LINEAR ANALYSIS 1

MATTHEW FERRIER

1. Introduction

Lecture notes for the 2017/18 course, delivered by Dr.Vladimir Kisil at the University of Leeds. This course covers the following, (taken straight from the Leeds module catalogue):

(1) Normed spaces, bounded linear operators on a Banach space, dual spaces, Hahn-Banach theorem, Zorn's lemma. Use of sequence spaces. The Banach space C(X) for a compact space X.

(2) Basic measure theory, up to the construction of the Lebesgue measure on the real line. Complex measures and measurable functions. Dominated convergence theorem. Product measures. Fubini theorem.

(3) Definition of spaces of Lebesgue integrable functions, and proof that with the standard norm they form a Banach space. Dual spaces. The Radon-Nikodym Theorem. The conjugate index theorem.

(4) The Banach space M(X) of regular Borel measures on a compact space X. Proof that the dual of C(X) is M(X).

(5) Applications to Fourier series.

Notation 1.1. We will be working with the fields $\mathbb{R}$ and $\mathbb{C}$. To avoid repetition we will use $\mathbb{K}$ to denote either $\mathbb{R}$ or $\mathbb{C}$

2. Normed and Banach Spaces

2.1. **Useful inequalities.** We begin by deriving some inequalities which will prove to be invaluable in our discussion of normed spaces.

Lemma 2.1. *(Young's Inequality)*
Let p and q be real numbers such that $1 < p, q < \infty$ and related by $\frac{1}{p} + \frac{1}{q} = 1$. Then;

$$|ab| \leqslant \frac{|a|^p}{p} + \frac{|b|^q}{q},$$

for any $a, b \in \mathbb{C}$.

Proof. We will display two proofs of this inequality. The first shall be analytic.

Note that it is sufficient to prove the inequality for positive reals $a = |a|$ and $b = |b|$, further note that if $1 < p < 2$, then $q > 2$, and similarly if $p > 2$, then $1 < q < 2$. Without loss of generality take $q \geqslant 2$ and define $\phi(t) = t^m - mt$, from the derivative $\phi' = m(t^{m-1} - 1)$, we find the function's only critical point $t = 1$ on $[0, \infty)$, which is its maximum. So $\phi(1) \geqslant \phi(t)$ for all $t \neq 1$. Put $t = \frac{|a|^p}{|b|^q}$ and $m = \frac{1}{p}$, then;

$$\phi(1) = 1 - \frac{1}{p} \geqslant \phi(t) = \left(\frac{|a|^p}{|b|^q}\right)^{\frac{1}{p}} - \frac{1}{p}\frac{|a|^p}{|b|^q}.$$

Since $1 - \frac{1}{p} = \frac{1}{q}$, it follows that

$$\frac{1}{q} \geqslant \frac{|a|}{|b|^{\frac{q}{p}}} - \frac{1}{p}\frac{|a|^p}{|b|^q}.$$

Multiplication by $|b|^q$ gives

$$\frac{1}{q}|b|^q \geqslant |a||b|^{q-\frac{q}{p}} - \frac{1}{p}|a|^p \iff \frac{1}{q}|b|^q + \frac{1}{p}|a|^p \geqslant |a||b|^{q-\frac{q}{p}} = |a||b|.$$

Where the last part follows since $q - \frac{q}{p} = q(1 - \frac{1}{p}) = q \cdot \frac{1}{q} = 1$. We now give a geometric proof.

Consider the curve $y = x^{p-1} \iff x = y^{q-1}$ in the xy plane. Comparing areas on the figure, we see that $S_1 + S_2 \geqslant ab$ for any positive $a, b \in \mathbb{R}$. We may integrate to find the values of S_1 and S_2, and we see

$$S_1 = \int_0^a x^{p-1}dx = \frac{a^p}{p},\ S_2 = \int_0^b y^{q-1}dy = \frac{b^q}{q}.$$

This completes the proof. $\square$

Proposition 2.2. *(Hölder's Inequality)*
Let p and q be real numbers such that $1 < p, q < \infty$ and related by $\frac{1}{p} + \frac{1}{q} = 1$. For $n \geqslant 1$ and $u, v \in \mathbb{K}^n$, we have

$$\sum_{j=1}^{n} |u_j v_j| \leqslant \left(\sum_{j=1}^{n} |u_j|^p\right)^{\frac{1}{p}} \left(\sum_{j=1}^{n} |v_j|^p\right)^{\frac{1}{p}}.$$

Proof. First we note that if u or v are 0, the inequality is obviously true, so we may assume that $u, v \neq 0$.

For clarity, we will write $\|u\|_p = \left(\sum_{j=1}^{n} |u_j|^p\right)^{\frac{1}{p}}$ and $\|v\|_q = \left(\sum_{j=1}^{n} |v_j|^q\right)^{\frac{1}{q}}$, for $u, v \in \mathbb{K}^n$. (We will see why we choose this notation later).

For $1 \leqslant i \leqslant n$, define $a_i = \frac{u_i}{\|u\|_p}$ and $b_i = \frac{v_i}{\|v\|_q}$. Then;

$$|a_i b_i| \leqslant \frac{|a_i|^p}{p} + \frac{|b_i|^q}{q} \text{ (by Young's inequality).}$$

Summing for $1 \leqslant i \leqslant n$, we get

$$\sum_{i=1}^{n} |a_i b_i| \leqslant \sum_{i=1}^{n} \frac{|a_i|^p}{p} + \sum_{i=1}^{n} \frac{|b_i|^q}{q}.$$

Putting in our definitions of a_i and b_i gives

$$\frac{1}{\|u||_p \|v||_q} \sum_{i=1}^{n} |u_i v_i| \leqslant \frac{1}{p(\|u\|_p)^p} \sum_{i=1}^{n} |u_i|^p + \frac{1}{q(\|v\|_q)^q} \sum_{i=1}^{n} |v_i|^q = \frac{(\|u\|_p)^p}{p(\|u\|_p)^p} + \frac{(\|v\|_q)^q}{q(\|v\|_q)^q}.$$

Finally multiply by $\|u\|_p \|v\|_q$ to get

$$\sum_{k=1}^{n} |u_i v_i| \leqslant \frac{1}{p}(\|u\|_p)(\|v\|_q) + \frac{1}{q}(\|v\|_q)(\|u\|_p) = (\frac{1}{p} + \frac{1}{q})(\|u\|_p \|v\|_q) = \|u\|_p \|v\|_q.$$

$\square$

Using Hölder's inequality, we may derive the following.

Proposition 2.3. *(Minkowski's Inequality)*
Let p be a real number such that $1 < p < \infty$, $n \geqslant 1$ and $u, v \in \mathbb{K}^n$. Then

$$\left(\sum_{j=1}^{n} |u_j + v_j|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{j=1}^{n} |u_j|^p\right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} |v_j|^p\right)^{\frac{1}{p}}.$$

Proof. We have

$$\sum_{k=1}^{n} |u_k + v_k|^p = \sum_{k=1}^{n} |u_k + v_k||u_k + v_k|^{p-1} \leqslant \sum_{k=1}^{n} |u_k||u_k + v_k|^{p-1} + \sum_{k=1}^{n} |v_k||u_k + v_k|^{p-1}$$

by the triangle inequality, and Hölder's inequality tells us that

$$\sum_{k=1}^{n} |u_k||u_k+v_k|^{p-1} \leqslant \left(\sum_{k=1}^{n} |u_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |u_k + v_k|^{(p-1)q}\right)^{\frac{1}{q}} = \|u\|_p \left(\sum_{k=1}^{n} |u_k + v_k|^p\right)^{\frac{1}{q}} = \|u\|_p \left(\|u + v\|_p\right)^{\frac{p}{q}}.$$

We may obtain a similar result for the second term and combining these gives us

$$(\|u+v\|_p)^p \leqslant \|u_p\|(\|u+v\|_p)^{\frac{p}{q}} + \|v_p\|(\|u+v\|_p)^{\frac{p}{q}}.$$

Division by $(\|u+v\|_p)^{\frac{p}{q}}$ gives us, $\|(\|u+v\|_p)^{p-\frac{p}{q}} = \|u+v\|_p \leqslant \|u\|_p + \|v\|_p$. □

Note 2.4. It would be a useful exercise to verify these inequalities for infinite series.

2.2. **Normed Spaces.** The main objects of study in this module are normed spaces, these spaces are where algebra and analysis meet.

Definition 2.5. A *norm* on a vector space V is a map $\|\cdot\| : V \to [0,\infty)$ such that;

(i) $\|u\| = 0$ if and only if $u = 0$
(ii) $\|\lambda u\| = |\lambda| \cdot \|u\|$ for $\lambda \in \mathbb{K}$ and $u \in V$
(iii) $\|u+v\| \leqslant \|u\| + \|v\|$ for all $u, v \in V$

We call a vector space equipped with a norm a *normed vector space*, or simply a *normed space*.

Suppose that $(E, \|\cdot\|_E)$ is a normed space and suppose that $F \subset E$ is a subspace. We can define a norm on F in a very natural way; simply view each element of F as an element of E, and take it's norm in E. Unless stated otherwise, we should consider this to be the norm on any given subspace.

Having just introduced them, we should probably see some examples of normed spaces.

Example 2.6. For $1 \leqslant p < \infty$, we may define a norm $||\cdot||_p$ on $\mathbb{K}^n$ by

$$||u||_p = \left(\sum_{j=1}^{n} |u_j|^p\right)^{\frac{1}{p}}, \text{ for } u \in \mathbb{K}^n.$$

Checking that $\|u\|_p \in [0,\infty)$ for all $u \in \mathbb{K}^n$ is easy. For the other properties, let $u \in \mathbb{K}^n$ and $\lambda \in \mathbb{K}$, then:

(i) $\|u\|_p = 0 \iff$ every coordinate of u is 0 $\iff u = 0$.
(ii) $\|\lambda u\|_p = \left(\sum_{j=1}^n |\lambda u_j|^p\right)^{\frac{1}{p}} = \left(\sum_{j=1}^n |\lambda|^p |u_j|^p\right)^{\frac{1}{p}} = \left(|\lambda|^p \sum_{j=1}^n |u_j|^p\right)^{\frac{1}{p}} = |\lambda| \left(\sum_{j=1}^n |u_j|^p\right)^{\frac{1}{p}}$, which is precisely $|\lambda|\|u\|_p$.
(iii) The triangle inequality follows immediately from Minkowski's inequality.

This should explain our choice of notation in the previous section.

Example 2.7. Let X be a compact metric space, then there is a norm on the vector space

$$C_{\mathbb{K}}(X) = \{f : X \to \mathbb{K};\ f \text{ is continuous}\}$$

given by the formula

$$\|f\| = \sup\{|f(x)|;\ x \in x\}.$$

First recall that if X is a compact metric space, and the function $g : X \to Y$ is continuous, then the image $g(X)$ is a compact in Y. Therefore, the set $g(X)$ is bounded. Since we are considering a space of continuous functions on a compact space, our norm is properly defined and only takes values in $[0,\infty)$.

All that remains now is to check the 3 properties of a norm. Let $f, g \in C_{\mathbb{K}}(X)$ and $\lambda \in \mathbb{K}$, then:

(i) $\|f\| = 0 \iff f(x) = 0$ for all points $x \in X \iff f = 0$.
(ii) $\|\lambda f\| = \sup\{|\lambda f|;\ x \in X\} = \sup\{|\lambda||f|;\ x \in X\} = |\lambda| \sup\{|f|;\ x \in X\} = |\lambda|\|f\|$.
(iii) $\|f+g\| = \sup\{|f(x)+g(x)|;\ x \in X\} \leqslant \sup\{|f(x)|+|g(x)|;\ x \in X\}$ by the triangle inequality. Then,

$$\|f+g\| \leqslant \sup\{|f(x)|;\ x \in X\} + \sup\{|g(x)|;\ x \in X\} = \|f\| + \|g\|.$$

Proposition 2.8. *Let V be a normed space, then there is a metric on the space defined by the formula*

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

Proof. All we need to do is check that our function d satisfies the 3 axioms for a metric. Let $u, v, w \in V$;

(i) $d(u, v) = 0 \iff \|u - v\| = 0 \iff u = v$.

(ii) $\|u - v\| = \|(-1)(v - u)\| = |-1| \cdot \|v - u\| = \|v - u\|$. And so $d(u, v) = d(v, u)$.

(iii) $\|u - w\| = \|(u - v) + (v - w)\| \leqslant \|u - v\| + \|v - w\|$. Hence $d(u, w) \leqslant d(u, v) + d(v, w)$.

When proving (ii) and (iii), notice where we have used the axioms for a norm. □

The metric described above is called the metric induced by the norm. The fact that every normed space is a metric space is very useful, as it allows us to talk about concepts such as continuity and convergence in normed spaces.

Definition 2.9. Let V be a normed vector space, a sequence (u_n) in V *converges in norm* to a limit $u \in V$ if for all $\epsilon > 0$, there exists $M \in \mathbb{N}$ such that $\|u_n - u\| < \epsilon$, whenever $n \geqslant M$.

Equivalently, (u_n) *converges in norm* to $u \in V$ if and only if

$$\lim_{n\to\infty} \|u_n - u\| = 0.$$

2.3. **Banach Spaces.** Recall what it means for a metric space to be complete. We may extend this notion to completeness of normed vector spaces.

Definition 2.10. Let V be a normed vector space, a sequence (u_n) in V is called a *Cauchy sequence* if for all $\epsilon > 0$, there exists $M \in \mathbb{N}$ such that $\|u_m - u_n\| < \epsilon$, whenever $m, n \geqslant M$.

Definition 2.11. *(Banach Space)*

(i) A normed space V is said to be *complete* if every Cauchy sequence in V, converges *in norm* to some limit in V.

(ii) A complete normed space is called a *Banach space.*

Definition 2.12. Let $1 < p < \infty$, we define

$$\ell_p = \left\{ (x_n) \text{ in } \mathbb{K};\ \sum_{n\in\mathbb{N}} |x_n|^p < \infty \right\}.$$

Note 2.13. Via careful use of Minkowski's inequality, we may show that ℓ_p is a vector space, and it becomes a normed space under the $\|\cdot\|_p$ norm.

Theorem 2.14. *For $1 < p < \infty$, the space ℓ_p is a Banach space.*

Proof. Most completeness proofs follow a similar recipe:

(i) Consider a Cauchy sequence in the space we are considering, construct a point which "should be" the limit of the sequence.

(ii) Prove that this point is an element of our space.

(iii) Prove that this point really is the limit of the sequence.

Take a Cauchy sequence $(x^{(n)})$ in ℓ_p, we must show that this converges to some element of ℓ_p. For each n, $x^{(n)} \in \ell_p$ is a sequence of scalars, say $x^{(n)} = (x_k^{(n)})_{k\in\mathbb{N}}$. As $(x^{(n)})$ is Cauchy, for all $\epsilon > 0$, there is $M \in \mathbb{N}$ such that $\|x^{(a)} - x^{(b)}\|_p < \epsilon$. whenever $a, b > M$.

For fixed k, we have

$$|x_k^{(a)} - x_k^{(b)}| \leqslant \left(\sum_{i=1}^{\infty} |x_i^{(a)} - x_i^{(b)}|^p \right)^{\frac{1}{p}} = \|x^{(a)} - x^{(b)}\|_p < \epsilon,$$

whenever $a, b > M$. So the scalar sequence $x^{(n)} = (x_k^{(n)})_{n \in \mathbb{N}}$ is Cauchy in $\mathbb{K}$. By completeness of $\mathbb{K}$, this sequence converges to some $x_k \in \mathbb{K}$. Let $x = (x_k)_{k \in \mathbb{N}}$, this x appears to be a good candidate for the limit of $(x^{(n)})$.

First we check that $x^{(n)} - x \in \ell_p$, for $\epsilon > 0$, find $N \in \mathbb{N}$ such that $\|x^{(n)} - x^{(m)}\|_p < \epsilon$ for all $n, m \geqslant N$. The for for fixed k, we have

$$\sum_{i=1}^{k} |x_i^{(n)} - x_i^{(m)}|^p \leqslant \|x^{(n)} - x^{(m)}\|_p^p < \epsilon^p.$$

Let $m \to \infty$, then; $\sum_{i=1}^{k} |x_i^{(n)} - x_i|^p < \epsilon^p$.
Let $k \to \infty$, then; $\sum_{i=1}^{\infty} |x_i^{(n)} - x_i|^p < \epsilon^p$.

Thus $x^{(n)} - x \in \ell_p$, and because ℓ_p is a vector space, $x = x^{(n)} - (x^{(n)} - x) \in \ell_p$. Finally, we saw above that for any $\epsilon > 0$, there is $N \in \mathbb{N}$ such that $\|x^{(n)} - x\|_p < \epsilon$, so $x^{(n)} \to x$. □

Definition 2.15. We define vector spaces and norms as follows;

$$\begin{aligned}
\ell_1 &= \left\{ (x_n); \sum_{n \in \mathbb{N}} |x_n| < \infty \right\} & \|(x_n)\|_1 &= \sum_{n=1}^{\infty} |x_n|. \\
\ell_\infty &= \left\{ (x_n); \sup_{n \in \mathbb{N}} |x_n| < \infty \right\} & \|(x_n)\|_\infty &= \sup_{n \in \mathbb{N}} |x_n| \\
c_0 &= \left\{ (x_n); \lim_{n \to \infty} x_n = 0 \right\} & \|(x_n)\|_\infty &= \sup_{n \in \mathbb{N}} |x_n|
\end{aligned}$$

Note 2.16. Since convergent sequences are bounded, $c_0 \subset \ell_\infty$, so we use the $\|\cdot\|_\infty$ norm on c_0.

Theorem 2.17. *$(\ell_\infty, \|\cdot\|_\infty)$, $(\ell_1, \|\cdot\|_1)$ and $(c_0, \|\cdot\|_\infty)$ are Banach spaces.*

2.4. **Bounded Linear Operators.** On their own, Banach spaces aren't particularly useful, however things start to become very interesting when we consider maps between spaces. Recall the following definition from linear algebra.

Definition 2.18. Let E and F be vector spaces over $\mathbb{K}$, a mapping $f : E \to F$ is said to be a *linear map* if for all $x, y \in E$ and $\lambda \in \mathbb{K}$, we have

(i) $f(x + y) = f(x) + f(y)$
(ii) $f(\lambda x) = \lambda f(x)$

Note 2.19. We will be calling linear maps *linear operators.*

It will be useful to recall the following result.

Lemma 2.20. *Let E and F be vector spaces and let $f : E \to F$ be a linear map, then $f(0_E) = 0_F$.*

Proof. $f(0_E) = f(0_E + 0_E) = f(0_E) + f(0_E)$ by linearity. Now subtract $f(0_E)$ from both sides to see that $0_F = f(0_E)$. □

Definition 2.21. A linear operator $T : E \to F$ between two normed $\mathbb{K}$-vector spaces is *bounded* if there exists $M \in \mathbb{R}^+$ such that

$$\|T(x)\|_F \leqslant M\|x\|_E, \text{ for all } x \in E.$$

We write $B(E, F)$, for the set of all *bounded linear operators* from E to F.

Remember that linear maps are between vector spaces over the same base field.

Definition 2.22. The set $B(E, F)$ is in fact a vector space, and we define a norm (the *operator norm*) by

$$\|T\| = \sup\left\{\frac{\|T(x)\|_F}{\|x\|_E}; x \in E, x \neq 0\right\}.$$

It would be worthwhile to check that $B(E, F)$ is really a vector space, and that this function really is a norm.

Proposition 2.23. *For a linear operator $T : E \to F$ between normed spaces $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$, the following are equivalent;*

(i) T is continuous.

(ii) T is continuous at 0.

(iii) T is bounded.

Proof. It is trivial that $(i) \Rightarrow (ii)$.

$(ii) \Rightarrow (iii)$ T is continuous at 0, so for any $\epsilon > 0$, there is $\delta > 0$ such that $\|x-0\|_E < \delta \Rightarrow \|T(x)-T(0)\|_F < \epsilon$. Taking $\epsilon = 1$, we can find $\delta > 0$ such that $\|x\|_E < \delta \Rightarrow \|T(x)\|_F < 1$. Now take $M = \frac{2}{\delta}$ and let $x \in E\backslash\{0\}$, then

$$\delta > \frac{\delta}{2} = \frac{\delta}{2}\left\|\frac{x}{\|x\|_E}\right\| = \left\|\frac{\delta x}{2\|x\|_E}\right\|_E.$$

So $\left\|T\left(\frac{\delta x}{2\|x\|_E}\right)\right\|_F < 1$, and by linearity of T, this is equivalent to

$$\frac{\delta}{2\|x\|_E}\|T(x)\|_F < 1 \iff \|T(x)\|_F < \frac{2\|x\|_E}{\delta} = M\|x\|_E.$$

Furthermore, if $x = 0$, then $\|T(x)\|_F = 0 \leqslant M\|x\|_E$, so the inequality holds for all $x \in E$, so T is bounded.

$(iii) \Rightarrow (i)$ T is bounded, so we can find $M > 0$ such that $\|T(x)\|_F \leqslant M\|x\|_E$ for all $x \in E$. Take $\epsilon > 0$ and let $\delta = \frac{\epsilon}{M}$. Then for $x, y \in E$, with $\|x - y\|_E \leqslant \delta$, we have

$$\|T(x) - T(y)\|_F = \|T(x - y)\|_F \leqslant M\|x - y\|_E \leqslant M\frac{\epsilon}{M} = \epsilon.$$

So T is continuous.

□

As the following theorem will show, our spaces of bounded (continuous!) linear operators may have special properties.

Theorem 2.24. *Let E be a normed space and F be a Banach space, then $B(E, F)$ is a Banach space.*

Proof. Let (T_n) be a Cauchy sequence of operators in $B(E, F)$. Given $x \in E$, consider the sequence $(T_n(x))$ in F, note that if $x = 0$, $\|T_n(x) - T_m(x)\|_F = 0 < \epsilon$ for all $\epsilon > 0$ and $n, m \in \mathbb{N}$.

For $x \in E\backslash\{0\}$, let $\epsilon > 0$ be given, then there is $N \in \mathbb{N}$ such that

$$n, m > N \Rightarrow \frac{\epsilon}{\|x\|_E} > \|T_n - T_m\| \geqslant \frac{\|T_n(x) - T_m(x)\|_F}{\|x\|_E}$$

So for each $x \in E$, given $\epsilon > 0$, we can find $M \in \mathbb{N}$ such that $\|T_n(x) - T_m(x)\|_F < \epsilon$ whenever $n, m > M$, so $(T_n(x))$ is a Cauchy sequence in F. By completeness of F, $(T_n(x))$ converges to some limit in F. Define a mapping $T : E \to F$ by

$$T(x) = \lim_{n\to\infty} T_n(x).$$

We will verify that this T is in fact the limit of the sequence (T_n). For this, we must check that our T is an element of $B(E, F)$, then check that (T_n) converges in norm to T. First we check linearity, let $\alpha, \beta \in \mathbb{K}$ and $u, v \in E$, then

$$T(\alpha u + \beta v) = \lim_{n\to\infty} T_n(\alpha u + \beta v) = \alpha \lim_{n\to\infty} T_n(u) + \beta \lim_{n\to\infty} T_n(v) = \alpha T(u) + \beta T(v).$$

Now we check that T is bounded. Since (T_n) is a Cauchy sequence, for all $\epsilon > 0$, there is $N \in \mathbb{N}$ such that

$$n, m > N \Rightarrow \epsilon > \|T_n - T_m\| > \big|\|T_n\| - \|T_m\|\big|,$$

where the last inequality follows from the reverse triangle inequality. Hence $(\|T_n\|)$ is a Cauchy sequence in $\mathbb{R}$ and by completeness, the limit $\lim_{n\to\infty} \|T_n\| = M$ exists and is a real number. For all $x \in E$, we have

$$\|T(x)\|_F = \lim_{n\to\infty} \|T_n(x)\| \leqslant \lim_{n\to\infty} \|T_n\|\|x\|_E = M\|x\|_E,$$

so T is bounded.

Finally we must check that (T_n) converges in norm to T. Since $(T_n(x))$ is a Cauchy sequence for all $x \in E$, it is certainly Cauchy for the $x \in E$ such that $\|x\|_E = 1$. So for $x \in E$ with $\|x\|_E = 1$, given any $\epsilon > 0$, we may find $N \in \mathbb{N}$ such that $\|T_n(x) - T_m(x)\|_F < \epsilon$ whenever $n, m > N$. Taking the limit as $m \to \infty$ gives $\|T_n(x) - T(x)\|_F < \epsilon$, and since this holds for all such x, it follows that

$$\sup\{\|T_n(x) - T_m(x)\|_F; x \in E, \|x\|_E = 1\} < \epsilon,$$

so $T_n \to T$ in norm. □

Definition 2.25 (*Invertible operator*)**.** An operator $T \in B(E, F)$ is called invertible if there exists $S \in B(F, E)$ such that,

(i) $S \circ T = I_E$

(ii) $T \circ S = I_F$

If this is the case, we write $S = T^{-1}$. An invertible operator is called an *isomorphism* of the vector spaces E and F.

Definition 2.26 (*Isometry*)**.** An operator $T \in B(E, F)$ is called an *isometry* if $\|T(x)\|_F = \|X\|_E$ for all $x \in E$, in particular this means that $\|T\| = 1$. However it is important to note that $\|T\| = 1$ does not imply that T is an isometry.

Lemma 2.27. *If $T \in B(E, F)$ is an isometry, T is injective.*

Proof. Suppose $T \in B(E, F)$ is an isometry and let $x, y \in E$ be such that $T(x) = T(y)$. Then $\|x - y\|_E = \|T(x - y)\|_F = \|T(x) - T(y)\|_F = \|0\|_F = 0$. So $x = y$ since $\|x - y\|_E = 0 \iff x = y$. □

2.5. **Dual spaces.**

Definition 2.28. For a normed space E, the space $E^* = B(E, \mathbb{K})$ is called the *dual space* of E. An element $\phi \in E^*$ is called a *bounded linear functional.*

Corollary 2.29. *For any normed space E, E^* is a Banach space.*

Proof. This follows immediately from theorem 2.24. □

Theorem 2.30. *Let $1 < p < \infty$ and q be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then the map $\ell_q \to \ell_p^* : u \mapsto \phi_u$ is an isometric isomorphism, where ϕ_u is defined for $u = (u_n) \in \ell_q$ by*

$$\phi_u(x) = \sum_{n=1}^{\infty} u_n x_n, \text{ where } x = (x_n) \in \ell_p.$$

Proof. Before beginning the proof, it is perhaps worth our time to make a checklist of what we have to show.

(i) The map ϕ_u as it is defined makes sense, is linear and is bounded.

(ii) The mapping $u \mapsto \phi_u$ is an isomorphism. That is it is both injective and surjective.

(iii) The mapping $u \mapsto \phi_u$ is an isometry. Note that this means we only need to check surjectivity of the map by lemma 2.27.

By Hölder's inequality and the generalised triangle inequality, we see that for $u = (u_n) \in \ell_q$ and $x = (x_n) \in \ell_p$,

$$|\phi_u(x)| = |\sum_{k=1}^{\infty} u_k x_k| \leqslant \sum_{k=1}^{\infty} |u_k x_k| \leqslant \left(\sum_{k=1}^{\infty} |u_k|^q\right)^{\frac{1}{q}} \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} = \|u\|_q \|x\|_p.$$

This shows that the series $\sum_{k=1}^{\infty} u_k x_k$ converges absolutely, so it converges and our definition of ϕ_u really does make sense. It should be clear that ϕ_u is linear and bounded, furthermore the above shows us that $\|\phi_u\| \leqslant \|u\|_q$, so the map $u \mapsto \phi_u$ is norm decreasing. It should also be clear that this map from $\ell_q \to \ell_p^*$ is linear.

Now let $\phi \in \ell_p^*$. For each n, let $e_n = (0, 0, \ldots, 0, 1, 0, \ldots)$ where the 1 is the nth term in the sequence. Then, for $x = (x_n) \in \ell_p$,

$$\left\| x - \sum_{k=1}^{n} x_k e_k \right\|_p = \left(\sum_{k=n+1}^{\infty} |x_k|^p\right)^{\frac{1}{P}} \to 0 \text{ as } n \to \infty.$$

Note 2.31. For $x = (x_n) \in \ell_p$, if we define $x^{(N)} = (x_1, x_2, \ldots, x_{N-1}, x_N, 0, 0, \ldots)$, then the above shows that $x^{(N)} \to x$ as $N \to \infty$.

Since $\phi \in \ell_p^*$, ϕ is continuous, so $\phi(x) = \lim_{n\to\infty} \sum_{k=1}^{n} \phi(x_k e_k) = \sum_{k=1}^{\infty} x_k \phi(e_k)$. Define a sequence (u_k) by $u_k = \phi(e_k)$. Fix $N \in \mathbb{N}$ and define $x = (x_k)$ by,

$$x_k = \begin{cases} 0 & \text{if } u_k = 0 \text{ or } k > N \\ \overline{u_k}|u_k|^{q-2} & \text{if } u_k \neq 0 \text{ and } k \leqslant N \end{cases}$$

Now we see that $\sum_{k=1}^{\infty} |x_k|^p = \sum_{k=1}^{N} |u_k|^{p(q-1)} = \sum_{k=1}^{N} |u_k|^q$. By our previous work, we have that $\phi(x) = \sum_{k=1}^{\infty} x_k \phi(e_k)$, so

$$\phi(x) = \sum_{k=1}^{\infty} x_k u_k = \sum_{k=1}^{N} x_k u_k = \sum_{k=1}^{N} \overline{u_k}|u_k|^{q-2} u_k = \sum_{k=1}^{N} |u_k|^{q-2}|u_k|^2 = \sum_{k=1}^{N} |u_k|^q.$$

Hence,

$$\|\phi\| \geqslant \frac{|\phi(x)|}{\|x\|_p} = \frac{\left|\sum_{k=1}^{N} |u_k|^q\right|}{\left(\sum_{k=1}^{N} |u_k|^q\right)^{\frac{1}{p}}} = \frac{\sum_{k=1}^{N} |u_k|^q}{\left(\sum_{k=1}^{N} |u_k|^q\right)^{\frac{1}{p}}} = \left(\sum_{k=1}^{N} |u_k|^q\right)^{1-\frac{1}{p}} = \left(\sum_{k=1}^{N} |u_k|^q\right)^{\frac{1}{q}}.$$

Since this holds for all $N \in \mathbb{N}$, when we take the limit $N \to \infty$ it follows that $u \in \ell_q$, with $\|u\|_q \leqslant \|\phi\|$.

Remember that we have $(u_k) = (\phi(e_k))$ and $\phi(x) = \sum_{k=1}^{\infty} x_k \phi(e_k)$. We have just shown that $(\phi(e_k)) \in \ell_q$, so it follows that $\phi = \phi_u$. Hence, $\|\phi\| = \|\phi_u\| \geqslant \|u\|_q$, so in fact $\|\phi_u\| = \|u\|_q$.

So we have shown that every element of ℓ_p^* arises as ϕ_u for some $u \in \ell_q$, so the map $u \mapsto \phi_u$ is surjective. We have also shown that the map $u \mapsto \phi_u$ is an isometry, so by lemma 2.27 it is injective. So the proposed map is an isometric isomorphism as claimed. $\square$

An immediate corollary of this theorem is the following.

Corollary 2.32. *(Riesz-Frechet self duality lemma)*
ℓ_2 is self dual, that is $\ell_2 = \ell_2^*$.

2.6. **Zorn's Lemma and the Hahn-Banach Theorem.** The Hahn-Banach theorem is a fundamental result in functional analysis, however it features a very non-constructive proof due to it's reliance on Zorn's Lemma. Zorn's Lemma can be shown to be equivalent to the axiom of choice and as such we just accept it as an axiom of mathematics and try not to worry too much about it.

Definition 2.33. A set P is said to be *partially-ordered* if there is a relation $\preceq$ such that for all $x, y, z \in P$ we have

(i) $x \preceq x$

(ii) If $x \preceq y$ and $y \preceq z$, then $x \preceq z$

(iii) If $x \preceq y$ and $y \preceq x$

We say that a poset $(P, \preceq)$ is *totally ordered* if for all $x, y \in P$ either $x \preceq y$ or $y \preceq x$. An element $x \in P$ is said to be *maximal* if for all $y \in P$, $x \preceq y \Rightarrow x = y$.

Theorem 2.34. *(Zorn's lemma)*
Let P be a poset, suppose that every totally ordered subset of P has an upper bound, then the set P has a maximal element.

Now we are ready to state and prove the Hahn-Banach theorem, we do so first for normed $\mathbb{R}$-vector spaces.

Theorem 2.35. *(The Hahn-Banach Theorem)*
Let E be a normed $\mathbb{R}$-vector space, let $F \subseteq E$ be a subspace and let $\phi \in F^*$ be a bounded linear functional. Then there exists $\psi \in E^*$ with $\|\psi\| \leqslant \|\phi\|$ and $\psi(x) = \phi(x)$ for all $x \in F$.

Proof. We will show that we can extend the functional ϕ one dimension at a time, then we apply Zorn's Lemma to get the full result.

An extension of ϕ is a bounded linear map $\phi_G : G \to \mathbb{R}$ where $F \subset G \subset E$, $\phi_G(x) = \phi(x)$ for $x \in F$ and $\|\phi_G\| \leqslant \|\phi\|$.

Choose a point $x_0 \in E \backslash F$ and let G_1 be the linear span of F and x_0. An extension of ϕ to G_1 must be of the form
$$\phi_1(\tilde{x} + ax_0) = \phi_1(\tilde{x}) + a\alpha \qquad (\tilde{x} \in F,\ a \in \mathbb{R})$$
for some $\alpha \in \mathbb{R}$. It should be clear that ϕ_1 is linear, and that $\phi_1(k) = \phi(k)$ if $k \in F$. All that remains for us to check is that $\|\phi_1\| \leqslant \|\phi\|$. Notice that the extension ϕ_1 is defined only by it's value on x_0, i.e. by the value of α. We will show that we can choose α such that the inequality of operator norms is satisfied.

To ensure $\|\phi_1\| \leqslant \|\phi\|$ we need, for all $x = \tilde{x} + ax_0 \in G_1$,
$$|\phi(\tilde{x}) + a\alpha| \leqslant \|\phi\| \|\tilde{x} + ax_0\|.$$

To simplify things, let $\tilde{x} = -ay$ where $y \in F$. So
$$|\phi(-ay) + a\alpha| = |a(\alpha - \phi(y))| \leqslant \|\phi\| \| - ay + ax_0\| = \|\phi\| \|a(x_0 - y)\|.$$
Dividing through by $a \in \mathbb{R}$ gives us $|\alpha - \phi(y)| \leqslant \|\phi\| \|x_0 - y\|$, and if we expand the modulus on the left hand side, we get
$$\phi(y) - \|\phi\| \|x_0 - y\| \leqslant \alpha \leqslant \phi(y) + \|\phi\| \|x_0 - y\|.$$

For distinct $y_1, y_2 \in F$, we have $\phi(y_1) - \phi(y_2) = \phi(y_1 - y_2) \leqslant \|\phi\| \|y_1 - y_2\|$, and by the triangle inequality we have
$$\phi(y_1 - y_2) \leqslant \|\phi\| (\|x_0 - y_2\| + \|x_0 - y_1\|).$$
Thus
$$\phi(y_1) - \|\phi\| \|x_0 - y_1\| \leqslant \phi(y_2) + \|\phi\| \|x_0 - y_2\|.$$

But remember that our choices of y_1 and y_2 were arbitrary, so

$$\sup_{y \in G}(\phi(y) - \|\phi\|\|x_0 - y\|) \leqslant \inf_{y \in G}(\phi(y) + \|\phi\|\|x_0 - y\|).$$

Hence we can pick α between the inf and the sup such that our inequality is satisfied. So $\|\phi_1\| \leqslant \|\phi\|$ and the first step of the proof is complete.

Now let P be the set of pairs (G, ψ) such that $G \subseteq E$ is a subspace containing F and $\psi : F \to \mathbb{R}$ is an extension of ϕ. We define a partial ordering on P by $(G, \psi) \preceq (G', \psi')$ if and only if $G \subseteq G'$ and ψ' extends ψ. Let $T = \{(G_i, \psi_i); i \in I\}$ be a totally ordered subset of P, then we can form the vector space $G = \cup_{i \in I} G_i$ and a linear functional $\psi : G \to \mathbb{R}$ by $\psi(x) = \psi_i(x)$ whenever $x \in G_i$. Then the pair (G, ψ) is an upper bound for T, so by Zorn's Lemma P has a maximal element, say (M, F).

Suppose that $M \neq E$, then by our previous work, we have a pair (M', F') such that $(M, F) \preceq (M', F')$ and $(M, F) \neq (M', F')$. This contradicts maximality of (M, F) and so we have an extension $\psi : E \to \mathbb{R}$ where $\psi(x) = \phi(x)$ if $x \in F$ and $\|\psi\| \leqslant \|\phi\|$. □

We may extend the Hahn-Banach theorem to normed $\mathbb{C}$-vector spaces with the next result.

Corollary 2.36. *Let E be a normed $\mathbb{C}$-vector space, let $F \subseteq E$ be a subspace and let $\phi \in F^*$ be a bounded linear functional. Then there exists $\psi \in E^*$ with $\|\psi\| \leqslant \|\phi\|$ and $\psi(x) = \phi(x)$ for all $x \in F$.*

Proof. Let $\phi \in F^*$, then since ϕ is complex valued, for $x \in F$ we may write $\phi(u) = f(x) + ig(x)$ where $f, g : F \to \mathbb{R}$. Since ϕ is $\mathbb{C}$-linear, for $x \in F$, we have

$$f(ix) + ig(ix) = \phi(ix) = i\phi(x) = if(x) - g(x),$$

and we see that $f(ix) = -g(x)$. So for $x \in F$ we have $\phi(x) = f(x) - if(ix)$. In other words, ϕ is completely determined by its real part $f(x) = \mathrm{Re}(\phi(x))$. Viewing F and E as $\mathbb{R}$-vector spaces, we see that f is $\mathbb{R}$-linear, since for $x, y \in F$ and $\mu, \lambda \in \mathbb{R}$,

$$f(\mu x + \lambda y) = \mathrm{Re}(\phi(\mu x + \lambda y)) = \mathrm{Re}(\mu\phi(x) + \lambda\phi(y)) = \mu f(x) + \lambda f(y).$$

Furthermore, for any $x \in F$ we have

$$\|\phi\|\|x\|_F \geqslant |\phi(x)| = \sqrt{|\mathrm{Re}(\phi(x))|^2 + |\mathrm{Im}(\phi(x))|^2} \geqslant |\mathrm{Re}(x)| = |f(x)|,$$

which shows that f is bounded, and $\|f\| \leqslant \|\phi\|$.

By the Hahn-Banach theorem for normed $\mathbb{R}$-vector spaces, we may extend f to a bounded linear functional $\tilde{f} : E \to \mathbb{R}$ such that $\|\tilde{f}\| \leqslant \|f\|$. Now define $\psi : E \to \mathbb{C}$ by

$$\psi(x) = \tilde{f}(x) - i\tilde{f}(ix), \text{ for } x \in E.$$

We can easily check that this is an extension of ϕ, since for $x \in F$, we have

$$\psi(x) = \tilde{f}(x) - i\tilde{f}(ix) = f(x) - if(ix) = \phi(x).$$

To see that ψ is $\mathbb{C}$-linear, first note that ψ is $\mathbb{R}$-linear, since for $x, y \in E$ and $\mu, \lambda \in \mathbb{R}$, we have

$$\psi(\mu x + \lambda y) = \tilde{f}(\mu x + \lambda y) - i\tilde{f}(i\mu x + i\lambda y) = \mu\tilde{f}(x) + \lambda\tilde{f}(y) - i\mu\tilde{f}(ix) - i\lambda\tilde{f}(iy) = \mu\psi(x) + \lambda\psi(y).$$

Now observe that for any $x \in E$,

$$\psi(ix) = \tilde{f}(ix) - i\tilde{f}(-x) = \tilde{f}(ix) + \tilde{f}(x) = i(\tilde{f}(x) - i\tilde{f}(ix)) = i\psi(x).$$

Combining we see that $\psi(\mu x + \lambda y) = \mu\psi(x) + \lambda\psi(y)$ for $x, y \in E$ and $\mu, \lambda \in \mathbb{C}$.

Finally, we must show that $\|\psi\| \leqslant \|\phi\|$. For $u \in E$, $\psi(u)$ is a complex number, so we may write $\psi(u) = re^{i\theta} = |\psi(u)|e^{i\theta}$. Multiplying by $e^{-i\theta}$, we see that $|\psi(u)| = \psi(u)e^{-i\theta} = \psi(ue^{-i\theta})$. Since $|\psi(u)|$ is real, it follows that $\psi(ue^{-i\theta})$ is real, so $\psi(ue^{-i\theta}) = \tilde{f}(ue^{-i\theta})$.

So,

$$|\psi(u)| = |e^{-i\theta}||\psi(u)| = |\psi(ue^{-i\theta})| = |\tilde{f}(ue^{-i\theta})| \leqslant \|\phi\|\|ue^{-i\theta}\|_E = \|\phi\|\|u\|_E.$$

And for $u \in E\backslash\{0\}$ we have $\frac{\psi(u)}{\|u\|_E} \leqslant \|\phi\|$, so $\|\psi\| \leqslant \|\phi\|$. □

Note 2.37. In the statement of the Hahn-Banach theorem, the inequality $\|\psi\| \leqslant \|\phi\|$ is actually an equality, because

$$\left\{\frac{|\phi(x)|}{\|x\|}; x \in F, x \neq 0\right\} \subseteq \left\{\frac{|\psi(x)|}{\|x\|}; x \in E, x \neq 0\right\} \text{ since F is a subset of E.}$$

And after taking supremums, it follows that $\|\phi\| \leqslant \|\psi\|$ so $\|\psi\| = \|\phi\|$.

We now present some corollaries of the Hahn-Banach theorem.

Corollary 2.38. *Let E be a normed space and let $x \in E$. Then there exists $\phi \in E^*$ such that $\|\phi\| = 1$ and $\phi(x) = \|x\|$.*

Proof. Let $F = \mathrm{Span}_{\mathbb{K}}\{x\}$ and define $\psi \in F^*$ by $\psi(\alpha x) = |\alpha|\|x\|_E$ where $\alpha \in \mathbb{K}$. It should be clear that ψ really is linear and bounded, and that $\psi(x) = \|x\|_E$. Also

$$\|\psi\| = \sup\left\{\frac{|\psi(\alpha x)|}{\|\alpha x\|_E}; \alpha \in \mathbb{K}\backslash\{0\}\right\} = \sup\left\{\frac{|\alpha|\|x\|_E}{\|\alpha x\|_E}; \alpha \in \mathbb{K}\backslash\{0\}\right\} = 1.$$

By the Hahn-Banach theorem, we may extend ψ to $\phi \in E^*$ such that $\|\phi\| = 1$ and $\phi(x) = \|x\|_E$, and we are done. □

Corollary 2.39. *Let E be a normed space and let $F \subseteq E$ be a subspace. For $x \in E$, the following are equivalent.*

(i) $x \in \overline{F}$, the closure of F

(ii) For each $\phi \in E^*$, with $\phi(y) = 0$ for $y \in F$, we have $\phi(x) = 0$

Proof. $(i) \Rightarrow (ii)$ If $x \in \overline{F}$, then we can find a sequence (y_n) in F such that $\lim_{n\to\infty} y_n = x$. Then, for each $\phi \in E^*$ with $\phi(y) = 0$ for all $y \in F$, we have

$$\phi(x) = \phi\left(\lim_{n\to\infty} y_n\right) = \lim_{n\to\infty} \phi(y_n) = 0.$$

$(ii) \Rightarrow (i)$ For this part, we shall prove the contrapositive. If $x \notin \overline{F}$ then write G for the linear span of F and x, and define $\psi : G \to \mathbb{K}$ by

$$\psi(y + tx) = t, \text{ for } y \in F,\ t \in \mathbb{K}.$$

We check that ψ is well defined and linear.

Furthermore, ψ is bounded, which we prove by contradiction. Suppose ψ is not bounded, so there exists a sequence $(y_n + t_n x)$ with $\|y_n + t_n x\| \leqslant 1$ for each n, and $|\psi(y_n + t_n x)| = |t_n| \to \infty$ as $n \to \infty$. Then

$$\left\|\frac{y_n}{t_n} + x\right\| \leqslant \frac{1}{|t_n|} \to \infty \text{ as } \to \infty,$$

so the sequence $(-\frac{y_n}{t_n})$ in F converges to x, so $x \in \overline{F}$ which is a contradiction. Thus ψ is bounded and by the Hahn-Banach theorem, we may find $\phi \in E^*$ extending ψ (and such that $\|\phi\| \leqslant \|\psi\|$).

Now for $y \in F$ we have $\phi(y) = \psi(y) = 0$ whilst $\phi(x) = \psi(x) = 1$, so (ii) doesn't hold.

□

Definition 2.40. Let E be a normed space, we define $E^{**} = (E^*)^* = \{T : E^* \to \mathbb{K};\ \text{T is linear and bounded}\}$ to be the *bidual* of E.

We define a map $J : E \to E^{**}$ as follows.

For $x \in E$, we need $J(x) \in E^{**}$, so $J(x)$ is a map $J(x) : E^* \to \mathbb{K}$. We define this to be the map which maps $\phi \mapsto \phi(x)$ for $\phi \in E^*$. We write this as

$$J(x)(\phi) = \phi(x) \text{ for } x \in E,\, \phi \in E^*.$$

Proposition 2.41. *The map J described above is an isometry.*

Proof. Fix some point $x \in E$, then

$$\|J(x)\| = \sup\{|J(x)(\phi)|;\, \phi \in E^*,\, \|\phi\| \leqslant 1\} = \sup\{|\phi(x)|;\, \phi \in E^*,\, \|\phi\| \leqslant 1\}.$$

For $\phi \in E^*$ with $\|\phi\| \leqslant 1$, we have $|\phi(y)| \leqslant \|y\|_E$ for all $y \in E$. By corollary 2.38, there is $\psi \in E^*$ with $\|\psi\| = 1$ and $\psi(x) = \|x\|_E$. Therefore, for our $x \in E$,

$$\|J(x)\| = \|x\|_E.$$

□

Definition 2.42. When the map J described above is surjective (so is an isometric isomophism), we say that the normed space E is *reflexive*.