321 2008–09

Chapter 3: Dual Spaces and the Hahn-Banach Theorem

Definition 3.1. If E is a normed space, the *dual space* of E is

$$E^* = \mathcal{B}(E, \mathbb{K}) = \{T \colon E \to \mathbb{K} : T \text{ continuous and linear}\}.$$

Elements of E^* are called (continuous) *linear functionals* on E.

The notation E' is sometimes used for E^*.

The letter $\mathcal{L}$ (for 'linear') is used by some people rather than $\mathcal{B}$ (for 'bounded', but really 'bounded linear').

Proposition 3.2. *If E is a normed space, then E^* is a Banach space.*

Proof. This follows from the following more general result (by taking $F = \mathbb{K}$). □

The following result should also be familiar from Exercises 4, question 3.

Theorem 3.3. *Let E be a normed space and F a Banach space. Let $\mathcal{B}(E, F)$ denote the space of all bounded linear operators $T \colon E \to F$. We can make $\mathcal{B}(E, F)$ a vector space by defining $T + S$ and λT ($T, S \in \mathcal{B}(E, F)$, $\lambda \in \mathbb{K}$) as follows: $(T + S)(x) = T(x) + S(x)$, $(\lambda T)(x) = \lambda(T(x))$ for $x \in E$.*

Then the operator norm is a norm on $\mathcal{B}(E, F)$ and makes it a Banach space.

Proof. The facts that $\mathcal{B}(E, F)$ is a vector space and that the operator 'norm' is actually a norm on the space are rather straightforward to check. The main point is to show that the space is complete.

Take a Cauchy sequence $\{T_n\}_{n=1}^\infty$ in $(\mathcal{B}(E, F), \|\cdot\|_{op})$. For any fixed $x \in E$ $\{T_n(x)\}_{n=1}^\infty$ is a Cauchy sequence in F because

$$\|T_n(x) - T_m(x)\|_F \leq \|T_n - T_m\|_{op} \|x\|_E$$

is small if n and m are both large. Since F is complete, it follows that

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

exists in F (for each $x \in E$). This allows us to define a map $T\colon E \to F$ by

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

We will be finished if we show that $T \in \mathcal{B}(E,F)$ and $\lim_{n\to\infty} \|T_n - T\|_{op} = 0$ (so that $T_n \to T$ in the norm of $\mathcal{B}(E,F)$).

By the Cauchy condition, we know that we can find N so that $\|T_n - T_m\|_{op} < 1$ for all $n, m \geq N$. Now take $x \in E$, $\|x\|_E \leq 1$. Then for $n, m \geq N$ we have

$$\|T_n(x) - T_m(x)\|_F = \|(T_n - T_m)(x)\|_F \leq \|T_n - T_m\|_{op} < 1.$$

Fix $n = N$ and let $m \to \infty$ and use continuity of the norm on F to conclude

$$\|T_N(x) - T(x)\|_F \leq 1.$$

This is true for all $x \in E$ of norm $\|x\|_E \leq 1$ and we can use that to conclude that

$$\begin{aligned} \sup_{\substack{x\in E\\ \|x\|\leq 1}} \|T(x)\|_F &\leq \sup_{\substack{x\in E\\ \|x\|\leq 1}} \|T_N(x)\|_F + \sup_{\substack{x\in E\\ \|x\|\leq 1}} \|T(x) - T_N(x)\|_F \\ &\leq \|T_N\|_{op} + 1. \end{aligned}$$

We see now that T is bounded. So $T \in \mathcal{B}(E,F)$.

We now repeat the last few steps with an arbitrary $\varepsilon > 0$ where we had 1 before. By the Cauchy condition we can find N so that $\|T_n - T_m\|_{op} < \varepsilon$ for all $n, m \geq N$. Now take $x \in E$, $\|x\|_E \leq 1$. As before we get

$$\|T_n(x) - T_m(x)\|_F = \|(T_n - T_m)(x)\|_F \leq \|T_n - T_m\|_{op} < \varepsilon$$

as long as $n, m \geq N$, $\|x\| \leq 1$. Fix any $n \geq N$ for the moment and let $m \to \infty$ to get

$$\|T_n(x) - T(x)\|_F \leq \varepsilon.$$

We have this for all x of norm at most 1 and all $n \geq N$. So

$$\|T_n - T\| = \sup_{x\in E, \|x\|\leq 1} \|T(x) - T_n(x)\|_F \leq \varepsilon$$

as long as $n \geq N$. This shows that $T_n \to T$ in $\mathcal{B}(E,F)$.

Thus every Cauchy sequence in the space converges in the space and so $\mathcal{B}(E,F)$ is a Banach space. $\square$

It is worth pointing out that the strategy of the above proof is a bit similar to the proof that (for a topological space X) $(BC(X), \|\cdot\|_\infty)$ is complete (which we did in 1.7.2 (iii)).

Examples 3.4. (i) If $E = \ell^p$ and $1 \leq p < \infty$, then its dual space

$$(\ell^p)^* = \ell^q \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right).$$

More exactly, ℓ^q is isometrically isomorphic to $(\ell^p)^*$ by the following isometry

$$T \colon \ell^q \to (\ell^p)^*,$$

where for $(b_n)_n \in \ell^q$

$$\begin{aligned} T((b_n)_n) \colon \ell^p &\to \mathbb{K} \\ \colon (a_n)_n &\mapsto \sum_{n=1}^{\infty} a_n b_n \end{aligned}$$

Proof. The first thing to do is to show that the map T makes sense. If we fix $b = (b_n)_n \in \ell^q$ then Hölders inequality tells us that

$$\left|\sum_{n=1}^{\infty} a_n b_n\right| \leq \|a\|_p \|b\|_q$$

for each $a = (a_n)_n \in \ell^p$. This implies that the series for $(T(b))(a)$ is always convergent and so $T(b) \colon \ell^p \to \mathbb{K}$ is a sensibly defined function. It is quite easy to see that it is a linear map.

The inequality we got from Hölders inequality $|T(b)(a)| \leq \|a\|_p \|b\|_q$ now implies that $T(b)$ is a bounded linear map and that $T(b)$ has (operator) norm at most $\|b\|_q$. So $T(b)$ is in the dual space of ℓ^p and $\|T(b)\|_{(\ell^p)^*} \leq \|b\|_q$.

Our next aim is to show that equality holds in this inequality. So we aim to show $\|T(b)\|_{(\ell^p)^*} \geq \|b\|_q$. If $b = 0$ then we certainly have that. So we assume $b \neq 0$ (which means there is some n with $b_n \neq 0$). The proof is a bit different for $p = 1$, $q = \infty$. So we consider the case $1 < p < \infty$ first.

If we take $a_n = |b_n|^q (b_n)^{-1}$ (interpreted as 0 if $b_n = 0$), then $a = (a_n)_n \in \ell^p$ and

$$\|a\|_p = \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p} = \left(\sum_{n=1}^{\infty} |b_n|^{qp-p}\right)^{1/p} = \left(\sum_{n=1}^{\infty} |b_n|^q\right)^{1/p} = \|b\|_q^{q/p}$$

because $(q-1)p = q$. Now $T(b)(a) = \sum_{n=1}^\infty a_n b_n = \sum_{n=1}^\infty |b_n|^q = \|b\|_q^q$. Thus (if $b \neq 0$)

$$\|T(b)\|_{(\ell^p)^*} \geq \frac{|T(b)(a)|}{\|a\|_p} = \frac{\|b\|_q^q}{\|b\|_q^{q/p}} = \|b\|_q^{q-q/p} = \|b\|_q.$$

From this we see that $\|T(b)\|_{(\ell^p)^*} = \|b\|_q$ (in the case $p > 1$).

If $p = 1$, take any n_0 with $b_{n_0} \neq 0$ and define $a = (a_n)_{n=1}^\infty$ by $a_n = |b_n|/b_n$ for $n = n_0$ and $a_n = 0$ otherwise. So $\|a\|_1 = \sum_n |a_n| = |a_{n_0}| = 1$ and

$$|T(b)(a)| = \left|\sum_{n=1}^\infty a_n b_n\right| = |a_{n_0} b_{n_0}| = |b_{n_0}|.$$

So

$$\|T(b)\|_{(\ell^p)^*} \geq \frac{|T(b)(a)|}{\|a\|_1} = \frac{|b_{n_0}|}{1} = |b_{n_0}|,$$

and that is true for and all choices of n_0 with $b_{n_0} \neq 0$. Thus

$$\|T(b)\|_{(\ell^1)^*} \geq \sup_{n_0, b_{n_0} \neq 0} |b_{n_0}| = \sup_n |b_n| = \|b\|_\infty.$$

This deals with the case $p = 1$, $q = \infty$.

So T is a norm-preserving map from ℓ^q to a subset of $(\ell^p)^*$. In fact T is also linear (and so is a linear isometry onto a vector subspace of $(\ell^p)^*$) because it is quite straightforward to check that for $b, \tilde{b} \in \ell^q$ and $\lambda \in \mathbb{K}$, we have $T(\lambda b + \tilde{b}) = \lambda T(b) + T(\tilde{b})$. (Apply both sides to an arbitrary $a \in \ell^p$ to see that this is true.)

To show it is surjective, take $\alpha \in (\ell^p)^*$. Then define $b = (b_n)_n$ by $b_n = \alpha(e_n)$ where by e_n we mean the sequence $(0, 0, \ldots, 0, 1, 0, \ldots)$ of all zeros except for a 1 in position n. We now show that $b \in \ell^q$ and that $T(b) = \alpha$.

Notice first that for any finitely nonzero sequence $a = (a_1, a_2, \ldots, a_n, 0, 0, \ldots)$ of scalars, $a \in \ell^p$ and a is a finite linear combination

$$a = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$$

of the e_j's. We can then apply α to both sides of this equation and use linearity of α to see that

$$\alpha(a) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

We will deal now with the case $1 < p < \infty$ only (as the case $p = 1$ would be a little different). Take $a_j = |b_j|^q (b_j)^{-1}$ for $j = 1, 2, \ldots, n$ (and $a_j = 0$ for $j > n$). We can calculate as we did little earlier that

$$\|a\|_p = \left(\sum_{j=1}^{n} |b_j|^q \right)^{1/p}$$

and

$$\alpha(a) = \sum_{j=1}^{n} a_j b_j = \sum_{j=1}^{n} |b_j|^q$$

so that

$$\|\alpha\|_{(\ell^p)^*} \geq \frac{|\alpha(a)|}{\|a\|_p} = \frac{\sum_{j=1}^{n} |b_j|^q}{\left(\sum_{j=1}^{n} |b_j|^q\right)^{\frac{1}{p}}} = \left(\sum_{j=1}^{n} |b_j|^q \right)^{1-\frac{1}{p}} = \left(\sum_{j=1}^{n} |b_j|^q \right)^{1/q}$$

(as long as there is some nonzero b_j in the range $1 \leq j \leq n$).

If we now let $n \to \infty$ in this inequality, we find that $b \in \ell^q$ (and that $\|b\|_q \leq \|\alpha\|_{(\ell^p)^*}$). This part of the argument was for $p > 1$. When $p = 1$, we can say $\|e_n\|_1 = 1$ and so

$$|b_n| = |\alpha(e_n)| \leq \|\alpha\|_{(\ell^1)^*} \|e_n\|_1 = \|\alpha\|_{(\ell^1)^*}.$$

Thus $\|b\|_q = \|b\|_\infty = \sup_n |b_n| \leq \|\alpha\|_{(\ell^1)^*}$ and $b \in \ell^q$ also holds in the case $p = 1$.

So now we can talk about $T(b) \in (\ell^p)^*$ and it is easy to see that $T(b)(e_n) = b_n = \alpha(e_n)$. It follows by linearity that for any finitely nonzero sequence $a = \sum_{j=1}^{n} a_j e_j \in \ell^p$ we have $T(b)(a) = \alpha(a)$.

Finally, these finitely nonzero sequences are dense in ℓ^p because for any $x = (x_n)_n \in \ell^p$ it is rather clear that

$$\|x - (x_1, x_2, \ldots, x_n, 0, 0, \ldots)\|_p = \left(\sum_{j=n+1}^{\infty} |x_j|^p \right)^{1/p} \to 0$$

as $n \to \infty$. Since $T(b)$ and α are continuous on ℓ^p and equal on a dense set, it must be that $T(b) = \alpha$. Thus T is surjective. $\square$

(ii) $(c_0)^* = \ell^1$ — similar proof to the proof of (i).

(iii) $L^p[0,1]^* = L^q[0,1]$ and $L^p(\mathbb{R})^* = L^q(\mathbb{R})$ for $1/p + 1/q = 1$, $1 \leq p < \infty$.

The precise meaning of "equality" here is that $g \in L^q$ corresponds to the linear functional

$$\begin{array}{rcl} L^p & \to & \mathbb{K} \\ f & \mapsto & \displaystyle\int f(x)g(x)\,dx. \end{array}$$

The proofs of the equalities here are somewhat similar to the proof of (i), but require some measure theory — proofs omitted).

Theorem 3.5 (Hahn-Banach Theorem for $\mathbb{R}$)**.** *Let E be a vector space over $\mathbb{R}$ and M a vector subspace. Suppose $p\colon E \to [0,\infty)$ is a seminorm on E and suppose*

$$\alpha\colon M \to \mathbb{R}$$

is a linear functional satisfying

$$|\alpha(x)| \leq p(x) \qquad \textit{for all } x \in M.$$

Then there exists an extension $\beta\colon E \to \mathbb{R}$ of α which is linear and satisfies

(i) $\beta(x) = \alpha(x)$ for all $x \in M$ (i.e. β extends α)

(ii) $|\beta(x)| \leq p(x)$ for all $x \in E$.

[The most frequently used case is where E is a normed space and α is a continuous linear functional on a vector subspace M — so that $\alpha \in M^*$ and $|\alpha(x)| \leq \|\alpha\|\,\|x\|$ ($x \in M$). To apply the theorem, we take $p(x) = \|\alpha\|\,\|x\|$.

The conclusion of the theorem is that there is a linear extension $\beta\colon E \to \mathbb{R}$ of α that satisfies

$$|\beta(x)| \leq p(x) = \|\alpha\|\,\|x\| \qquad (x \in E),$$

or, in other terms, an extension $\beta \in E^*$ with norm $\|\beta\| \leq \|\alpha\|$.

In fact the extension β cannot have a smaller norm than α and so $\|\beta\| = \|\alpha\|$.]

Proof. The proof is rather nice, being a combination of an application of Zorn's Lemma and a clever argument.

With Zorn's lemma, we show we can find a maximal extension of α satisfying the inequality. What we do is consider all possible linear extensions $\gamma\colon M_\gamma \to \mathbb{R}$

of α to linear subspaces M_γ of E containing M. To be more precise we demand that $M \subseteq M_\gamma \subseteq E$, that γ is linear, that $\gamma(x) = \alpha(x)$ for all $x \in M$ and that $|\gamma(x)| \leq p(x)$ for all $x \in M_\gamma$. We partially order this collection of extensions by saying that $\gamma \colon M_\gamma \to \mathbb{R}$ is less that $\delta \colon M_\delta \to \mathbb{R}$ if $M_\gamma \subset M_\delta$ and $\delta(x) = \gamma(x)$ for all $x \in M_\gamma$ (that is δ extends γ). We write $\gamma \leq \delta$ when this happens.

Let S denote the set of all these linear extensions $\gamma \colon M_\gamma \to \mathbb{R}$ and it is not difficult to see that $(S, \leq)$ is partially ordered. Note that $S \neq$ as $\alpha \colon M \to \mathbb{R}$ belongs to S. Take a chain C in S. If $C = \emptyset$, then α is an upper bound for C. If $C \neq \emptyset$, then we can take $M_\delta = \bigcup_{\gamma \in C} M_\gamma$ and define $\delta \colon M_\gamma \to \mathbb{R}$ as follows. If $x \in M_\delta$, then $x \in M_\gamma$ for some $\gamma \in C$. We define $\delta(x) = \gamma(x)$. Before we can do that we need to show that this makes δ a well-defined function. For $x \in M_\delta$ we could well have several $\gamma \in C$ where $x \in M_\gamma$ and then we might have ambiguous definitions of $\delta(x)$. If $x \in M_{\gamma_1}$ and $x \in M_{\gamma_2}$ for $\gamma_1, \gamma_2 \in C$, then we must gave $\gamma_1 \leq \gamma_2$ or else $\gamma_2 \leq \gamma_1$ (since C is a chain). We consider only the case $\gamma_1 \leq \gamma_2$ because the other case is similar (or even the same if we re-number γ_1 and γ_2). Then $M_{\gamma_1} \subset M_{\gamma_2}$ and $\gamma_1(x) = \gamma_2(x)$. So $\delta(x)$ is well-defined.

To show that $\delta \in S$, we first need to know that M_δ is a subspace, then that δ is linear. If $x, y \in M_\delta$ and $\lambda \in \mathbb{R}$, then there are $\gamma_1, \gamma_2 \in C$ with $x \in M_{\gamma_1}$, $y \in M_{\gamma_2}$. As C is a chain, $\gamma_1 \leq \gamma_2$ or $\gamma_2 \leq \gamma_1$. If $\gamma_1 \leq \gamma_2$ let $\gamma = \gamma_2$, but if $\gamma_2 \leq \gamma_1$ let $\gamma = \gamma_1$. Then $x, y \in M_\gamma$. As we know M_γ is linear, we have $\lambda x + y \in M_\gamma \subset M_\delta$. So M_δ a subspace. Moreover we can say that $\delta(\lambda x + y) = \gamma(\lambda x + y) = \lambda\gamma(x) + \gamma(y) = \lambda\delta(x) + \delta(y)$ (because γ is linear). So δ is linear.

To complete the verification that $\delta \in S$, we take $x \in M$. as C is not empty, there is $\gamma \in C$ and then we have $\alpha(x) = \gamma(x) = \delta(x)$. So $\delta \colon M_\delta \to \mathbb{R}$ is a linear extension of α. Finally $|\delta(x)| \leq p(x)$ $(\forall x \in M_\delta)$ because

$$x \in M_\delta \Rightarrow \exists \gamma \in C \text{ with } x \in M_\gamma \Rightarrow |\delta(x)| = |\gamma(x)| \leq p(x)$$

and so $\delta \in S$.

Now δ is an upper bound for C since $\gamma \in C$ implies $M_\gamma \subset M_\delta$ and for $x \in M_\gamma$ we do have $\gamma(x) = \delta(x)$.

By Zorn's lemma, we conclude that there must be a maximal element $\gamma \colon M_\gamma \to \mathbb{R}$ in S.

If $M_\gamma = E$ for this maximal γ then we take $\beta = \gamma$. If not, there is a point $x_0 \in E \setminus M_\gamma$ and what we do is show that we can extend γ to the space spanned by x_0 and γ while keeping the extension linear and still obeying the inequality. In other words, we can get a contradiction to the maximality of γ if $M_\gamma \neq E$.

The linear span of M_γ and x_0 is

$$N = \{x + \lambda x_0 : x \in M_\gamma, \lambda \in \mathbb{R}\}.$$

Also each element of $y \in N$ can be uniquely expressed as $y = x + \lambda x_0$ (with $x \in M_\gamma$, $\lambda \in \mathbb{R}$) because if

$$y = x + \lambda x_0 = x' + \lambda' x_0 \quad (x, x' \in M_\gamma, \lambda, \lambda' \in \mathbb{R})$$

then $x - x' = (\lambda' - \lambda)x_0$. So $\lambda' - \lambda = 0$ must hold. Otherwise $x_0 = (1/(\lambda' - \lambda))(x - x') \in M_\gamma$, which is false. Thus $\lambda = \lambda'$ and $x = x'$.

We define $\delta\colon N \to \mathbb{R}$ by $\delta(x + \lambda x_0) = \gamma(x) + \lambda c$ where we still have to say how to choose $c \in \mathbb{R}$. No matter how we choose c, δ will be linear on N and will extend γ, but the issue is to choose c so that we have

$$|\gamma(x) + \lambda c| \leq p(x + \lambda x_0) \quad (x \in M_\gamma, \lambda \in \mathbb{R}).$$

In fact it is enough to choose c so that this works for $\lambda = 1$. If $\lambda = 0$, we already know $|\gamma(x)| \leq p(x)$, while if $\lambda \neq 0$ we can write

$$|\gamma(x) + \lambda c| = |\lambda| \left| \gamma\left(\frac{1}{\lambda} x\right) + x_0 \right|, \quad p(x + \lambda x_0) = |\lambda| p\left(\frac{1}{\lambda} x + x_0\right),$$

with $(1/\lambda)x \in M_\gamma$. Thus we want to choose c so that $|\gamma(x) + c| \leq p(x + x_0)$ holds for $x \in M_\gamma$, which is the same as

$$-p(x + x_0) \leq \gamma(x) + c \leq p(x + x_0)$$

or

$$c \leq p(x + x_0) - \gamma(x) \text{ and } -p(x + x_0) - \gamma(x) \leq c \quad (\forall x \in M_\gamma) \tag{1}$$

Notice that for $x, x_1 \in M_\gamma$ we have

$$\begin{aligned} \gamma(x) - \gamma(x_1) &= \gamma(x - x_1) \\ &\leq p(x - x_1) = p((x + x_0) + (-x_1 - x_0)) \\ &\leq p(x + x_0) + p(-x_1 - x_0) \\ &= p(x + x_0) + p(x_1 + x_0) \end{aligned}$$

Thus

$$-p(x_1 + x_0) - \gamma(x_1) \leq p(x + x_0) - \gamma(x) \quad (\forall x, x_1 \in M_\gamma).$$

For fixed $x \in M_\gamma$, we then have

$$\sup_{x_1 \in M_\gamma} -p(x_1 + x_0) - \gamma(x_1) \leq p(x + x_0) - \gamma(x).$$

Hence

$$\sup_{x_1 \in M_\gamma} -p(x_1 + x_0) - \gamma(x_1) \leq \inf_{x \in M_\gamma} p(x + x_0) - \gamma(x).$$

We choose c so that

$$\sup_{x_1 \in M_\gamma} -p(x_1 + x_0) - \gamma(x_1) \leq c \leq \inf_{x \in M_\gamma} p(x + x_0) - \gamma(x).$$

Then c satisfies the requirement (1).

With this choice of c, we have $\delta\colon N \to \mathbb{R}$ in S, $\gamma \leq \delta$ but $\delta \neq \gamma$ (because $N \neq M_\gamma$). This is a contradiction to the maximality of γ. So we must have $M_\gamma = E$. □

Theorem 3.6 (Hahn-Banach Theorem for $\mathbb{C}$)**.** *Let E be a vector space over $\mathbb{C}$ and M a vector subspace. Suppose $p\colon E \to [0, \infty)$ is a seminorm on E and suppose*

$$\alpha\colon M \to \mathbb{C}$$

is a linear functional satisfying

$$|\alpha(x)| \leq p(x) \qquad \textit{for all } x \in M.$$

Then there exists an extension $\beta\colon E \to \mathbb{C}$ of α which is linear and satisfies

(i) $\beta(x) = \alpha(x)$ for all $x \in M$

(ii) $|\beta(x)| \leq p(x)$ for all $x \in E$.

Proof. Observe that the statement is the same as for the case of $\mathbb{R}$ (and it is usually applied in exactly the same way). This is one of the few theorems where separate proofs are needed for the two cases. In fact the complex case follows from the real case.

We have $\alpha\colon M \to \mathbb{C}$. Write

$$\alpha(x) = \alpha_1(x) + i\alpha_2(x),$$

where $\alpha_1\colon M \to \mathbb{R}$, $\alpha_2\colon M \to \mathbb{R}$ are the real and imaginary parts of α.

Note that α_1 is $\mathbb{R}$-linear and $|\alpha_1(x)| \leq |\alpha(x)| \leq p(x)$. So, by Theorem 3.5, we can extend α_1 to get an $\mathbb{R}$-linear $\beta_1: E \to \mathbb{R}$ satisfying $|\beta_1(x)| \leq p(x)$ for all $x \in E$.

The key observation is that there is a relationship between the real and imaginary parts of a $\mathbb{C}$-linear functional, found as follows:

$$\begin{aligned} \alpha(ix) &= i\alpha(x) \\ \alpha_1(ix) + i\alpha_2(ix) &= i(\alpha_1(x) + i\alpha_2(x)) \\ &= i\alpha_1(x) - \alpha_2(x) \end{aligned}$$

We deduce that $\alpha_2(x) = -\alpha_1(ix)$ for all $x \in M$.

Since a similar analysis could be applied to a $\mathbb{C}$-linear functional on E (such as the extension we are seeking), this leads us to define

$$\begin{aligned} \beta: E &\to \mathbb{C} \\ \beta(x) &= \beta_1(x) - i\beta_1(ix) \end{aligned}$$

Then β is easily seen to be $\mathbb{R}$-linear on E and, by our observation, it extends α. To show β is $\mathbb{C}$-linear, we need

$$\begin{aligned} \beta(ix) &= \beta_1(ix) - i\beta_1(i^2x) \\ &= \beta_1(ix) - i\beta_1(-x) \\ &= \beta_1(ix) + i\beta_1(x) \\ &= i(\beta_1(x) - i\beta_1(ix)) \\ &= i\beta(x). \end{aligned}$$

It remains to show that $|\beta(x)| \leq p(x)$ for $x \in E$. To do this, fix $x \in E$ and choose $e^{i\theta}\beta(x) \in \mathbb{R}$. Then

$$\begin{aligned} &|\beta(x)| = |e^{i\theta}\beta(x)| = |\beta(e^{i\theta}x)| \\ &= |\beta_1(e^{i\theta}x)| \\ &\leq p(e^{i\theta}x) = p(x). \quad \square \end{aligned}$$

Corollary 3.7. *If E is a normed space and $x \in E$ is a nonzero element, then there exists $\alpha \in E^*$ with*

$$\|\alpha\| = 1 \text{ and } \alpha(x) = \|x\|.$$

Proof. Let $M = \{\lambda x : \lambda \in \mathbb{K}\}$, a one-dimensional subspace of E. Define $\alpha\colon M \to \mathbb{K}$ by $\alpha(\lambda x) = \lambda\|x\|$. Then α is linear, and

$$\|\alpha\| = \sup_{\lambda \neq 0} \frac{|\alpha(\lambda x)|}{\|\lambda x\|} = 1.$$

Also $\alpha(x) = \|x\|$. By the Hahn-Banach theorem, we can extend α to a linear functional on the whole space E of norm 1. □

Corollary 3.8. *Let E be a normed space and $x, y \in E$ two distinct elements ($x \neq y$). Then there exists $\alpha \in E^*$ with $\alpha(x) \neq \alpha(y)$.*

Proof. Apply Corollary 3.7 to $x - y$ and observe that $\alpha(x-y) \neq 0 \Rightarrow \alpha(x) \neq \alpha(y)$. □

Corollary 3.9. *If E is any normed space, then there is a natural linear map*

$$J\colon E \to E^{**} = (E^*)^*$$

given by

$$J(x) = \text{ point evaluation at } x.$$

In other words, for each $x \in E$

$$\begin{aligned} J(x)\colon E^* &\to \mathbb{K} \\ J(x)(\alpha) &= \alpha(x). \end{aligned}$$

The map J is injective and

$$\|J(x)\|_{E^{**}} = \|x\|_E.$$

In short, J is a natural isometric isomorphism from E onto its range $J(E) \subset E^{**}$.

Proof. The map $J(x)\colon E^* \to \mathbb{K}$ is clearly linear, and

$$|J(x)(\alpha)| = |\alpha(x)| \leq \|\alpha\|\,\|x\|$$

shows that $J(x)$ is bounded on E^*. Thus $J(x) \in (E^*)^*$ and, in fact,

$$\|J(x)\|_{(E^*)^*} \leq \|x\|.$$

By Corollary 3.7, given $x \in E$, there exists $\alpha \in E^*$, with $\|\alpha\| = 1$ and $|\alpha(x)| = \|x\| = \|\alpha\|\,\|x\|$. This shows that $\|J(x)\| \geq \|x\|$. Therefore $\|J(x)\| = \|x\|$ and J is an isometry onto its range. It follows that J must be injective. □

Definition 3.10. A normed space E is called **reflexive** if the natural map $J \colon E \to E^{**}$ is surjective (and is then an isometric isomorphism because of Corollary 3.9).

Example 3.11. (i) *If* $1 < p < \infty$*, then* ℓ^p*,* $L^p[0,1]$ *and* $L^p(\mathbb{R})$ *are all reflexive.*

Proof. The idea is basically that

$$(\ell^p)^* = \ell^q, \qquad (\ell^q)^* = \ell^p \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right).$$

This shows that $(\ell^p)^{**} = \ell^p$ but is not quite a proof that ℓ^p is reflexive, because we still must check that the identification of $(\ell^p)^{**}$ with ℓ^p is the natural one. □

(ii) *Every finite-dimensional normed space is reflexive* (because E^* is the algebraic dual space by Corollary 1.8.10; so $\dim E^{**} = \dim E^* = \dim E$ and so the injective $J \colon E \to E^{**}$ must be surjective).

(iii) *If a normed space* E *is reflexive, then* E *must be a Banach space* (because dual spaces are complete — see Proposition 3.2).

Theorem 3.12. *If* E *is a normed space and* $M \subset E$ *is a closed subspace and if* $x_0 \in E$*,* $x_0 \notin M$*, then there exists* $\alpha \in E^*$ *with* $\alpha(x_0) = 1$ *and* $\alpha(x) = 0$ *for all* $x \in M$*.*

Proof. Let

$$d = \mathrm{dist}(x_0, M) = \inf\{\|x - x_0\| : x \in M\}.$$

Since M is closed and $x_0 \notin M$, $d > 0$. Next let

$$M_1 = \{x + \lambda x_0 : x \in M, \quad \lambda \in \mathbb{K}\}$$

and define a linear functional on M_1 by

$$\begin{aligned} \alpha \colon M_1 &\to \mathbb{K} \\ \alpha(x + \lambda x_0) &= \lambda. \end{aligned}$$

Observe that (for $\lambda \neq 0$)

$$\|x + \lambda x_0\| = |\lambda| \left\| \frac{x}{\lambda} + x_0 \right\| \geq |\lambda| d$$

and consequently

$$|\alpha(x + \lambda x_0)| = |\lambda| \leq \frac{1}{d}\|x + \lambda x_0\|$$

(true for all $x + \lambda x_0 \in M_1$, even when $\lambda = 0$). This shows that α is a continuous linear functional on M_1, that is $\alpha \in M_1^*$. By the Hahn-Banach theorem, we can extend α to get an element of E^* with the required properties. $\square$

Corollary 3.13. $(\ell^\infty)^* \neq \ell^1$.

More precisely, it is not possible to identify $(\ell^\infty)^*$ *with* ℓ^1 *in the same way as we identified* $(\ell^p)^*$ *with* ℓ^q *and* $(c_0)^*$ *with* ℓ^1 *in Examples 3.4.*

Proof. We know that $(c_0)^* = \ell^1$, c_0 a closed subspace of ℓ^∞ and $c_0 \neq \ell^\infty$.

In more detail, there is a map $T\colon \ell^1 \to (c_0)^*$ given by

$$(T(b))(a) = (T((b_n)_{n=1}^\infty))((a_n)_{n=1}^\infty) = \sum_{n=1}^\infty a_n b_n$$

for $b = (b_n)_{n=1}^\infty \in \ell^1$, $a = (a_n)_{n=1}^\infty \in c_0$. This map is an isometric isomorphism from ℓ^1 onto $(c_0)^*$. By the same argument (using Hölders inequality) used in Examples 3.4 to show that T is a bounded linear operator, we can see that it makes sense to define $T\colon \ell^1 \to (\ell^\infty)^*$ by the same formula. Also $\|T(b)\|_{(\ell^\infty)^*} \leq \|b\|_1$. Since $c_0 \subset \ell^\infty$ we know that this new variant of T satisfies $\|T(b)\|_{(\ell^\infty)^*} = \|b\|_1$ also. Our result is that T does not map ℓ^1 onto $(\ell^\infty)^*$.

By Theorem 3.12, there exists $\alpha \in (\ell^\infty)^*$, $\alpha \neq 0$ with $\alpha = 0$ on c_0. This α cannot be represented by an element of ℓ^1 — that is there is no choice of $(b_n)_n \in \ell^1$ so that

$$\alpha((a_n)_n) = \sum_n a_n b_n \qquad \text{for all } (a_n)_n \in \ell^\infty.$$ $\square$

Theorem 3.14. *If* E *is any infinite dimensional Banach space, then there exists a discontinuous linear transformation*

$$\alpha\colon E \to \mathbb{K}.$$

Proof. Since E is a vector space over $\mathbb{K}$, it must have an algebraic basis (or Hamel basis) — see Theorem 2.19. Let $\{e_i : i \in I\}$ be such a basis and recall that each $x \in E$ can be expressed (in a unique way) as a *finite* linear combination

$$x = x^{i_1} e_{i_1} + x^{i_2} e_{i_2} + \cdots + x^{i_n} e_{i_n}$$

of basis elements. Write this as

$$x = \sum_{i \in I} x^i e_i$$

where $x^i = 0$ if $i \neq i_j$ for some j.

Consider now the coefficient functionals

$$\begin{aligned} \alpha_i \colon E &\to \mathbb{K} \\ \colon x &\mapsto x^i \end{aligned}$$

We *claim* that at least one of the linear transformations α_i ($i \in I$) must be discontinuous.

If all the α_i are continuous we take an infinite sequence $(i_n)_n$ of distinct elements of I. Let

$$x_n = \sum_{j=1}^{n} \frac{1}{2^j} \frac{e_{i_j}}{\|e_{i_j}\|}.$$

It is easy to check that $(x_n)_n$ is a Cauchy sequence in E, and hence there exists $\lim_{n\to\infty} x_n = x \in E$. If each α_{i_j} is continuous

$$\alpha_{i_j}(x) = \lim_n \alpha_{i_j}(x_n) = \frac{1}{2^j \|e_{i_j}\|}$$

is nonzero for infinitely many i_j. This is a contradiction because

$$x = \sum_{i \in I} \alpha_i(x) e_i$$

is a finite sum. □