking the square root across the inequality. □

Remark. This norm naturally induces the metric

$$d(x,y) = \sqrt{(x-y, x-y)}$$

from which we can introduce convergence, completeness and density to the inner product space.

Proposition 3.9 (Parallelogram Law)**.** *Let V be an inner product space with inner product given by $(\cdot,\cdot)$. Then for all $x, y \in V$ we have*

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2)$$

Proof. We have that

$$\begin{aligned}||x+y||^2 + ||x-y||^2 &= (x+y, x+y) + (x-y, x-y)\\ &= (x+y, x) + (x+y, y) + (x-y, x) - (x-y, y)\\ &= (x,x) + (x,y) + (y,x) + (y,y) + (x,x) - (x,y) - (y,x) + (y,y)\\ &= 2(x,x) + 2(y,y)\\ &= 2(||x||^2 + ||y||^2)\end{aligned}$$

□

Proposition 3.10 (Polarisation Identity)**.** *Let V be an inner product space with inner product given by $(\cdot,\cdot)$. Let $x, y \in V$. If V is an $\mathbb{R}$-vector space then*

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

If V is a $\mathbb{C}$-vector space then

$$(x,y) = \frac{1}{4}(||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2)$$

Proof. We shall verify only the real case, the complex case follows by a similar argumentation. Expanding the right hand side of the claimed identity gives

$$\begin{aligned}\frac{1}{4}[(x+y, x+y) - (x-y, x-y)] &= \frac{1}{4}[(x+y, x) + (x+y, y) - (x-y, x) + (x-y, y)]\\ &= \frac{1}{4}[(x,x) + (x,y) + (y,x) + (y,y) - (x,x) + (x,y)\\ &+ (y,x) - (y,y)]\\ &= \frac{1}{4}[2(x,y) + 2(y,x)]\\ &= (x,y)\end{aligned}$$

□

Theorem 3.11. *Let V be a normed vector space over $\mathbb{R}$ (or $\mathbb{C}$) with norm given by $||\cdot||$. Then $||\cdot||$ is induced by an inner product if and only if the Parallelogram Law holds in V.*

Proof. First suppose that $||\cdot||$ is induced by an inner product. Then V is clearly an inner product space and Proposition 3.9 implies that the Parallelogram Law holds.

Now suppose that the Parallelogram Law holds. We shall only prove the case where V is a $\mathbb{R}$-vector space. Define (x, y) by the Polarisation Identity

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

We claim that $(\cdot,\cdot)$ is an inner product that induces $||\cdot||$. First note that

$$(x, x) = \frac{1}{4}(||x+x||^2) = \frac{1}{4}(4||x||^2) = ||x||^2$$

and so $(\cdot,\cdot)$ induces $||\cdot||$.

We now check each axiom of inner products. We clearly have that $(x, x) \geq 0$ with equality if and only if $x = 0$. It is also clearly symmetric. To prove linearity in the second argument, let $x, y, z \in V$. By the Parallelogram Identity, we have

$$\begin{aligned}(x, y+z) &= \frac{1}{4}(||x+y+z||^2 - ||x-(y+z)||^2)\\ &= \frac{1}{4}(2||x+y||^2 + 2||z||^2 - ||x+y-z||^2 - 2||y||^2 - 2||x-z||^2 + ||x+y-z||^2)\\ &= \frac{1}{4}(2||z||^2 - 2||y||^2 + 2||x+y||^2 - 2||x-z||^2)\end{aligned}$$

Now note that these expressions should be symmetric in y, z. In other words, we have

$$\begin{aligned}||x+y+z||^2 - ||x-(y+z)||^2 &= 2||z||^2 - 2||y||^2 + 2||x+y||^2 - 2||x-z||^2\\ &= 2||y||^2 - 2||z||^2 + 2||x+z||^2 - 2||x-y||^2\end{aligned}$$

Adding these two expressions together and dividing by two, we have

$$||x+y-z||^2 - ||x-(y+z)||^2 = ||x+y||^2 + ||x+z||^2 - ||x-z||^2 - ||x-y||^2$$

and so

$$\begin{aligned}(x, y+z) &= \frac{1}{4}(||x+y||^2 + ||x+z||^2 - ||x-z||^2 - ||x-y||^2)\\ &= \frac{1}{4}(||x+y||^2 - ||x-y||^2) + \frac{1}{4}(||x+z||^2 - ||x-z||^2)\\ &= (x, y) + (x, z)\end{aligned}$$

Finally, we must check that $(x, \alpha y) = \alpha(x, y)$ for all $\alpha \in \mathbb{R}$. First suppose that $\alpha \in \mathbb{N}$. We shall prove the claim by induction on α. If $\alpha = 1$ then there is nothing to prove so assume that the claim holds for arbitrary natural α. We have

$$(x, (\alpha+1)y) = (x, \alpha y + y) = (x, \alpha y) + (x, y) = \alpha(x, y) + (x, y) = (\alpha+1)(x, y)$$

and so the claim holds for all $\alpha \in \mathbb{N}$. Observe that

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

It then follows that the claim holds true for all integers. Now suppose that $\alpha \in \mathbb{Q}$. We may assume, without loss of generality, that $\alpha = 1/b$ where b is a stricly positive integer. Then

$$\begin{aligned}\left(x, \frac{1}{b}y\right) &= \frac{1}{4}\left(\left|\left|x + \frac{1}{b}y\right|\right|^2 - \left|\left|x - \frac{1}{b}y\right|\right|^2\right) \\ &= \frac{1}{b^2}\frac{1}{4}(||bx+y||^2 - ||bx-y||^2) \\ &= \frac{1}{b^2}(bx, y) = \frac{1}{b^2}(y, bx) = \frac{1}{b}(x, y)\end{aligned}$$

Now, given any $x \in V$, $(x, \cdot)$ is clearly a continuous function in the second argument. Indeed, this function is expressed in terms of basic arithmetic operations of $||\cdot||$ which is continuous. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, the claim must hold for all $\alpha \in \mathbb{R}$ and we are done. □

Corollary 3.12. *Let $p \in [1, \infty]$. Then ℓ^p is an inner product space if and only if $p = 2$.*

Proof. By Theorem 3.11, ℓ^p is an inner product space if and only if the Parallelogram Law holds in ℓ^p. Let $x = e_1$ and $y = e_2$ where the e_i are the standard basis vectors for ℓ^p. We have

$$||x+y||_p^2 + ||x-y||_p^2 = 2||x||_p^2 + 2||y||_p^2 \iff 2^{2/p} + 2^{2/p} = 4 \iff 2^{2/p} = 2 \iff p = 2$$

If p is infinite then the condition reduces to $2 = 4$ which is clearly absurd. Hence the Parallelogram Law does not hold if $p \neq 2$.

It remains to show that ℓ^2 is indeed an inner product space. Define

$$\begin{aligned}(\cdot, \cdot) : \ell^2 \times \ell^2 &\to \mathbb{R} \\ (x, y) &\mapsto \sum_{i=1}^{\infty} \overline{x_i} y_i\end{aligned}$$

Clearly, $(x, x) \geq 0$ for all $x \in \ell^2$ with equality holding if and only if $x = 0$. Next, conjugate symmetry indeed holds:

$$(x, y) = \sum_{i=1}^{\infty} \overline{x}_i y_i = \overline{\sum_{i=1}^{\infty} x_i \overline{y_i}} = \overline{(y, x)}$$

Finally, let $x, y, z \in \ell^2$ and $\alpha \in \mathbb{C}$. Then

$$(x, y + \alpha z) = \sum_{i=1}^{\infty} \overline{x_i}(y_i + \alpha z_i) = \sum_{i=1}^{\infty} \overline{x_i} y_i + \alpha \sum_{i=1}^{\infty} \overline{x_i} z_i$$

and so $(\cdot, \cdot)$ is an inner product on ℓ^2. □

Corollary 3.13. *Let $p \in [1, \infty]$. Then $L^p(\mathbb{R})$ is an inner product space if and only if $p = 2$.*

Proof. First suppose that p is finite. Let $X, Y \subseteq \mathbb{R}$ be disjoint sets of finite measure. Define

$$f(x) = \left(\frac{1}{\mu(X)}\right)^{1/p} \chi_X(x), \quad g(x) = \left(\frac{1}{\mu(Y)}\right)^{1/p} \chi_Y(x)$$

We have that

$$\begin{aligned}||f+g||_p &= \left(\int \left| \left(\frac{1}{\mu(X)}\right)^{1/p} \chi_X + \left(\frac{1}{\mu(Y)}\right)^{1/p} \chi_Y \right|^p d\mu \right)^{1/p} \\ &= \left(\int \frac{1}{\mu(X)}\chi_X + \frac{1}{\mu(Y)}\chi_Y \, d\mu\right)^{1/p} \\ &= 2^{1/p}\end{aligned}$$

where we have used the fact that X and Y to see that $\chi_X \cdot \chi_Y = 0$. A similar argument also shows that $||f-g||_p = 2^{1/p}$. Note that $||f||_p = 1$ and $||g||_p = 1$. Now, in order for $L^p(\mathbb{R})$ to be an inner product space, the Parallelogram Law must hold. We have that

$$||f+g||_p^2 + ||f-g||_p^2 = 2||f||_p^2 + 2||g||_p^2 \iff 2^{1/p} + 2^{1/p} = 4 \iff 2^{1/p} = 2 \quad \iff p = 2$$

Now suppose that $p = \infty$. Let $f = \chi_X$ and $g = \chi_Y$ where X and Y are disjoint and have non-zero measure. Then

$$||f||_\infty = \sup\{\, t \mid \mu(\{\, x \mid |f(x)| > t\,\}) > 0\,\} = 1$$

Similarly, we see that $||g||_\infty = ||f+g||_\infty = ||f-g||_\infty = 1$. Now, checking the Parallelogram Law, we have

$$||f+g||_\infty^2 + ||f-g||_\infty = 2||f||_\infty^2 + 2||g||_\infty^2 \iff 2 = 4$$

which is absurd. Hence $L^p(\mathbb{R})$ is not an inner product space except possibly at $p = 2$.

To show that $L^2(\mathbb{R})$ is an inner product space, define

$$\begin{aligned}(\cdot,\cdot) : L^2(\mathbb{R}) &\to L^2(\mathbb{R}) \\ (f,g) &\mapsto \int \overline{f} g \, d\mu\end{aligned}$$

We must first check that this function is well-defined. We have

$$\int \overline{f} g \, d\mu \le \int |\overline{f} g| \, d\mu \le \int \frac{1}{2}|f(x)|^2 + \frac{1}{2}|g(x)|^2 \, d\mu < \infty$$

Clearly, $(f,f) \ge 0$ with equality if and only if $f = 0$. Next, conjugate symmetry indeed holds:

$$(f,g) = \int \overline{f} g \, d\mu = \overline{\int f \overline{g}} = \overline{(g,f)}$$

Finally, let $f, g, h \in L^2(\mathbb{R})$ and $\alpha \in \mathbb{C}$. Then

$$(f, g+\alpha h) = \int \overline{f}(g+\alpha h)\, d\mu = \int \overline{f} g \, d\mu + \alpha \int \overline{f} h \, d\mu = (f,g) + \alpha(f,h)$$

□

Definition 3.14. Let V be an inner product space. We say that V is a **Hilbert space** if it is complete.

Example 3.15. ℓ^2 is a Hilbert space.

Example 3.16. $L^2(\mathbb{R})$ is a Hilbert space.

Definition 3.17. Let V_1 and V_2 be inner product spaces with inner products given by $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$ respectively. If $U : V_1 \to V_2$ is a linear operator such that

$$(Ux, Uy)_2 = (x, y)_1$$

for all $x, y \in V_1$ then we say that U is **unitary**

Definition 3.18. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. We say that $\mathcal{H}_1$ and $\mathcal{H}_2$ are **isomorphic** if there exists a unitary linear operator from $\mathcal{H}_1$ to $\mathcal{H}_2$.

Definition 3.19. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces with inner products given by $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_2$ respectively. We define their **direct sum**, denoted $\mathcal{H}_1 \oplus \mathcal{H}_2$, to be the Hilbert space consisting of the collection of pairs $\langle x, y\rangle$ together with the inner product

$$(\langle x_1, y_1\rangle, \langle x_2, y_2\rangle) = (x_1, x_2)_1 + (y_1, y_2)_2$$

Definition 3.20. Let $\{\mathcal{H}_i\}_{i\in\mathbb{N}}$ be a countable collection of Hilbert spaces where $\mathcal{H}_i$ is equipped with the inner product $(\cdot,\cdot)_1$. We define their **countable direct sum**, denoted

$$\bigoplus_{i=1}^{\infty} \mathcal{H}_i$$

to be the Hilbert space consisting of infinite sequences $\{x_i\}_{i\in\mathbb{N}}$ with each $x_i \in \mathcal{H}_i$ satisfying

$$\sum_{i=1}^{\infty} ||x_i||_1^2 < \infty$$

equipped with the natural inner product.

3.2 The Riesz Lemma

Proposition 3.21. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{M} \subseteq \mathcal{H}$ a closed (with respect to the natural topology induced by the norm) subspace. Then $\mathcal{M}$ is a Hilbert space whose inner product is inherited from $\mathcal{H}$.*

Proof. Let $(\cdot,\cdot)$ be the inner product on $\mathcal{H}$. $\mathcal{M}$ is clearly an inner product space with $(\cdot,\cdot)$ restricted to the subspace. Now, any closed subspace of a complete normed space is necessarily complete whence $\mathcal{M}$ is a Hilbert space. $\square$

Definition 3.22. Let $\mathcal{H}$ be a Hilbert space with inner product given by $(\cdot,\cdot)$ and $\mathcal{M} \subseteq \mathcal{H}$ a subspace. We define the **orthogonal complement**, denoted $\mathcal{M}^\perp$ to be the following set

$$\mathcal{M}^\perp = \{ v \in \mathcal{H} \mid (v, m) = 0 \;\forall m \in \mathcal{M} \}$$

Proposition 3.23. *Let $\mathcal{H}$ be a Hilbert space with inner product given by $(\cdot,\cdot)$ and $\mathcal{M} \subseteq \mathcal{H}$ a subspace. Then $\mathcal{M}^\perp$ is a closed subspace of $\mathcal{H}$.*

Proof. It follows directly from the linearity property of the inner product that $\mathcal{M}^\perp$ is a subspace of $\mathcal{H}$.

We now show that $\mathcal{M}^\perp$ is closed. Indeed, consider a sequence $\{ x_n \} \subseteq \mathcal{M}^\perp$ such that $x_n \to x$ as $n \to \infty$ for some $x \in \mathcal{H}$. Fix some $m \in \mathcal{M}$. Then for all $n \in \mathbb{N}$ we have $(x_n, m) = 0$. By the continuity of the inner product, we have

$$\begin{aligned} \lim_{n\to\infty} (x_n, m) &= 0 \\ \left(\lim_{n\to\infty} x_n, m \right) &= 0 \\ (x, m) &= 0 \end{aligned}$$

whence $x \in \mathcal{M}^\perp$. The orthogonal complement thus contains all its limit points and is therefore closed. $\square$

Lemma 3.24. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{M} \subseteq \mathcal{H}$ a closed subspace. For any $x \in \mathcal{H}$, there exists a unique $m \in \mathcal{M}$ which is closest to x.*

Proof. Fix $x \in \mathcal{H}$ and set $d = \inf_{y\in\mathcal{M}} ||x - y||$. Choose a sequence $\{ y_n \} \subseteq \mathcal{M}$ such that $||x - y_n|| \to d$. By the Parallelogram Law we have

$$\begin{aligned} ||y_n - y_m||^2 &= ||(y_n - x) - (y_m - x)||^2 \\ &= 2||y_n - x||^2 + 2||y_m - x||^2 - || - 2x + y_n + y_m||^2 \\ &= 2||y_n - x||^2 + 2||y_m - x||^2 - 4 \left|\left| x - \frac{1}{2}(y_n + y_m) \right|\right|^2 \end{aligned}$$

Now note that $\frac{1}{2}(y_n - y_m) \in \mathcal{M}$ and thus

$$\begin{aligned} ||y_n - y_m||^2 &= 2||y_n - x||^2 + 2||y_m - x||^2 - 4 \left|\left| x - \frac{1}{2}(y_n + y_m) \right|\right|^2 \\ &\le 2||y_n - x||^2 + 2||y_m - x||^2 - 4d^2 \end{aligned}$$

Passing to the limit $n, m \to \infty$ on both sides yields $\lim_{n,m\to\infty} ||y_n - y_m||^2 \le 2d^2 + 2d^2 - 4d^2 = 0$. Therefore, $\{ y_n \}$ is a Cauchy sequence of elements in $\mathcal{M}$. But $\mathcal{M}$ is closed whence $\{ y_n \} \to z$ for some $z \in \mathcal{M}$. Hence $||x - z|| = d$.

It remains to show that such a z is unique. Let $z_1, z_2 \in \mathcal{M}$ be such that $||x - z_1|| = ||x - z_2|| = d$. Then, by the Parallelogram Law, we have

$$\begin{aligned} ||z_1 - z_2|| &= ||(z_1 - x) - (z_2 - x)||^2 \\ &= 2||z_1 - x||^2 + 2||z_2 - x||^2 - || - 2x + z_1 + z_2|| \\ &= 2||z_1 - x||^2 + 2||z_2 - x||^2 - 4 \left|\left| x - \frac{1}{2}(z_1 + z_2) \right|\right|^2 \\ &= 4d^2 - 4 \left|\left| x - \frac{1}{2}(z_1 + z_2) \right|\right|^2 \end{aligned}$$

Note that $\frac{1}{2}(z_1 + z_2) \in \mathcal{M}$ and thus

$$\begin{aligned} ||z_1 - z_2|| &= 4d^2 - 4 \left|\left| x - \frac{1}{2}(z_1 + z_2) \right|\right|^2 \\ &\le 4d^2 - 4d^2 \\ &= 0 \end{aligned}$$

Now by the properties of the norm, we can see that $z_1 = z_2$. $\square$

Theorem 3.25 (Projection Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space with inner product given by $(\cdot,\cdot)$ and $\mathcal{M} \subseteq \mathcal{H}$ a closed subspace. Then for all $x \in \mathcal{H}$ there exists unique $z \in \mathcal{M}$ and $w \in \mathcal{M}^\perp$ such that $x = z + w$.*

Proof. Fix $x \in \mathcal{H}$. By Lemma 3.24, there exist a $z \in \mathcal{M}$ closest to x. Define $w = x - z$. We claim that $w \in \mathcal{M}^\perp$. First, set $d = ||x - z||$. For all $y \in \mathcal{M}$ and $t \in \mathbb{R}$ we have

$$\begin{aligned} d^2 &\le ||x - (z + ty)||^2 \\ &= ||w - ty||^2 \\ &= ||w||^2 + t^2||y||^2 - 2t\mathrm{Re}(w, y) \\ &= d^2 + t^2||y||^2 - 2t\mathrm{Re}(w, y) \end{aligned}$$

which implies that

$$0 \le t^2||y||^2 - 2t\mathrm{Re}(w, y)$$

for all $t \in \mathbb{R}$. Hence $Re(w, y) = 0$. A similar argument using ti instead of t shows that $\mathrm{Im}(w, y) = 0$. Hence $w \in \mathcal{M}^\perp$. It remains to show that such a $w \in \mathcal{M}^\perp$ is unique. Suppose that $x = z + w_1$ and $x = z + w_2$ for some $w_2 \in \mathcal{M}^\perp$. Then clearly, $w_1 = w_2$. ☐

Remark. The projection theorem implies that that there is a natural isomorphism between $\mathcal{M} \oplus \mathcal{M}^\perp$ and $\mathcal{H}$ given by

$$\langle z, w \rangle \mapsto z + w$$

We will often surpress the isomorphosm and just write $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$

Definition 3.26. Let V be a normed vector space over $\mathbb{C}$ with norm given by $||\cdot||_X$. We say that $f : V \to \mathbb{C}$ is a **linear continuous functional** on V if

1. $f(x + \lambda y) = f(x) + \lambda f(y)$ for all $x, y \in V$ and $\alpha \in \mathbb{C}$

2. f is bounded. In other words, there exists a $C \in \mathbb{R}$ such that $|f(x)| \le C||x||_X$ for all x

The vector space of all linear continuous functionals on V is called the **dual space** of V and is denoted V^*. We can endow X^* with the norm

$$||f||_{X^*} = \sup_{||x||_X = 1} |f(x)|$$

Theorem 3.27 (Riesz Representation Theorem)**.** *Let $\mathcal{H}$ be a Hilbert space with inner product given by $(\cdot,\cdot)$. Let $T \in \mathcal{H}^*$. Then there exists a unique $y_T \in \mathcal{H}$ such that $T(x) = (y_T, x)$ for all $x \in \mathcal{H}$. Furthermore, $||y_T||_{\mathcal{H}} = ||T||_{\mathcal{H}^*}$*

Proof. Let $\mathcal{N}$ be the subset of $\mathcal{H}$ consisting of elements x such that $T(x) = 0$. By linearity and continuity, $\mathcal{N}$ is a closed subspace. Suppose first that $\mathcal{N} = \mathcal{H}$, Then $T(x) = 0 = (0, x)$ for all $x \in \mathcal{H}$.

Now suppose that $\mathcal{N} \neq \mathcal{H}$. By the Projection Theorem, $\mathcal{H} = \mathcal{N} \oplus \mathcal{N}^\perp$ and there exists a non-zero vector $x_0 \in \mathcal{N}^\perp$. We define $y_T = \overline{T(x_0)}||x_0||^{-2}x_0$. We claim that y_T is the desired element of $\mathcal{H}$. If $x \in \mathcal{N}$ then $T(x) = 0 = (y_T, x)$. Now set $x = \alpha x_0$ for some scalar α. Then

$$T(x) = T(\alpha x_0) = \alpha T(x_0) = \alpha T(x_0)\frac{||x_0||^2}{||x_0||^{-2}} = (\overline{T(x_0)}||x_0||^{-2}x_0, \alpha x_0) = (y_T, \alpha x_0) = (y_T, x)$$

Now, $T(\cdot)$ and $(y_T, \cdot)$ are both linear and they agree on $\mathcal{N}$ and at x_0. Hence they must also agree on the space spanned by $\mathcal{N} \cup \{ x_0 \}$. We claim that $\mathcal{H} = \operatorname{span}\{ \mathcal{N} \cup \{ x_0 \} \}$. Indeed, we can always write $y \in \mathcal{H}$ in the form

$$y = \left(y - \frac{T(y)}{T(x_0)} x_0 \right) + \frac{T(y)}{T(x_0)} x_0$$

which is clearly in $\operatorname{span}\{ \mathcal{N} \cup \{ x_0 \} \}$. Hence $T(x) = (y_T, x)$ for all $x \in \mathcal{H}$.

It remains to show that $||T||_{\mathcal{H}^*} = ||y_T||_{\mathcal{H}}$. We have that

$$||T||_{\mathcal{H}^*} = \sup_{||x||_{\mathcal{H}}=1} |T(x)| = \sup_{||x||_{\mathcal{H}}=1} |(y_T, x)| \leq \sup_{||x||_{\mathcal{H}}=1} ||y_T|| \, ||x||_{\mathcal{H}} = ||y_T||_{\mathcal{H}}$$

where we have used the Cauchy-Schwarz inequality. Conversely,

$$||T||_{\mathcal{H}^*} = \sup_{||x||_{\mathcal{H}}=1} |T(x)| \geq \left| T\left(\frac{y_T}{||y_T||} \right) \right| = \left(y_T, \frac{y_T}{||y_T||} \right) = ||y_T||_{\mathcal{H}}$$

whence $||T||_{\mathcal{H}^*} = ||y_T||_{\mathcal{H}}$. □

3.3 Orthonormal Bases

Definition 3.28. Let $\mathcal{H}$ be a Hilbert space with inner product given by $(\cdot, \cdot)$ and $\{ e_n \} \subseteq \mathcal{H}$ a subset. We say that $\{ e_n \}$ is a **complete system** if $(x, e_n) = 0$ for all $x \in \mathcal{H}$ and for all$e_n \in \{ e_n \}$ implies that $x = 0$.

Remark. Let $u_1, u_2, \ldots$ be linearly independent vectors. We can construct a set of orthonormal vectors from these vectors using the following:

$$w_n = u_n - \sum_{k=1}^{n-1} (v_k, u_n) v_k, \qquad v_n = \frac{w_n}{||w_n||}$$

The family of vectors v_n is then an orthonormal set. For example,

$$\begin{aligned} w_1 &= u_1 \\ w_2 &= u_2 - (v_1, u_2) v_1 \\ &\vdots \end{aligned}$$

This is referred to as the Gram-Schmidt process.

Theorem 3.29. *Let $\mathcal{H}$ be a separable Hilbert space. Then $\mathcal{H}$ contains a complete countable orthonormal system.*

Proof. Suppose $\mathcal{H}$ is separable and fix a countable dense subset $\{ x_n \}$. By removing certain elements from $\{ x_n \}$, we can obtain a linearly independent set whose span is equal to the span of $\{ x_n \}$. Applying the Gram-Schmidt process to the subcollection, we obtain a complete countable orthonormal system. □

Theorem 3.30. *Let $\mathcal{H}$ be a separable Hilbert space and $S = \{ x_n \}_{n \in \mathbb{N}}$ a complete countable orthonormal system. Then for all $y \in \mathcal{H}$ we have*

$$y = \sum_{n \in \mathbb{N}} (x_n, y) x_n \tag{3.1}$$

$$||y||^2 = \sum_{n \in \mathbb{N}} |(x_n, y)|^2 \tag{3.2}$$

Conversely, given a countable subset $\{ c_n \} \subseteq \mathbb{C}$ *such that*

$$\sum_{n \in \mathbb{N}} |c_n|^2 < \infty$$

Then

$$\sum_{n \in \mathbb{N}} c_n x_n$$

converges to an element of $\mathcal{H}$.

Proof. By Bessel's Inequality, we have for any finite set $N' \subseteq \mathbb{N}$

$$||y||^2 \geq \sum_{n \in N'} |(x_n, y)|^2$$

Hence for at most countably many $n \in \mathbb{N}$ we have $(x_n, y) \neq 0$. Order these $N_1, N_2, \ldots$. Now,

$$\sum_{i=1}^{M} |(x_{N_i}, y)|^2$$

is a bounded, monotone increasing sequence and hence converges as $M \to \infty$. Let $\{ y_n \}$ be the sequence given by

$$y_n = \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}$$

Then for any $n > m$ we have

$$||y_n - y_m||^2 = \left|\left| \sum_{j=m+1}^{n} (x_{N_j}, y) x_{N_j} \right|\right|^2 = \sum_{j=m+1}^{n} |(x_{N_j}, y)|^2$$

Therefore $\{ y_n \}$ is a Cauchy sequence by the previous result and must converge to some limit y' in $\mathcal{H}$. Now, given any N_l we have

$$\begin{aligned}
(y - y', x_{N_l}) &= \lim_{n \to \infty} \left(y - \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, x_{N_l} \right) \\
&= \lim_{n \to \infty} \left[(y, x_{N_l}) - \left(\sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, x_{N_l} \right) \right] \\
&= \lim_{n \to \infty} \left[(y, x_{N_l}) - \sum_{j=1}^{n} \overline{(x_{N_j}, y)} (x_{N_j}, x_{N_l}) \right] \\
&= \lim_{n \to \infty} [(y, x_{N_l}) - \overline{(x_{N_l}, y)}] \\
&= 0
\end{aligned}$$

Moreover, if we have $N \neq N_l$ for any l then

$$(y - y', x_N) = 0$$

Hence for all $x_N \in \{ x_n \}$ we have $(y - y', x_N)$. Since $\{ x_n \}$ is a complete orthonormal system, it follows that $y - y' = 0$ and, in particular, $y = y'$. Hence

$$y = \sum_{j=1}^{\infty} (x_{N_j}, y) x_{N_j}$$

Furthermore,

$$\begin{aligned}
0 &= \lim_{n\to\infty} \left|\left| y - \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j} \right|\right|^2 \\
&= \lim_{n\to\infty} \left(y - \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, y - \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j} \right) \\
&= \lim_{n\to\infty} \left[\left(y - \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, y \right) - \left(y - \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j} \right) \right] \\
&= \lim_{n\to\infty} \left[(y, y) - \left(y, \sum_{j=1}^{N} (x_{N_j}, y) x_{N_j} \right) - \left(\sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, y \right) + \left(\sum_{j=1}^{n} (x_{N_j}, y) x_{N_j}, \sum_{j=1}^{n} (x_{N_j}, y) x_{N_j} \right) \right] \\
&= \lim_{n\to\infty} \left[||y||^2 - \sum_{j=1}^{n} (x_{N_j}, y)(y, x_{N_j}) - \sum_{j=1}^{n} \overline{(x_{N_j}, y)} (x_{N_j}, y) + \sum_{i=1}^{n} \overline{(x_{N_i}, y)} \sum_{j=1}^{n} (x_{N_j}, y)(x_{N_i}, x_{N_j}) \right] \\
&= \lim_{n\to\infty} \left[||y||^2 - 2 \sum_{j=1}^{n} \overline{(x_{N_j}, y)} (x_{N_j}, y) + \sum_{i=1}^{n} \overline{(x_{N_i}, y)} (x_{N_i}, y) \right] \\
&= \lim_{n\to\infty} \left[||y||^2 - \sum_{j=1}^{n} \overline{(x_{N_j}, y)} (x_{N_j}, y) \right] \\
&= \lim_{n\to\infty} \left[||y||^2 - \sum_{j=1}^{n} |(x_{N_j}, y)|^2 \right]
\end{aligned}$$

and so $||y||^2 = \sum_{n\in\mathbb{N}} |(x_n, y)|^2$.

Conversely, suppose that $\{ c_n \} \subseteq \mathbb{C}$ is such that $\sum_{n\in\mathbb{N}} |c_n|^2 < \infty$. We need to show that

$$\sum_{n\in\mathbb{N}} c_n x_n$$

converges to an element of $\mathcal{H}$. Since $\mathcal{H}$ is a Hilbert space, it is complete so we just need to show that the sequence given by

$$y_n = \sum_{i=1}^{n} c_i x_i$$

is Cauchy. Note that

$$\sum_{i=1}^{M} |c_i|^2$$

is a bounded, monotone increasing sequence which converges to some limit as $M \to \infty$. We have that

$$||y_n - y_m||^2 = || \sum_{i=m+1}^{n} c_i x_i ||^2 = \sum_{i=m+1}^{n} |c_i|^2$$

and thus $\{ y_n \}$ is a Cauchy sequence. □

Chapter 4

Banach Spaces

4.1 Dual Spaces

Proposition 4.1. *Let X be a normed vector space over $\mathbb{R}$ (or $\mathbb{C}$) with norm given by $||\cdot||$. Let $f \in X^*$ be a continuous linear functional. Then the following are equivalent definitions for $||f||_{X^*}$:*

1. $I = \inf \{ c \mid |f(x)| \leq c||x|| \; \forall\, x \in X \}$
2. $S_1 = \sup_{||x|| \leq 1} |f(x)|$
3. $S_2 = \sup_{||x||=1} |f(x)|$
4. $S_3 = \sup_{x \neq 0} \frac{|f(x)|}{||x||}$

Proof. We clearly have $S_2 \leq S_1$. Furthermore, observe that

$$\frac{|f(x)|}{||x||} = \left| f\left(\frac{x}{||x||} \right) \right|$$

and so $S_3 \leq S_2$. Now suppose that $||x|| \leq 1$. Then

$$|f(x)| \leq \frac{|f(x)|}{||x||}$$

and so $S_3 \leq S_1$. It then follows that $S_1 = S_2 = S_3$. Now note that

$$|f(x)| \leq S_3 ||x||$$

for all $x \in X$ and thus $I \leq 3$. Conversely, by the definition of the supremum we have

$$I \geq \frac{|f(x)|}{||x||} \geq S_3 - \varepsilon_x$$

for some $\varepsilon_x \geq 0$. In particular, we can find a sequence x_n such that $\varepsilon_{x_n} = \frac{1}{n}$ and passing to the limit $n \to \infty$ yields $I \geq S_3$ and we are done. $\square$

Theorem 4.2. *Let X be a normed vector space over $\mathbb{C}$ with norm given by $||\cdot||_X$. Then X^* is a Banach space.*

Proof. We must show that X^* is complete with respect to its norm. Let $\{ f_n \} \subseteq X^*$ be a Cauchy sequence. By definition we have that for all $\varepsilon > 0$ there exists an N_ε such that for all $m, n > N_\varepsilon$,

$$||f_n - f_m||_{X^*} < \varepsilon$$

Fix $x \in X$. From the definition of the norm we have

$$\begin{aligned} |f_n(x) - f_m(x)| &= |(f_n - f_m)(x)| && (4.1) \\ &\leq ||f_n - f_m||_{X^*} ||x||_X = \varepsilon ||x||_X && (4.2) \end{aligned}$$

and hence $\{ f_n(x) \}$ is a Cauchy sequence of complex numbers. Since $\mathbb{C}$ is complete, $\{ f_n(x) \}$ converges to some point $f(x)$. We claim that $f(x)$ is the limit point of the original Cauchy sequence. We must first show that the function is linear. Let $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{C}$. Then

$$\begin{aligned} f(\alpha x_1 + \beta x_2) &= \lim_{n\to\infty} [f_n(\alpha x_1 + \beta x_2)] \\ &= \lim_{n\to\infty} [\alpha f_n(x_1) + \beta f_n(x_2)] \\ &= \alpha f(x_1) + \beta f(x_2) \end{aligned}$$

as required. Now passing to the limit $m \to \infty$ in Equation 4.2 we have

$$|f_n(x) - f(x)| \leq \varepsilon ||x||_X$$

for all $n \geq N_\varepsilon$. Now since x is arbitrary, $f(x)$ must be bounded. Hence

$$||f_n - f||_{X^*} \leq \varepsilon$$

for all $n \geq N_\varepsilon$. Hence the original Cauchy sequence $\{ f_n \}$ converges to the function f whence X^* is complete. □

Remark. If H is a Hilbert space, given any element $g \in H$ the map $f \mapsto (f, g)$ is a continuous functional. By the Riesz representation theorem, all functionals on H have this form and thus $H = H^*$.

Definition 4.3. Let X and Y be vector spaces over $\mathbb{R}$ (or $\mathbb{C}$) with norms given by $||\cdot||_X$ and $||\cdot||_Y$ respectively. We say that a linear map $f : X \to Y$ is an **isometry** if $||f(x)||_Y = ||x||_X$ for all $x \in X$. We say that X and Y are **isometric** if there exists a bijective isometry between them.

Lemma 4.4. *Let X and Y be vector spaces over $\mathbb{R}$ (or $\mathbb{C}$) with norms given by $||\cdot||_X$ and $||\cdot||_Y$ respectively. If $f : X \to Y$ is an isometry then f is injective.*

Proof. Since f is a linear map, it suffices to show that if $T(x) = 0$ for some $x \in X$ then $x = 0$. We have that

$$||x||_X = ||T(x)||_Y = ||0|| = 0$$

By the properties of the norm, it follows that $x = 0$ as required. □

Proposition 4.5. $(\ell^1)^*$ *is isometric to* ℓ^∞.

Proof. We claim that the following mapping is an isometry

$$J : \ell^\infty \to (\ell^1)^*$$
$$y \mapsto \lambda_y(x) = \sum_{n=1}^{\infty} y_n x_n$$

where $x \in \ell^1$. We must first show that this map is well defined. λ_y is clearly linear. Indeed, given $w, v \in \ell^1$ and $\alpha, \beta \in \mathbb{C}$ we have

$$\begin{aligned}\lambda_y(\alpha w + \beta v) &= \sum_{n=1}^{\infty} y_n(\alpha w_n + \beta v_n) \\ &= \alpha \sum_{n=1}^{\infty} y_n w_n + \beta \sum_{n=1}^{\infty} y_n v_n \\ &= \alpha \lambda_y(w) + \beta \lambda_y(v)\end{aligned}$$

λ_y is also bounded in ℓ^1:

$$|\lambda_y(x)| = |\sum_{n=1}^{\infty} y_n x_n| \leq \sum_{n=1}^{\infty} |y_n||x_n| \leq \sum_{n=1}^{\infty} ||y||_\infty |x_n| = ||y||_\infty ||x||_1$$

Now by the definition of the norm on $(\ell^1)^*$ we have $||\lambda_y||_{(\ell^1)^*} \leq ||y||_\infty$. To show the opposite inequality, we fix some $y \in \ell^\infty$. Fix $\varepsilon > 0$ and $n \in \mathbb{N}$ be such that $|y_n| \geq ||y||_\infty - \varepsilon$. Then

$$|\lambda_y(e_n)| = |y_n| \geq ||y||_\infty - \varepsilon \geq (||y||_\infty - \varepsilon)||e_n||_1$$

This implies that

$$||\lambda_y||_{(\ell^1)^*} = \sup_{x \in X \setminus \{0\}} \frac{|\lambda_y(x)|}{||x||_{\ell^1}} \geq \frac{|\lambda_y(e_n)|}{||e_n||_{\ell^1}} \geq ||y||_\infty - \varepsilon$$

But ε is arbitrary whence $||\lambda_y||_{(\ell^1)^*} = ||y||_\infty$.

Since J is an isometry, it must be injective so it remains to show that J is surjective. Fix some $\lambda \in (\ell^1)^*$. Setting $y_n = \lambda(e_n)$ we can see that $y = (y_1, y_2, \dots) \in \ell^\infty$. Consider $\mu = \lambda - \lambda_y \in (\ell^1)^*$. Clearly, $\mu(e_n) = 0$ for all n. Hence μ vanishes on the set of all linear combinations of e_n. But such a set is dense in ℓ^1 and, since μ is continuous, we must have that $\mu \equiv 0$ whence $\lambda = \lambda_y$. □

Proposition 4.6. *Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then $(\ell^p)^*$ is isometric to ℓ^q.*

Proof. We claim that the following map is an isometry

$$J : \ell^q \to (\ell^p)^*$$
$$y \mapsto \lambda_y(x) = \sum_{n=1}^{\infty} y_n x_n$$

where $x \in \ell^p$. We first claim that λ_y is linear. To this end, let $w, v \in \ell^p$ and $\alpha, \beta \in \mathbb{C}$. Then

$$\lambda_y(\alpha w + \beta v) = \sum_{n=1}^{\infty} y_n(\alpha w_n + \beta v_n) = \alpha \sum_{n=1}^{\infty} y_n w_n + \beta \sum_{n=1}^{\infty} y_n v_n = \alpha \lambda_y(w) + \beta \lambda_y(v)$$

To see that it is bounded, we apply Hölder's inequality:

$$|\lambda_y(x)| = \left|\sum_{n=1}^{\infty} y_n x_n\right| \leq \sum_{n=1}^{\infty} |y_n x_n| \leq \left(\sum_{k=1}^{\infty} |y_k|^q\right)^{\frac{1}{q}} \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}}$$
$$= ||y||_q ||x||_p$$

It thus follows that

$$||\lambda_y||_{(\ell^p)^*} \leq ||y||_q$$

Now fix non-zero $y \in \ell^p$ and define

$$x_n = \begin{cases} \overline{y_n}|y_n|^{q-2} & \text{if } y_n \neq 0 \\ 0 & \text{if } y_n = 0 \end{cases}$$

Then

$$\begin{aligned} ||x||_p^p = \sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} ||\overline{y_n}|y_n|^{q-2}|^p &= \sum_{n=1}^{\infty} |\overline{y_n}|^p \, |y_n|^{pq-2p} \\ &= \sum_{n=1}^{\infty} |\overline{y_n}|^p \, |y_n|^{q-p} \\ &= \sum_{n=1}^{\infty} |y_n|^q \\ &= ||y||_q^q \end{aligned}$$

Furthermore,

$$\lambda_y(x) = \sum_{n=1}^{\infty} y_n x_n = \sum_{n=1}^{\infty} y_n \overline{y_n} |y_n|^{q-2} = |y_n|^2 |y_n|^{q-2} = |y_n|^q = ||y||_q^q$$

Now applying this to the definition of the norm, we see that

$$||\lambda_y(x)||_{(\ell^p)^*} \geq \frac{|\lambda_y(x)|}{||x||_p} = \frac{||y||_q^q}{||y||_q^{\frac{q}{p}}} = ||y||_q^{q-\frac{q}{p}} = ||y||_q$$

whence $||\lambda_y(x)||_{(\ell^p)^*} = ||y||_q$. Hence J is an isometry and must be injective. It remains to show that J is surjective.

Fix some $\lambda \in (\ell^p)^*$. Set $y_n = \lambda(e_n)$ and let x_n be as previously defined. We claim that $J(y) = \lambda$. We must first show that $y \in \ell^q$. Choose some $N \in \mathbb{N}$ and let $x = (x_1, x_2, \ldots, x_N, 0, \ldots)$. Then

$$\sum_{n=1}^{N} |y_n|^q = \lambda(x) \leq ||\lambda||_{(\ell^p)^*} ||x||_p = ||\lambda||_{(\ell^p)^*} \left(\sum_{n=1}^{N} |y_n|^q\right)^{\frac{q}{p}}$$

whence $\sum_{n=1}^{N} |y_n|^q \leq ||\lambda||_{(\ell^p)^*}^q$. Now since N is arbitrary, we see that $y \in \ell^q$. Consider $\mu = \lambda - \lambda_y$. We have that $\mu(e_n) = 0$ for all n. μ thus vanishes on the set of all finite linear combinations of e_n. But this set is dense in ℓ^1 and μ is continuous so we must have that $\mu \equiv 0$ whence $\lambda = \lambda_y$. $\square$

Example 4.7. Similarly, $(L^p[a, b])^*$ is isometric to $L^q[a, b]$ when $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.

4.2 Hahn-Banach Theorem

Theorem 4.8. *Let X be a Banach space and $Y \subseteq X$ a linear subspace. Let $\lambda : Y \to \mathbb{C}$ be a linear functional such that $|\lambda(x)| \leq a||x||_X$ for all $x \in Y$. Then there exists some $\Lambda \in X^*$ such that $||\Lambda||_{X^*} \leq a$ and $\Lambda(x) = \lambda(x)$ for all $x \in Y$. In other words, any bounded linear functional on a subspace of a Banach space can be extended onto the whole Banach space without increasing the norm.*

Proof. For simplicity, we only consider the case where $a = 1$ and X is separable. We first reduce to the real case. Set $\ell(x) = \text{Re}\lambda(x)$. Then $\ell(x)$ is a bounded $\mathbb{R}$-linear functional. Clearly, $\lambda(x) = \ell(x) - i\ell(ix)$. Hence, if we can extend ℓ to L without increasing the norm then $\Lambda(x) = L(x) - iL(ix)$ extends $\lambda(x)$. Furthermore, Λ would be bounded with norm at most 1. Indeed, given $x \in X$, set $\alpha = \Lambda(x)/|\Lambda(x)|$ then

$$|\Lambda(x)| = \overline{\alpha}\Lambda(x) = \Lambda(\overline{\alpha}x) = L(\overline{\alpha}x) \leq ||\overline{a}x|| = ||x||$$

Were we have used the fact that $|\Lambda(x)|$ is real to discard the imaginary part when going from $\Lambda(\overline{a}x)$ to $L(\overline{a}x)$. We now extend ℓ by one real dimension. Fix some $z \in X\backslash Y$. We shall extend ℓ onto the $\mathbb{R}$-linear set

$$Z = \{\, y + tz \mid y \in Y, t \in \mathbb{R} \,\}$$

without increasing the norm. Since ℓ is $\mathbb{R}$-linear, we have that $\ell(y + tz) = \ell(y) + t\ell(z)$. Hence it suffices to define the value of $\ell(z)$. For convenience, we set $p = \ell(z)$. We need to find a p such that

$$|\ell(y) + tp| \leq ||y + tz|| \tag{4.3}$$

for all $y \in Y$ and $t \in \mathbb{R}$. To this end, let $y_1, y_2 \in Y$ and $\alpha, \beta \in \mathbb{R}$ positive. Then

$$\begin{aligned} |\ell(\beta y_1 - \alpha y_2)| \leq ||\beta y_1 - \alpha y_2|| &\leq ||\beta y_1 - \beta\alpha z|| + ||\beta\alpha z - \alpha y_2|| \\ &= \beta||y_1 - \alpha z|| + \alpha||\beta z - y_2|| \end{aligned}$$

From which we get

$$\frac{1}{\alpha}(\ell(y_1) - ||y_1 - \alpha z||) \leq \frac{1}{\beta}(\ell(y_2) + ||y_2 - \beta z||)$$

Now, y_1, y_2, α and β are arbitrary so

$$\sup_{\alpha>0, y_1\in Y} \frac{1}{\alpha}(\ell(y_1) - ||y_1 - \alpha z||) \leq \inf_{\beta>0, y_2\in Y} \frac{1}{\beta}(\ell(y_2) + ||y_2 - \beta z||)$$

Hence we may replace y_1 and y_2 with y and α and β with t to get

$$\frac{1}{t}(\ell(y) - ||y - tz||) \leq p \leq \frac{1}{t}(\ell(y) + ||y - tz||) \tag{4.4}$$

for some $p \in \mathbb{R}$. Replacing y by $-y$ also yields

$$-\frac{1}{t}(\ell(y) + ||y + tz||) \leq p \leq -\frac{1}{t}(\ell(y) - ||y + tz||) \tag{4.5}$$

Multiplying Equations (4.4) and (4.5) through by t we have

$$\ell(y) - ||y - tz|| \leq tp \leq \ell(y) + ||y - tz|| \tag{4.6}$$

$$-\ell(y) - ||y + tz|| \leq tp \leq -\ell(y) + ||y + tz|| \tag{4.7}$$

Multiplying Equation (4.6) through by -1 we get

$$-\ell(y) - ||y - tz|| \leq -tp \leq -\ell(y) + ||y - tz|| \tag{4.8}$$

Now, Equations (4.7) and (4.8) imply that the same inequality holds for all positive and negative real numbers. Since the case where $t = 0$ is trivial, the inequality holds for all $t \in \mathbb{R}$. It is clear that this inequality is equivalent to Equation (4.3) and hence we have extended ℓ by one real dimension.

With this in mind, we can easily extend ℓ onto the set

$$\{\, y + itz \mid y \in Y, t \in \mathbb{R} \,\}$$

without increasing the norm. This is done in exactly the same way as the extension by the one real dimension by replacing z with iz.

To finish off the proof, we fix some countable dense subset $x_1, x_2, \ldots$ in X. We can define a sequence of linear sets $Y \subseteq Y_1 \subseteq Y_2 \subseteq \cdots \subseteq X$ such that $x_n \subseteq Y_n$. By the previous results, we can extend ℓ onto $\bigcup_n Y_n$ without increasing the norm. Clearly, $\bigcup_n Y_n$ is dense in X. Now since ℓ is continuous, we may extend ℓ onto X and we are done. □

Corollary 4.9. *Let X be a Banach space and let $x \in X$. Then there exists an $\ell \in X^*$ such that $||\ell||_{X^*} = 1$ and $\ell(x) = ||x||_X$.*

Proof. Let Y be the one dimensional subspace of X spanned by x and define $\ell(tx) = t||x||_X$ for all $t \in \mathbb{C}$. We can then extend ℓ to all of X by the Hahn-Banach theorem. □

4.3 Second Dual space

Proposition 4.10. *Let X be a Banach space. Denote X^{**} as the dual of X^*. Then there exists an isometric embedding $J : X \to X^{**}$.*

Proof. Fix $x \in X$. Then x generates a linear functional $\lambda_x \in X^{**}$ given by $\lambda_x(\ell) = \ell(x)$ for some $\ell \in X^*$. We have that $|\lambda_x(\ell)| = |\ell(x)| \leq ||l||_{X^*}||x||_X$ whence $||\lambda_x||_{X^{**}} \leq ||x||$.

Now we may choose ℓ by Corollary 4.9 such that $\ell(x) = ||x||_X$ and $||\ell||_{X^*} = 1$. It follows that $|\lambda_x(\ell)| = |\ell(x)| = ||x||_X$. Hence

$$||\lambda_x||_{X^{**}} = \sup_{||\ell||_{X^*}=1} |\lambda_x(\ell)| = \sup_{||\ell||_{X^*}=1} |\ell(x)| \geq ||x||_X$$

If we write $J(x) = \lambda_x$ then $||J(x)||_{X^{**}} = ||\lambda_x||_{X^{**}} = ||x||_X$ whence J is an isometry. □

Definition 4.11. Let X be a Banach space. We say that X is **reflexive** if X is isometric to its second dual space X^{**}. In other words, X is reflexive if the above map J is surjective.

Example 4.12. For all $1 < p < \infty$, ℓ^p and L^p are reflexive. However they are not reflexive for $p = 1$ and ∞. Consider the subspace c_0 of ℓ^∞ consisting of all sequences whose limit is 0. Then it can be shown that $c_0^* = \ell^1$ but $(\ell^1)^* = \ell^\infty$ as we have seen.

4.4 Bounded linear operators

Definition 4.13. Let X and Y be normed vector spaces over $\mathbb{R}$ (or $\mathbb{C}$) with norms given by $||\cdot||_X$ and $||\cdot||_Y$ respectively. We say that $f : X \to Y$ is a **bounded linear operator** if

1. $f(x + \lambda y) = f(x) + \lambda f(y)$ for all $x_1, x_2 \in V$ and $\alpha \in \mathbb{C}$
2. f is bounded. In other words, there exists a $C \in \mathbb{R}$ such that $||f(x)||_Y \leq C||x||_X$ for all $x \in X$

We denote the vector space of all bounded linear operators from X to Y by $\mathcal{B}(X, Y)$. We can make $\mathcal{B}(X, Y)$ into a normed vector space with the operator norm

$$||f|| = \sup_{x \neq 0} \frac{||f(x)||_Y}{||x||_X}$$

Proposition 4.14. *Let X and Y be normed vector spaces over $\mathbb{R}$ (or $\mathbb{C}$) with norms given by $||\cdot||_X$ and $||\cdot||_Y$ respectively. Let $f \in \mathcal{B}(X, Y)$ be a bounded linear operator. Then the following are equivalent definitions for $||f||$:*

1. $I = \inf\{\, c \mid ||f(x)||_Y \leq c||x||_X \;\forall\, x \in X \,\}$
2. $S_1 = \sup_{||x||_X \leq 1} ||f(x)||_Y$
3. $S_2 = \sup_{||x||_X = 1} ||f(x)||_Y$
4. $S_3 = \sup_{x \neq 0} \frac{||f(x)||_Y}{||x||_X}$

Proof. This proof follows the same argument as the proof for Proposition 4.1. □

Theorem 4.15. *Let X be a normed vector space and Y a Banach space over $\mathbb{R}$ (or $\mathbb{C}$). Then $\mathcal{B}(X, Y)$ is a Banach space.*

Proof. This proof follows the same argument as the proof for Theorem 4.2. □

Theorem 4.16 (Banach-Steinhaus theorem)**.** *Let X be a Banach space and $\mathcal{F}$ a collection of bounded linear operators from X into a normed space Y such that the set $\{\, Tx \mid T \in \mathcal{F} \,\}$ is bounded for all $x \in X$. Then the set of norms $\{||T|| \mid T \in \mathcal{F}\}$ is bounded.*

4.5 Weak convergence

Definition 4.17. Let X be a Banach space and $\{\, x_n \,\} \subseteq X$ a sequence. We say that $\{\, x_n \,\}$ **weakly converges** to $x \in X$, denoted $\text{w-lim}\, x_n = x$, if $\ell(x_n) \to \ell(x)$ for all $\ell \in X^*$.

Remark. Let $\{\, x_n \,\} \subseteq X$ be a sequence in a Banach space. If $||x_n - x|| \to 0$ as $n \to \infty$ then clearly, x_n converges weakly to x. Indeed, let $f \in X^*$ be a continuous linear functional. Since f is continuous, we have

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

and so $\text{w-lim}_{n\to\infty} x_n = x$.

The converse does not hold. Indeed, consider the sequence of basis vectors $\{\, e_n \,\} \subseteq \ell^2$. Since ℓ^2 is an inner product space, any linear continuous functional takes the form $(y, \cdot)$ for some $y \in \ell^2$. Recall that y can be expressed as an infinite sum

$$y = \sum_{i=1}^{\infty} (y, e_i) e_i$$

and so

$$\lim_{n\to\infty}(y,e_n)=\lim_{n\to\infty}\left(\sum_{i=1}^{\infty}(y,e_i)e_i,e_n\right)=\lim_{n\to\infty}\sum_{i=1}^{\infty}\overline{(y,e_i)}(e_i,e_n)=\lim_{n\to\infty}\sum_{i=1}^{\infty}\overline{(y,e_i)}\delta_{in}=\lim_{n\to\infty}\overline{(y,e_n)}$$

Now let $c_i=(y,e_i)$. $y\in\ell^2$ means that

$$||y||^2=(y,y)=\left(\sum_{i=1}^{\infty}(y,e_i)e_i,\sum_{i=1}^{\infty}(y,e_i)e_i\right)=\sum_{i=1}^{\infty}\overline{(y,e_i)}\sum_{j=1}^{\infty}(y,e_i)(e_i,e_j)=\sum_{i=1}^{\infty}|(y,e_i)|^2<\infty$$

and so $\lim_{i\to\infty}\overline{(y,e_i)}=0$. Hence w-$\lim_{n\to\infty}(e_n)=0$.

On the other hand, $||e_n||_2=1$ for all n and so $\{e_n\}$ does not converge in ℓ^2.

Theorem 4.18. *Let X be a Banach space and $\{x_n\}\subseteq X$ a sequence. If $\{x_n\}$ converges weakly then $\{x_n\}$ is bounded.*

Proof. Recall that X embeds isometrically into X^{**} so we may consider each x_n and x as elements of X^{**}. Fix an $\ell\in X^*$. Then $\ell(x_n)=x_n(\ell)$ converges and is thus bounded. Now the Banach-Steinhaus theorem implies that the norms $||x_n||_{X^{**}}=||x_n||_X$ are bounded. $\square$

Chapter 5

Compactness in Banach Spaces

5.1 Preliminaries

Definition 5.1. Let (X, ρ) be a metric space and $K \subseteq X$ a subset. A collection $\hat{S}$ of subsets of X is said to be a **cover** of K if

$$K \subseteq \bigcup_{S \in \hat{S}} S$$

If every member of $\hat{S}$ is open then $\hat{S}$ is said to be an **open cover**. If a subset $\hat{S}_0 \subseteq \hat{S}$ is also a cover of K then $\hat{S}_0$ is said to be a **subcover** of K.

Definition 5.2. Let (X, ρ) be a metric space and $K \subseteq X$ a subset. We say that K is **compact** if every open cover of K has a finite subcover.

Definition 5.3. Let (X, ρ) be a metric space and $K \subseteq X$ a subset. We say that X is **sequentially compact** if any sequence of elements in K has a subsequence that converges to a limit in K.

Theorem 5.4. *Let (X, ρ) be a metric space and $K \subseteq X$ a subset. Then K is compact if and only if K is sequentially compact.*

Theorem 5.5. *Let (X, ρ) be a metric space and $K \subseteq X$ a subset. If K is compact then K is closed and bounded.*

Theorem 5.6. *Let $K \subseteq \mathbb{R}^n$ (or $\mathbb{C}^n$) be a subset. Then K is compact if and only if K is closed and bounded.*

Proposition 5.7. *Let (X, ρ) be a metric space and $K \subseteq X$ a compact set. If $K_0 \subseteq K$ is closed then K_0 is compact.*

Theorem 5.8. *Let (X, ρ) and (Y, σ) be metric spaces and $f : X \to Y$ a continuous mapping. If $K \subseteq X$ is compact then $f(X)$ is compact in Y.*

Corollary 5.9. *Let (X, ρ) be a metric space and $f : X \to \mathbb{R}$ a continuous mapping. If $K \subseteq X$ is compact then f attains its maximum and minimum on K.*

Proof. The image of K in $\mathbb{R}$ is compact and is thus closed and bounded. Hence $\max f(K)$ and $\min f(X)$ exist and are finite. $\square$

Definition 5.10. Let (X, ρ) be a metric space and $f : X \to \mathbb{R}$ (or $\mathbb{C}$) a function. We say that f is **uniformly continuous** if for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$.

Theorem 5.11. *Let (X, ρ) be a metric space and $K \subseteq X$ a compact subset. If $f : K \to \mathbb{R}$ is continuous then it is uniformly continuous.*

If a subset of a metric space is compact then it is closed and bounded. We shall soon see that the converse holds in a Banach space if and only if the metric space is finite dimensional. The following are counter examples for the infinite dimensional cases.

Example 5.12. We claim that for all $p \in [1, \infty]$, the closed unit ball B centered at zero in ℓ^p is not compact. Consider the sequence $\{ e_n \}$ consisting of the standard basis elements. Clearly, $e_n \in B$ for all n. Now if p is finite, we have

$$||e_n - e_m||_p = \left(\sum_{i=1}^{\infty} |e_n^{(i)} - e_m^{(i)}|^p \right)^{1/p} = 2^{1/p}$$

If p is infinite then

$$||e_n - e_m||_\infty = \sup_{i \in \mathbb{N}} |e_n^{(i)} - e_m^{(i)}| = 1$$

Hence the sequence is not Cauchy and cannot admit a convergent subsequence.

Example 5.13. We claim that the unit ball B centered at zero is not compact in $C[a, b]$. Let $I_n \subseteq [a, b]$ be a sequence of disjoint open intervals. For each n, let f_n be a function that is zero on $[a, b]\backslash I_n$ satisfying $0 \leq f(x) \leq 1$ for all $x \in I_n$ and $f(x) = 1$ for at least one $x \in I_n$ (for example, let $f(x)$ be a piecewise linear function in the shape of a triangle). We have that

$$||f_n|| = \sup_{x \in [a,b]} |f_n(x)| = 1$$

and so each $f_n \in B$. Furthermore,

$$||f_n - f_m|| = \sup_{x \in [a,b]} |f_n(x) - f_m(x)|| = 1$$

Hence the sequence $\{ f_n \}$ is not Cauchy and cannot admit a convergent subsequence.

5.2 Finite Dimensional Subspaces

Lemma 5.14. *Let V be an n-dimensional vector space over $\mathbb{R}^n$ (or $\mathbb{C}^n$). Then all norms on V are equivalent.*

Proof. We shall prove the lemma for $\mathbb{C}^n$. Let $|| \cdot ||$ be a norm on $\mathbb{C}^n$. We claim that $|| \cdot ||$ is equivalent to the norm $|| \cdot ||_1$. To this end, fix an $x \in \mathbb{C}^n$. We can always write $x = \sum_{i=1}^n x_i e_i$ for some $x_i \in \mathbb{C}$ where the e_i are the standard basis vectors. Then

$$||x|| = \left|\left| \sum_{i=1}^n x_i e_i \right|\right| \leq \sum_{i=1}^n |x_i| \, ||e_i|| \leq \left(\sum_{i=1}^n |x_i| \right) \left(\sum_{i=1}^n ||e_i|| \right) = C||x||_1$$

where $C = \sum_{i=1}^{\infty} ||e_i||$.

Conversely, let K be the set

$$K = \{ x \in \mathbb{C}^n \mid ||x||_1 = 1 \}$$

Then K is clearly closed and bounded and is thus compact. Since $|| \cdot ||$ is a norm, it is a continuous function and so $|| \cdot ||$ attains its minimum, say m, on the compact set K. Furthermore, $0 \notin K$ so, necessarily, $m > 0$. Hence for all $x \in K$ we have

$$||x|| \geq m$$

Now let $x \in \mathbb{C}^n$. Clearly, $x/||x||_1 \in K$ and so

$$\left|\left| \frac{x}{||x||_1} \right|\right| \geq m$$

and so $||x|| \geq m||x||_1$ for all $x \in \mathbb{C}^n$. Hence $|| \cdot ||$ is equivalent to $|| \cdot ||_1$. □

Proposition 5.15. *Let $(V, ||\cdot||)$ be an n-dimensional Banach space over $\mathbb{C}$. Then any closed ball in V is compact.*

Proof. We may assume that V is isomorphic to $\mathbb{C}^n$. Let B be a closed ball in V with respect to $|| \cdot ||$. By Lemma 5.14, all norms on V are equivalent so $|| \cdot ||$ induces the same topology on V as $|| \cdot ||_2$. Since B is closed and bounded in the $|| \cdot ||_2$ topology, it is compact. Hence B is compact with respect to $|| \cdot ||$. □

Proposition 5.16. *Let $(V, || \cdot ||)$ be a Banach space over $\mathbb{C}$. Then any finite dimensional subspace of V is closed.*

Proof. We may assume that V is isomorphic to $\mathbb{C}^n$. Then any linear subspace of V will be isomorphic to $\mathbb{C}^m$ for some $1 \leq m \leq n$. Since $\mathbb{C}^m$ is closed with respect to $|| \cdot ||_2$, it is then closed with respect to all norms. □

Lemma 5.17 (Riesz' Lemma)**.** *Let $(V, || \cdot ||)$ be a normed vector space over $\mathbb{R}$ (or $\mathbb{C}$) and $V_0 \subsetneq V$ a closed linear subspace. Then for all $\varepsilon \in (0, 1)$, there exists $x_0 \in V \backslash V_0$ such that $||x_0|| = 1$ and $||x_0 - x|| \geq 1 - \varepsilon$ for all $x \in V_0$.*

Proof. Fix some $x_1 \in V \backslash V_0$. The following number is greater than 0:

$$d = \inf_{x \in V_0} ||x_1 - x||$$

Indeed, if d were to equal zero then there would exist a sequence $y_n \in V_0$ such that $||x_1 - y_n|| \to 0$. Since V_0 is closed, this would imply that $x_1 \in V_0$ which is a contradiction. Now, $d < d/(1 - \varepsilon)$ so there exists $y \in V_0$ such that $d \leq ||x_1 - y|| \leq d(1 - \varepsilon)$. Define

$$x_0 = \frac{x_1 - y}{||x_1 - y||}$$

We claim that x_0 is the desired element. Clearly, $||x_0|| = 1$. Furthermore, since V_0 is a linear subspace, $y + ||x_1 - y||x \in V_0$ for all $x \in X_0$. Hence

$$||x_0 - x|| = \frac{x_1 - (y + ||x_1 - y||x)}{||x_1 - y||} = \frac{1 - \varepsilon}{d} ||x_1 - (y + ||x_1 - y||x)|| \geq \frac{1 - \varepsilon}{d} = 1 - \varepsilon$$

□

Theorem 5.18. *Let $(V, ||\cdot||)$ be a Banach space. Then a closed ball in V is compact if and only if V is finite dimensional over its base field.*

Proof. Assume that V is infinite dimensional. We shall exhibit the existence of a bounded sequence in the unit ball about zero that does not have any Cauchy subsequences. To this end, fix $x_1 \in V$ such that $||x_1|| = 1$. Denote by V_1 the one-dimensional linear subspace spanned by x_1. By Proposition 5.16, V_1 is closed and $V_1 \neq V$ since V is infinite dimensional. By Riesz' Lemma with $\varepsilon = 1/2$, there exists $x_2 \in V$ such that $||x_2|| = 1$ and $||x_2 - x|| \geq 1/2$ for all $x \in V_1$. In particular, $||x_2 - x_1|| \geq 1/2$.

Proceeding by induction, suppose that we have already constructed $x_1, \ldots, x_n \in V$ such that

$$||x_k|| = 1, \quad ||x_k - x_j|| \geq 1/2$$

for all $j, k = 1, \ldots, n$ and $j \neq k$. Denote by V_n the n-dimensional linear subspace spanned by $x_1, \ldots, x_n$. Proposition 5.16 again implies that V_n is closed and $V_n \neq V$ since V is infinite dimensional. Appealing again to Riesz' Lemma with $\varepsilon = 1/2$, there exists $x_{n+1} \in V$ such that $||x_{n+1}|| = 1$ and $||x_{n+1} - x|| \geq 1/2$ for all $x \in V_n$. We have thus constructed an infinite sequence $x_k \in V$ such that $||x_k|| = 1$ for all k and $||x_k - x_j|| \geq 1/2$ for all $j, k = 1, \ldots, n$ and $j \neq k$. Clearly, this sequence cannot contain any Cauchy subsequences. Furthermore, the sequence is contained in the unit ball centered at zero. Such a ball is therefore not sequentially compact whence it is not compact.

The converse is exactly Proposition 5.15. □

5.3 Total Boundedness

Definition 5.19. Let (X, ρ) be a metric space and $K \subseteq X$ a subset. We say that K is **totally bounded** if, given any $\varepsilon > 0$, K is contained in a union of finitely many balls of radius ε.

Example 5.20. Consider the subset of ℓ^∞ given by all the standard basis vectors e_i. Then the distance between any two e_i is 1 meaning the set is bounded. However, any open unit ball around an e_i will not contain any of the other e_i and so this set cannot possibly be totally bounded.

Theorem 5.21. *Let (X, ρ) be a complete metric space and $K \subseteq X$ a subset. Then K is compact if and only if it is closed and totally bounded.*

Proof. First suppose that K is compact. Then K is closed. Now given $\varepsilon > 0$, consider the open cover $\{ B_\varepsilon(x) \mid x \in K \}$ of K. Since K is compact, such a cover necessarily has a finite subcover. But this is exactly what it means for K to be totally bounded.

Conversely, assume that K is closed and totally bounded. Fix a sequence $\{ y_n \}$ in K. We need to construct a subsequence of $\{ y_n \}$ that converges to an element of K. Since K is totally bounded, there exists a finite cover of K by balls of radius 1. At least one of these balls contains infinitely many points of $\{ y_n \}$, label it $B_1(x_1)$. Let $K_1 = K \cap B_1(x_1)$. K_1 is also clearly totally bounded so there exists a finite cover of K by balls of radius $1/2$. At least one of these balls contains infinitely many points of $\{ y_n \}$, label it $B_{1/2}(x_2)$. Let $K_2 = K_1 \cap B_{1/2}(x_2)$. Continuing in this fashion, we construct a sequence of sets K_n satisfying

1. $\operatorname{diam} K_n \to 0$ as $n \to \infty$

2. for all n, K_n contains infinitely many points of the sequence $\{ y_n \}$

The second property allows us to select a subsequence $\{ y_{n_k} \}$ such that $y_{n_i} \in K_i$ for all $i \in \mathbb{N}$. The first condition ensures that $\{ y_{n_k} \}$ is Cauchy. Since X is complete, this sequence must converge to some limit in X. But K is closed and must contain such a limit point. Therefore K is sequentially compact whence it is compact. $\square$

Example 5.22 (Hilbert's Brick)**.** Consider the subset $K \subseteq \ell^2$ consisting of sequences satisfying

$$|x_1| \leq 2^{-1}, |x_2| \leq 2^{-2}, |x_3| \leq 2^{-3}, \ldots$$

We claim that K is compact. It suffices to show that K is closed and totally bounded. K is clearly closed as it is the infinite intersection of the closed sets

$$K_n = \{ x \in \ell^2 \mid |x_n| \leq 2^{-n} \}$$

Now fix $\varepsilon > 0$. We need to find a finite set $S \subseteq \ell^2$ such that $K \subseteq \bigcup_{x \in S} B_\varepsilon(x)$. Choose $n \in \mathbb{N}$ such that $2^{-n-1} \leq \varepsilon$. Consider the mapping

$$x = (x_1, x_2, \ldots, x_n, \ldots) \mapsto x^* = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$$

We first observe that $||x - x^*||_2 \leq \varepsilon/2$. Indeed, we have

$$||x - x^*||_2 = \left(\sum_{i=1}^{\infty} |x_i - x_i^*|^2 \right)^{1/2} = \left(\sum_{i=n+1}^{\infty} |x_i|^2 \right)^{1/2} \leq \left(\sum_{i=1}^{\infty} |2^{-i}|^2 \right)^{1/2} = \left(\sum_{i=1}^{\infty} 4^{-i} \right)^{1/2} = \left(\frac{4^{-n}}{3} \right)^{1/2}$$
$$= \frac{2^{-n}}{\sqrt{3}} = \frac{2^{-n-1}}{2\sqrt{3}} \leq \frac{\varepsilon}{2}$$

The set K^* of all points x^* is totally bounded since it is a closed bounded set in a finite dimensional space. Hence there exists a finite set S such that $K^* \subseteq \bigcup_{x \in S} B_{\varepsilon/2}(x)$. Since $||x - x^*|| \leq \varepsilon/2$, it is then clear that $K \subseteq \bigcup_{x \in S} B_\varepsilon(x)$.

5.4 Arzela-Ascoli Theorem

Definition 5.23. Let $\Phi \subseteq C[a, b]$ be a family of functions. We say that Φ is **uniformly bounded** if there exists $M \in \mathbb{R}$ such that $|\Phi(x)| \leq M$ for all $\phi \in \Phi$ and $x \in [a, b]$.

Definition 5.24. Let $\Phi \subseteq C[a, b]$ be a family of functions. We say that Φ is **equicontinuous** if, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|\phi(x_1) - \phi(x_2)| \leq \varepsilon$ for all $\phi \in \Phi$ and $x_1, x_2 \in [a, b]$ such that $|x_1 - x_2| < \delta$.

Theorem 5.25 (Arzela-Ascoli)**.** *Let $\Phi \subseteq C[a, b]$ be a subset of the continuous real valued functions on $[a, b]$. Then Φ is totally bounded if and only if it is uniformly bounded and equicontinuous.*

Proof. First suppose that Φ is totally bounded. Then, clearly, Φ is bounded. In other words, $\Phi \subseteq B_r(0)$ for some $r > 0$. Hence Φ is uniformly bounded. We must now show that Φ is equicontinuous. Fix some $\varepsilon > 0$. Since Φ is totally bounded, there exist finitely many functions $f_1, \ldots, f_n)$ such that $\Phi \subseteq \bigcup_{k=1}^{n} B_{\varepsilon/3}(f_k)$. Now, each f_k is continuous by hypothesis and $[a, b]$ is compact. Hence each f_k is uniformly continuous and thus, for each k, there exists a $\delta_k > 0$ such that

$$|f_k(x) - f_k(y)| \leq \varepsilon/3 \quad \text{whenever } |x - y| \leq \delta_k$$

Let $\delta = \min_k \delta_k$. Then for all k we have

$$|f_k(x) - f_k(y)| \leq \varepsilon/3 \quad \text{whenever } |x - y| \leq \delta$$

$\square$