4. Dual spaces and weak topologies

Recall that if X is a Banach space, we write X^* for its dual. This was defined as the space of all continuous (or bounded) linear functionals $F : X \to \mathbb{C}$. We know from the special case $Y = \mathbb{C}$ of Theorem 2.12 that X^* itself is a Banach space, too, if we use the operator norm

$$\|F\| = \sup_{\|x\|=1} |F(x)| \qquad (F \in X^*).$$

The following fundamental result makes sure that there is a large supply of bounded linear functionals on every normed space.

Theorem 4.1 (Hahn-Banach)**.** *Let X be a normed space and let M be a subspace of X. Suppose that $F : M \to \mathbb{C}$ is a linear map satisfying $|F(x)| \le C\|x\|$ ($x \in M$). Then there exists a linear extension $G : X \to \mathbb{C}$ of F satisfying $|G(x)| \le C\|x\|$ for all $x \in X$.*

In other words, a bounded linear functional on a subspace can always be extended to the whole space *without increasing the norm.* This latter property is the point here; it is easy, at least in principle, to linearly extend a given functional. (Sketch: Fix a basis of M as a vector space, extend to a basis of the whole space and assign arbitrary values on these new basis vectors.)

Proof. We first prove a real version of theorem. So, for the time being, let X be a *real* vector space, and assume that $F : M \to \mathbb{R}$ is a bounded linear functional on a subspace.

Roughly speaking, the extension will be done one step at a time. So our first goal is to show that F can be extended to a one-dimensional extension of M in such a way that the operator norm is preserved. We are assuming that $|F(x)| \le C\|x\|$; in fact, since $C\|x\|$ defines a new norm on X (if $C > 0$), we can assume that $C = 1$ here.

Now let $x_1 \in X$, $x_1 \notin M$. We want to define a linear extension F_1 of F on the (slightly, by one dimension) bigger space

$$M_1 = \{x + cx_1 : x \in M, c \in \mathbb{R}\}.$$

Such a linear extension is completely determined by the value $f = F_1(x_1)$ (and, conversely, every $f \in \mathbb{R}$ will define an extension). Since we also want an extension that still satisfies $|F_1(y)| \le \|y\|$ ($y \in M_1$), we're looking for an $f \in \mathbb{R}$ such that

$$-\|x + cx_1\| \le F(x) + cf \le \|x + cx_1\| \tag{4.1}$$

for all $c \in \mathbb{R}$, $x \in M$. By assumption, we already have this for $c = 0$, and by discussing the cases $c > 0$ and $c < 0$ separately, we see that

(4.1) is equivalent to

$$-\left\|\frac{x}{c}+x_1\right\| - F\left(\frac{x}{c}\right) \le f \le \left\|\frac{x}{c}+x_1\right\| - F\left(\frac{x}{c}\right)$$

for all $c \neq 0$, $x \in M$. In other words, there will be an extension F_1 with the desired properties if (and only if, but of course that is not our concern here)

$$\|z+x_1\| - F(z) \ge -\|y+x_1\| - F(y)$$

for arbitrary $y, z \in M$. This is indeed the case, because

$$F(z) - F(y) = F(z-y) \le \|z-y\| \le \|z+x_1\| + \|x_1+y\|.$$

We now use Zorn's Lemma to obtain a norm preserving extension to all of X (this part of the proof can be safely skipped if you're not familiar with this type of argument). We consider the set of all linear extensions G of F that satisfy $|G(x)| \le \|x\|$ on the subspace on which they are defined. This set can be partially ordered by declaring $G \prec G'$ if G' is an extension of G. Now if $\{G_\alpha\}$ is a totally ordered subset (any two G_α's can be compared) and if we denote the domain of G_α by M_α, then $G : \bigcup M_\alpha \to \mathbb{R}$, $G(x) = G_\alpha(x)$ defines an extension of all the G_α's, that is, $G \succ G_\alpha$ for all α. Note that there are no consistency problems in the definition of G because if there is more than one possible choice for α for any given x, then the corresponding G_α's must give the same value on x because one of them is an extension of the other.

We have verified the hypotheses of Zorn's Lemma. The conclusion is that there is a G that is maximal in the sense that if $H \succ G$, then $H = G$. This G must be defined on the whole space X because otherwise the procedure described above would give an extension H to a strictly bigger space. We have proved the real version of the Hahn-Banach Theorem.

The original, complex version can be derived from this by some elementary, but ingenious trickery, as follows: First of all, we can think of X and M as real vector spaces also (we just refuse to multiply by non-real numbers and otherwise keep the algebraic structure intact). Moreover, $L_0(x) = \text{Re } F(x)$ defines an $\mathbb{R}$-linear functional $L_0 : M \to \mathbb{R}$. By the real version of the theorem, there exists an $\mathbb{R}$-linear extension $L : X \to \mathbb{R}$, $|L(x)| \le \|x\|$.

I now claim that the following definition will work: $G(x) = L(x) - iL(ix)$ Indeed, it is easy to check that $G(x+y) = G(x) + G(y)$, and if

$c = a + ib \in \mathbb{C}$, then

$$\begin{aligned} G(cx) &= L(ax + ibx) - iL(iax - bx) \\ &= aL(x) + bL(ix) - iaL(ix) + ibL(x) \\ &= (a + ib)L(x) + (b - ia)L(ix) = c(L(x) - iL(ix)) = cG(x). \end{aligned}$$

So G is $\mathbb{C}$-linear. It is also an extension of F because if $x \in M$, then $L(x) = L_0(x) = \text{Re } F(x)$ and thus

$$\begin{aligned} G(x) &= \text{Re } F(x) - i \text{ Re } F(ix) = \text{Re } F(x) - i \text{ Re}(iF(x)) \\ &= \text{Re } F(x) + i \text{ Im } F(x) = F(x). \end{aligned}$$

Finally, if we write $G(x) = |G(x)|e^{i\varphi(x)}$, we see that

$$\begin{aligned} |G(x)| &= G(x)e^{-i\varphi(x)} = G(e^{-i\varphi(x)}x) = \text{Re } G(e^{-i\varphi(x)}x) \\ &= L(e^{-i\varphi(x)}x) \leq \|e^{-i\varphi(x)}x\| = \|x\|. \end{aligned}$$

□

Here are some immediate important consequences of the Hahn-Banach Theorem. They confirm that much can be learned about a Banach spaces by studying its dual. For example, part (b) says that norms can be computed by testing functionals on the given vector x.

Corollary 4.2. *Let X, Y be normed spaces.*
(a) X^* separates the points X, that is, if $x, y \in X$, $x \neq y$, then there exists an $F \in X^*$ with $F(x) \neq F(y)$.
(b) For all $x \in X$, we have

$$\|x\| = \sup\{|F(x)| : F \in X^*, \|F\| = 1\}.$$

(c) If $T \in B(X, Y)$, then

$$\|T\| = \sup\{|F(Tx)| : x \in X, F \in Y^*, \|F\| = \|x\| = 1\}.$$

Proof of (b). If $F \in X^*$, $\|F\| = 1$, then $|F(x)| \leq \|x\|$. This implies that $\sup |F(x)| \leq \|x\|$. On the other hand, $F_0(cx) = c\|x\|$ defines a linear functional on the one-dimensional subspace $L(x)$ that satisfies $|F_0(y)| \leq \|y\|$ for all $y = cx \in L(x)$ (in fact, we have equality here). By the Hahn-Banach Theorem, there exists an extension $F \in X^*$, $\|F\| = 1$ of F_0; by construction, $|F(x)| = \|x\|$, so $\sup |F(x)| \geq \|x\|$. This completes the proof, and, in fact, this argument has also shown that the supremum is attained; it is a maximum. □

Exercise 4.1. Prove parts (a) and (c) of Corollary 4.2.

Let X be a Banach space. Since X^* is a Banach space, too, we can form *its* dual $X^{**} = (X^*)^*$. We call X^{**} the *bidual* or *second dual* of X. We can identify the original space X with a closed subspace of X^{**} in a natural way, as follows: Define a map $j : X \to X^{**}$, $j(x)(F) = F(x)$ ($x \in X$, $F \in X^*$). In other words, vectors $x \in X$ act in a natural way on functionals $F \in X^*$: we just evaluate F on x.

Proposition 4.3. *We have $j(x) \in X^{**}$, and the map j is a (linear) isometry. In particular, $j(X) \subseteq X^{**}$ is a closed subspace of X^{**}.*

An operator $I : X \to Y$ is called an *isometry* if $\|Ix\| = \|x\|$ for all $x \in X$.

Exercise 4.2. (a) Show that an isometry I is always injective, that is, $N(I) = \{0\}$.
(b) Show that $S : \ell^1 \to \ell^1$, $Sx = (0, x_1, x_2, \dots)$ is an isometry that is not onto, that is $R(S) \neq \ell^1$.

Proof. We will only check that j is an isometry and that $j(X)$ is a closed subspace.

Exercise 4.3. Prove the remaining statements from Proposition 4.3. More specifically, prove that $j(x)$ is a linear, bounded functional on X^* for every $x \in X$, and prove that the map $x \mapsto j(x)$ is itself linear.

By the definition of the operator norm and Corollary 4.2(b), we have

$$\begin{aligned} \|j(x)\| &= \sup\{|j(x)(F)| : F \in X^*, \|F\| = 1\} \\ &= \sup\{|F(x)| : F \in X^*, \|F\| = 1\} = \|x\|, \end{aligned}$$

so j indeed is an isometry. Clearly, $j(X)$ is a subspace (being the image of a linear map). If $y_n \in j(X)$, that is, $y_n = j(x_n)$, and $y_n \to y$, then also $x_n \to x$ for some $x \in X$, because y_n is a Cauchy sequence, and since j preserves norms, so is x_n. Since j is continuous, it follows that $j(x) = \lim j(x_n) = y$, so $y \in j(X)$. □

A linear isometry preserves all structures on a Banach space (the vector space structure and the norm), and thus provides an identification of its domain with its image. Using j and Proposition 4.3, we can therefore think of X as a closed subspace of X^{**}. If, in this sense, $X = X^{**}$, we call X *reflexive.* This really means that $j(X) = X^{**}$. In particular, note that for X to be reflexive, it is not enough to have X isometrically isomorphic to X^{**}; rather, we need this isometric isomorphism to be specifically j.

We now use dual spaces to introduce new topologies on Banach spaces. If $\mathcal{T}_1$, $\mathcal{T}_2$ are two topologies on a common space X, we say that $\mathcal{T}_1$ is *weaker* than $\mathcal{T}_2$ (or $\mathcal{T}_2$ is *stronger* than T_1) if $\mathcal{T}_1 \subseteq \mathcal{T}_2$. In topology, *coarse* and *fine* mean the same thing as weak and strong, respectively, but it would be uncommon to use these alternative terms in functional analysis.

Given a set X and a topological space $(Y, \mathcal{T})$ and a family $\mathcal{F}$ of maps $F : X \to Y$, there exists a weakest topology on X that makes all $F \in \mathcal{F}$ continuous. Let us try to give a description of this topology (and, in fact, we also need to show that such a topology exists). We will denote it by $\mathcal{T}_w$.

Clearly, we must have $F^{-1}(U) \in \mathcal{T}_w$ for all $F \in \mathcal{F}$, $U \in \mathcal{T}$. Conversely, any topology that contains these sets will make the F's continuous. So we could stop here and say that $\mathcal{T}_w$ is the topology generated by these sets. (Given any collection of sets, there always is a weakest topology containing these sets.) However, we would like to be somewhat more explicit. It is clear that finite intersections of sets of this type have to be in $\mathcal{T}_w$, too; in other words,

$$\{x \in X : F_1(x) \in U_1, \ldots, F_n(x) \in U_n\} \tag{4.2}$$

belongs to $\mathcal{T}_w$ for arbitrary choices of $n \in \mathbb{N}$, $F_j \in \mathcal{F}$, $U_j \in \mathcal{T}$. If these sets are open, then arbitrary unions of such sets need to belong to $\mathcal{T}_w$, and, fortunately, the process stops here: we don't get additional sets if we now take finite intersections again. So the claim is that

$$\mathcal{T}_w = \{ \text{ arbitrary unions of sets of the type (4.2) } \}. \tag{4.3}$$

We must show that $\mathcal{T}_w$ is a topology; by its construction, any other topology that makes all $F \in \mathcal{F}$ continuous must then be stronger than $\mathcal{T}_w$. This verification is quite straightforward, but a little tedious to write down, so I'll make this an exercise:

Exercise 4.4. Prove that (4.2), (4.3) define a topology.

We now apply this process to a Banach space X, with $Y = \mathbb{C}$ and $\mathcal{F} = X^*$. Of course, we already have a topology on X (the norm topology); this new topology will be different, unless X is finite-dimensional. Here's the formal definition:

Definition 4.4. Let X be a Banach space. The *weak topology* on X is defined as the weak topology $\mathcal{T}_w$ generated by X^*.

If we denote the norm topology by $\mathcal{T}$, then, since all $F \in X^*$ are continuous if we use $\mathcal{T}$ (by definition of X^*!), we see that $\mathcal{T}_w \subseteq \mathcal{T}$; in other words, the weak topology is weaker than the norm topology. By the discussion above, (4.3) gives a description of $\mathcal{T}_w$. A slightly more convenient variant of this can be obtained by making use of the vector space structure. First of all, the sets

$$U(F_1, \ldots, F_n; \epsilon_1, \ldots, \epsilon_n) = \{x \in X : |F_j(x)| < \epsilon_j \ (j = 1, \ldots, n)\} \tag{4.4}$$

are in $\mathcal{T}_w$ for arbitrary $n \in \mathbb{N}$, $F_j \in X^*$, $\epsilon_j > 0$. In fact, they are of the form (4.2), with $U_j = \{z : |z| < \epsilon_j\}$. I now claim that $V \in \mathcal{T}_w$ if and only if for every $x \in V$, there exists a set $U = U(F_j; \epsilon_j)$ of this form with $x + U \subseteq V$.

Exercise 4.5. Prove this claim.

We can rephrase this as follows: The U's form a neighborhood base at 0 (that is, any neighborhood of $x = 0$ contains some U) and the neighborhoods of an arbitrary $x \in X$ are precisely the translates $x + W$ of the neighborhoods W of 0.

We'll make two more observations on the weak topology and then leave the matter at that. First of all, $\mathcal{T}_w$ is a Hausdorff topology: If $x, y \in X$, $x \neq y$, then there exist $V, W \in \mathcal{T}_w$ with $x \in V$, $y \in W$, $V \cap W = \emptyset$. To prove this, we use the fact that X^* separates the points of X; see Corollary 4.2(a). So there is an $F \in X^*$ with $F(x) \neq F(y)$. We can now take $V = x + U(F; \epsilon)$, $W = y + U(F; \epsilon)$ with a sufficiently small $\epsilon > 0$.

Exercise 4.6. Provide the details for this last step. You can (and should) make use of the description of $\mathcal{T}_w$ established above, in Exercise 4.5.

Finally, if x_n is a sequence from X, then $x_n \to x$ in $\mathcal{T}_w$ (this is usually written as $x_n \xrightarrow{w} x$; x_n goes to x weakly) if and only if $F(x_n) \to F(x)$ for all $F \in X^*$.

Exercise 4.7. Prove this. Again, by the results of Exercise 4.5, $x_n \xrightarrow{w} x$ means that for every U of the form (4.4), we eventually have $x_n - x \in U$.

This gives a characterization of convergent sequences and thus some idea of what the topology does. However, it can happen that $\mathcal{T}_w$ is not metrizable and then the topological notions (closed sets, compactness, continuity etc.) can *not* be characterized using sequences.

Definition 4.5. Let X be a Banach space. The *weak-$*$ topology* $\mathcal{T}_{w^*}$ on X^* is defined as the weak topology generated by X, viewed as a subset of X^{**}. Put differently, $\mathcal{T}_{w^*}$ is the weakest topology that turns all point evaluations $j(x) : X^* \to \mathbb{C}$, $F \mapsto F(x)$ $(x \in X)$ into continuous functions on X^*.

We have an analogous description of $\mathcal{T}_{w^*}$. The sets

$$U(x_1, \ldots, x_n; \epsilon_1, \ldots, \epsilon_n) = \{F \in X^* : |F(x_j)| < \epsilon_j\}$$

are open, and $V \subseteq X^*$ is open in the weak-$*$ topology if and only for every $F \in V$, there exists such a U so that $F + U \subseteq V$.

Exercise 4.8. Prove that $\mathcal{T}_{w^*}$ is a Hausdorff topology. *Hint:* If $F \neq G$, then $F(x) \neq G(x)$ for some $x \in X$. Now you can build disjoint neighborhoods of F, G as above, using this x; see also Exercise 4.6.

Exercise 4.9. Let X be a Banach space, and let $F_n, F \in X^*$. Show that $F_n \to F$ in the weak-$*$ topology if and only if $F_n(x) \to F(x)$ for all $x \in X$.

Since X^* is a Banach space, we can also define a weak topology on X^*. This is the topology generated by X^{**}. The weak-$*$ topology is generated by X, which in general is a smaller set of maps, so the weak-$*$ topology is weaker than the weak topology. It can only be defined on a dual space. If X is reflexive, then there's no difference between the weak and weak-$*$ topologies.

Despite its clumsy and artificial looking definition, the weak-$*$ topology is actually an extremely useful tool. All the credit for this is due to the following fundamental result.

Theorem 4.6 (Banach-Alaoglu)**.** *Let X be a Banach space. Then the closed unit ball $\overline{B}_1(0) = \{F \in X^* : \|F\| \le 1\}$ is compact in the weak-$*$ topology.*

Proof. This will follow from *Tychonoff's Theorem:* The product of compact topological spaces is compact in the product topology. To get this argument started, we look at the Cartesian product set

$$K = \prod_{x \in X} \{z \in \mathbb{C} : |z| \le \|x\|\}.$$

As a set, this is defined as the set of maps $F : X \to \mathbb{C}$ with $|F(x)| \le \|x\|$. The individual factors $\{|z| \le \|x\|\}$ come with a natural topology, and we endow K with the product topology, which, by definition, is the weak topology generated by the projections $p_x : K \to \{|z| \le \|x\|\}$, $p_x(F) = F(x)$ (equivalently, you can also produce it from cylinder sets, if this is more familiar to you). By Tychonoff's Theorem, K is compact.

Now $\overline{B}_1(0) \subseteq K$; more precisely, $\overline{B}_1(0)$ consists of those maps $F \in K$ that are also linear. I now claim that the topology induced by K on $\overline{B}_1(0)$ is the same as the induced topology coming from the weak-$*$ topology on $X^* \supseteq \overline{B}_1(0)$. This should not come as a surprise because both the product topology and $\mathcal{T}_{w^*}$ are weak topologies generated by the point evaluations $x \mapsto F(x)$. Writing it down is a slightly unpleasant task that is best delegated to the reader.

Exercise 4.10. Show that K and $(X^*, \mathcal{T}_{w^*})$ indeed induce the same topology on $\overline{B}_1(0)$. Come to think of it, we perhaps really want to prove the following abstract fact: Let $(Z, \mathcal{T})$ be a topological space, $\mathcal{F}$ a family of maps $F : X \to Z$ and let $Y \subseteq X$. Then we can form the weak topology $\mathcal{T}_w$ on X; this induces a relative topology on Y. Alternatively, we can restrict the maps $F \in \mathcal{F}$ to Y and let the restrictions generate a weak topology on Y. Prove that both methods lead to the same topology. As usual, this is mainly a matter of unwrapping definitions. You could use the description (4.2), (4.3) of the weak topologies and look at what happens when these induce relative topologies.

Exercise 4.11. (a) Let Y be a compact topological space and let $A \subseteq Y$ be closed. Prove that then A is compact, too.
(b) Let Y be a topological space. Show that a subset $B \subseteq Y$ is compact if and only if B with the relative topology is a compact topological space. (Sometimes compactness is defined in this way; recall that we defined compact sets by using covers by open sets $U \subseteq Y$. It is now in fact almost immediate that we get the same condition from both variants, but this fact will be needed here, so I thought I'd point it out.)

With these preparations out of the way, it now suffices to show that $\overline{B}_1(0)$ is closed in K. So let $F \in K \setminus \overline{B}_1(0)$. We want to find a neighborhood of F that is contained in $K \setminus \overline{B}_1(0)$ (note that we cannot use sequences here because there is no guarantee that our topologies are metrizable). Since $F \notin \overline{B}_1(0)$, F is not linear and thus there are $c, d \in \mathbb{C}$, $x, y \in X$ such that

$$\epsilon \equiv |F(cx + dy) - cF(x) - dF(y)| > 0.$$

But then

$$V = \Big\{ G \in K : |G(cx + dy) - F(cx + dy)| < \frac{\epsilon}{3}, \\ |c|\,|G(x) - F(x)| < \frac{\epsilon}{3}, |d|\,|G(y) - F(y)| < \frac{\epsilon}{3} \Big\}$$

is an open set in K with $F \in V$, and if $G \in V$, then still

$$|G(cx + dy) - cG(x) - dG(y)| > 0,$$

so V does not contain any linear maps and thus is the neighborhood of F we wanted to find. $\square$

We have already seen how the fact that $\mathcal{T}_{w^*}$ need not be metrizable makes this topology a bit awkward to deal with. The following result provides some relief. We call a metric space X *separable* if X has a countable dense subset (that is, there exist $x_n \in X$ such that if $x \in X$ and $\epsilon > 0$ are arbitrary, then $d(x, x_n) < \epsilon$ for some $n \in \mathbb{N}$).

Exercise 4.12. (a) Show that ℓ^p (with index set $\mathbb{N}$, as usual) is separable for $1 \leq p < \infty$. You can use the result of Exercise 4.13 below if you want.
(b) Show that ℓ^∞ is not separable. *Suggestion:* Consider all $x \in \ell^\infty$

that only take the values 0 and 1. How big is this set? What can you say about $\|x - x'\|$ for two such sequences?

Exercise 4.13. Let X be a Banach space. Show that X will be separable if there is a countable *total* subset, that is, if there is a countable set $M \subseteq X$ such that the (finite) linear combinations of elements from M are dense in X (in other words, if $x \in X$ and $\epsilon > 0$, we must be able to find $m_j \in M$ and coefficients $c_j \in \mathbb{C}$ such that $\left\| \sum_{j=1}^{N} c_j m_j - x \right\| < \epsilon$.)

Theorem 4.7. *If X is a separable Banach space, then the weak-$*$ topology on $\overline{B}_1(0) \subseteq X^*$ (more precisely: the relative topology induced by $\mathcal{T}_{w^*}$) is metrizable.*

We don't want to prove this in detail, but the basic idea is quite easy to state. The formula

$$d(F, G) = \sum_{n=1}^{\infty} 2^{-n} \frac{|F(x_n) - G(x_n)|}{1 + |F(x_n) - G(x_n)|}$$

(say), where $\{x_n\}$ is a dense subset of X, defines a metric that generates the desired topology.

Corollary 4.8. *Let X be a separable Banach space. If $F_n \in X^*$, $\|F_n\| \le 1$, then there exist $F \in X^*$, $\|F\| \le 1$ and a subsequence $n_j \to \infty$ such that $F_{n_j}(x) \to F(x)$ for all $x \in X$.*

Proof. This follows by just putting things together. By the Banach-Alaoglu Theorem, $\overline{B}_1(0)$ is compact in the weak-$*$ topology. By Theorem 4.7, this can be thought of as a metric space. By Theorem 1.7(c), compactness is therefore equivalent to sequences having convergent subsequences. By using Exercise 4.9, we now obtain the claim. □

To make good use of the results of this chapter, we need to know what the dual space of a given space is. We now investigate this question for the Banach spaces from our list that was compiled in Chapter 2.

Example 4.1. If $X = \mathbb{C}^n$ with some norm, then Corollary 2.16 implies that all linear functionals on X are bounded, so in this case X^* coincides with the algebraic dual space. From linear algebra we know that as a vector space, X^* can be identified with $\mathbb{C}^n$ again; more precisely, $y \in \mathbb{C}^n$ can be identified with the functional $x \mapsto \sum_{j=1}^{n} y_j x_j$. It also follows from this that X is reflexive. The norm on X^* depends on the norm on X; Example 4.3 below will throw some additional light on this.

Exercise 4.14. Show that the weak topology on $X = \mathbb{C}^n$ coincides with the norm topology. *Suggestion:* It essentially suffices to check that open balls $B_r(0)$ (say) are in $\mathcal{T}_w$. Show this and then use the definition of the norm topology $\mathcal{T}$ to show that $\mathcal{T} \subseteq \mathcal{T}_w$. Since always $\mathcal{T}_w \subseteq \mathcal{T}$, this will finish the proof.

This Exercise says that we really don't get anything new from the theory of this chapter on finite-dimensional spaces; recall also in this context that $\mathcal{T}_w = \mathcal{T}_{w^*}$ on $\mathbb{C}^n$, thought of as the dual space X^* of $X = \mathbb{C}^n$, because X is reflexive.

Example 4.2. Let K be a compact Hausdorff space and consider the Banach space $C(K)$. Then $C(K)^* = \mathcal{M}(K)$, where $\mathcal{M}(K)$ is defined as the space of all complex, regular Borel measures on K. Here, we call a (complex) measure μ (inner and outer) *regular* if its total variation $\nu = |\mu|$ is regular in the sense that

$$\nu(B) = \sup_{L \subseteq B: L \text{ compact}} \nu(L) = \inf_{U \supseteq B: U \text{ open}} \nu(U)$$

for all Borel sets $B \subseteq K$. $\mathcal{M}(K)$ becomes a vector space if the vector space operations are introduced in the obvious way as $(\mu + \nu)(B) = \mu(B) + \nu(B)$, $(c\mu)(B) = c\mu(B)$. In fact, $\mathcal{M}(K)$, equipped with the norm

$$\|\mu\| = |\mu|(K), \tag{4.5}$$

is a Banach space. This is perhaps most elegantly deduced from the main assertion of this Example, namely the fact that $C(K)^*$ can be identified with $\mathcal{M}(K)$, and, as we will see in a moment, the operator norm on $\mathcal{M}(K) = C(K)^*$ turns out to be exactly (4.5). More precisely, the claim is that every $\mu \in \mathcal{M}(K)$ generates a functional $F_\mu \in C(K)^*$ via

$$F_\mu(f) = \int_K f(x)\, d\mu(x), \tag{4.6}$$

and we also claim that the corresponding map $\mathcal{M}(K) \to C(K)^*$, $\mu \mapsto F_\mu$ is an isomorphism between Banach spaces (in other words, a bijective, linear isometry).

The *Riesz Representation Theorem* does the lion's share of the work here; it implies that $\mu \mapsto F_\mu$ is a bijection from $\mathcal{M}(K)$ onto $C(K)^*$; see for example, Folland, Real Analysis, Corollary 7.18. It is also clear that this map is linear.

Exercise 4.15. Suppose we introduce a norm on $\mathcal{M}(K)$ by just declaring $\|\mu\| = \|F_\mu\|$ (operator norm), that is, we just move the norm on $C(K)^*$ over to $\mathcal{M}(K)$. Show that this leads to (4.5); put differently, show that the operator norm of F_μ from (4.6) equals $|\mu|(K)$.

One pleasing consequence of this identification of $C(K)^*$ as $\mathcal{M}(K)$ is the fact that this latter space, being a dual space, can now be equipped with a weak-$*$ topology. This, in turn, has implications of the following type:

Exercise 4.16. Show that $C[a,b]$ (so $K=[a,b]$, with the usual topology) is separable. *Suggestion:* Deduce this from the Weierstraß approximation theorem.

Exercise 4.17. Let μ_n be a sequence of complex Borel measures on $[a,b]$ with $|\mu_n|([a,b]) \le 1$ (in particular, these could be arbitrary positive measures with $\mu_n([a,b]) \le 1$). Show that there exists another Borel measure μ on $[a,b]$ with $|\mu|([a,b]) \le 1$ and a subsequence n_j such that

$$\lim_{j\to\infty} \int_{[a,b]} f(x)\, d\mu_{n_j}(x) = \int_{[a,b]} f(x)\, d\mu(x)$$

for all $f \in C[a,b]$.

Hint: This in fact follows quickly from Corollary 4.8 and Exercise 4.16.

Example 4.3. We now move on to ℓ^p spaces. We first claim that if $1 \le p < \infty$, then $(\ell^p)^* = \ell^q$, where $1/p + 1/q = 1$. More precisely, the claim really is that every $y \in \ell^q$ generates a functional $F_y \in (\ell^p)^*$, as follows:

$$F_y(x) = \sum_{j=1}^{\infty} y_j x_j \tag{4.7}$$

Moreover, the corresponding map $y \mapsto F_y$ is an isomorphism between the Banach spaces ℓ^q and $(\ell^p)^*$.

Let us now prove these assertions. First of all, Hölder's inequality shows that the series from (4.7) converges, and in fact $|F_y(x)| \le \|y\|_q \|x\|_p$. Since (4.7) is also obviously linear in x, this shows that $F_y \in (\ell^p)^*$ and $\|F_y\| \le \|y\|_q$. We will now explicitly discuss only the case $1 < p < \infty$. If $p = 1$ (and thus $q = \infty$), the same basic strategy works and actually the technical details get easier, but some slight adjustments are necessary.

To compute $\|F_y\|$, we set

$$x_n = \begin{cases} \frac{|y_n|^q}{y_n} & n \le N, y_n \ne 0 \\ 0 & \text{else} \end{cases}, \tag{4.8}$$

with $N \in \mathbb{N}$. It is then clear that $x \in \ell^p$,

$$\|x\|_p^p = \sum_{n=1}^{N} |y_n|^{(q-1)p} = \sum_{n=1}^{N} |y_n|^q,$$

and thus

$$F_y(x) = \sum_{n=1}^{N} |y_n|^q = \left(\sum_{n=1}^{N} |y_n|^q \right)^{1/q} \|x\|_p.$$

Thus $\|F_y\| \geq \left(\sum_{n=1}^{N} |y_n|^q \right)^{1/q}$, and this holds for arbitrary $N \in \mathbb{N}$, so it follows that $\|F_y\| \geq \|y\|_q$, so $\|F_y\| = \|y\|_q$. This says that the identification $y \mapsto F_y$ is isometric, and it is obviously linear (in y!), so it remains to show that it is surjective, that is, every $F \in (\ell^p)^*$ equals some F_y for suitable $y \in \ell^q$. To prove this, fix $F \in (\ell^p)^*$. It is clear from (4.7) that only $y_n = F(e_n)$ can work, so define a sequence y in this way (here, $e_n(j) = 1$ if $j = n$ and $e_n(j) = 0$ otherwise). Consider again the $x \in \ell^p$ from (4.8). Then

$$F(x) = F\left(\sum_{n=1}^{N} \frac{|y_n|^q}{y_n} e_n \right) = \sum_{n=1}^{N} |y_n|^q.$$

As above, $\|x\|_p = \left(\sum_{n=1}^{N} |y_n|^q \right)^{1/p}$, so

$$\sum_{n=1}^{N} |y_n|^q \leq \|F\| \left(\sum_{n=1}^{N} |y_n|^q \right)^{1/p}$$

or $\left(\sum_{n=1}^{N} |y_n|^q \right)^{1/q} \leq \|F\|$. Again, N is arbitrary here, so $y \in \ell^q$. By construction of y, we have $F(z) = F_y(z)$ if z is a finite linear combination of the e_n's. These vectors z, however, are dense in ℓ^p, so it follows from the continuity of both F and F_y that $F(w) = F_y(w)$ for *all* $w \in \ell^p$; in other words, $F = F_y$.

As a consequence, ℓ^p is reflexive for $1 < p < \infty$, basically because $(\ell^p)^{**} = (\ell^q)^* = \ell^p$, by applying the above result on the dual of ℓ^r twice.

Exercise 4.18. Give a careful version of this argument, where the identification $j : X \to X^{**}$ from the definition of reflexivity is taken seriously.

We can't be sure about ℓ^1 and ℓ^∞ at this point because we don't know yet what $(\ell^\infty)^*$ is. It will turn out that these spaces are not reflexive.

Example 4.4. Similar discussions let us identify the duals of c_0 and c. We claim that $c_0^* = \ell^1 = \ell^1(\mathbb{N})$ and $c^* = \ell^1(\mathbb{N}_0)$; as usual, we really mean that there are Banach space isomorphisms (linear, bijective isometries) $y \mapsto F_y$ that provide identifications between these spaces. We can make this more explicit:

$$F_y(x) = \sum_{j=1}^{\infty} y_j x_j \quad (y \in \ell^1(\mathbb{N}), x \in c_0),$$

$$F_y(x) = y_0 \cdot \left(\lim_{n\to\infty} x_n \right) + \sum_{j=1}^{\infty} y_j \left(x_j - \lim_{n\to\infty} x_n \right) \quad (y \in \ell^1(\mathbb{N}_0), x \in c)$$

Since this discussion is reasonably close to Example 4.3, I don't want to do it here. The above representations of the dual spaces as $\ell^1(\mathbb{N})$ and $\ell^1(\mathbb{N}_0)$ seem natural, especially since $c = c_0 \dot{+} L(e)$, with $e = (1, 1, 1, \ldots)$. This then also allows us to similarly identify c_0^* with a codimension 1 subspace M of c^*; we can see this in abstract style by taking $M = \{F \in c^* : F(e) = 0\}$ (and a careful proof would use the Hahn-Banach theorem). However, it is of course easier to read it off from the above representation of the dual spaces as $\ell^1(\mathbb{N})$ and $\ell^1(\mathbb{N}_0)$, respectively.

On the other hand, $\ell^1(\mathbb{N})$ and $\ell^1(\mathbb{N}_0)$ are of course isometrically isomorphic as Banach spaces: the map $(y_n)_{n\ge1} \mapsto (y_{n+1})_{n\ge0}$ is an isometry. The most interesting aspect of this is that $c \ncong c_0$; this can be seen by considering extreme points of the closed unit ball in both spaces. So the space can not, in general, be recovered from its dual.

Example 4.5. We now discuss $(\ell^\infty)^*$. We will obtain an explicit looking description of this dual space, too, but, actually, this result will not be very useful. This is so because the objects that we will obtain are not particularly well-behaved and there is no well developed machinery that would recommend them for further use.

It will turn out that $(\ell^\infty)^* = \mathcal{M}_{fa}(\mathbb{N})$, the space of bounded, finitely additive set functions on $\mathbb{N}$. More precisely, the elements of $\mathcal{M}_{fa}(\mathbb{N})$ are set functions $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{C}$ that satisfy $\sup_{M\subseteq\mathbb{N}} |\mu(M)| < \infty$ and if $M_1, M_2 \subseteq \mathbb{N}$ are disjoint, then $\mu(M_1 \cup M_2) = \mu(M_1) + \mu(M_2)$. Note that the complex measures on the σ-algebra $\mathcal{P}(\mathbb{N})$ are precisely those $\mu \in \mathcal{M}_{fa}(\mathbb{N})$ that are σ-additive rather than just finitely additive.

Finitely additive bounded set functions will act on vectors $x \in \ell^\infty$ by an integration of sorts. We discuss this new integral first (as far as I can see, this integral does not play any major role in analysis except in this particular context). Let $x \in \ell^\infty$. We subdivide the disk $\{z \in \mathbb{C} : |z| \le \|x\|\}$ into squares (say) Q_j, and we fix a number $z_j \in Q_j$ from each square. Let $M_j = \{n \in \mathbb{N} : x_n \in Q_j\}$ be the inverse image under x_n. It's quite easy to check that for $\mu \in \mathcal{M}_{fa}(\mathbb{N})$, the sums $\sum z_j \mu(M_j)$ will approach a limit as the subdivision gets finer and finer, and this limit is independent of the choice of the Q_j and z_j. We call this limit the *Radon integral* of x_n with respect to μ, and we denote it by

$$R - \int_{\mathbb{N}} x_n \, d\mu(n).$$

Next, we show how to associate a set function $\mu \in \mathcal{M}_{fa}(\mathbb{N})$ with a given functional $F \in (\ell^\infty)^*$. Define $\mu(M) = F(\chi_M)$ ($M \subseteq \mathbb{N}$). Then $|\mu(M)| \le \|F\| \, \|\chi_M\| \le \|F\|$, so μ is a bounded set function. Also, if $M_1 \cap M_2 = \emptyset$, then

$$\begin{aligned} \mu(M_1 \cup M_2) &= F(\chi_{M_1 \cup M_2}) = F(\chi_{M_1} + \chi_{M_2}) \\ &= F(\chi_{M_1}) + F(\chi_{M_2}) = \mu(M_1) + \mu(M_2). \end{aligned}$$

Thus $\mu \in \mathcal{M}_{fa}(\mathbb{N})$. Moreover, if squares Q_j and points $z_j \in Q_j$ are chosen as above and if we again set $M_j = \{n \in \mathbb{N} : x_n \in Q_j\}$, then $\left\| x - \sum z_j \chi_{M_j} \right\|$ is bounded by the maximal diameter of the Q_j's, so this will go to zero if we again consider a sequence of subdivisions becoming arbitrarily fine. It follows that

$$F(x) = \lim F\left(\sum z_j \chi_{M_j} \right) = \lim \sum z_j \mu(M_j) = R - \int_{\mathbb{N}} x_n \, d\mu(n).$$

Conclusion: Every $F \in (\ell^\infty)^*$ can be represented as a Radon integral. Conversely, one can show that every $\mu \in \mathcal{M}_{fa}(\mathbb{N})$ generates a functional F_μ on ℓ^∞ by Radon integration:

$$F_\mu(x) = R - \int_{\mathbb{N}} x_n \, d\mu(n)$$

(The boundedness of F_μ requires some work; Exercise 4.19 below should help to clarify things.)

We obtain a bijection $\mathcal{M}_{fa}(\mathbb{N}) \to (\ell^\infty)^*$, $\mu \mapsto F_\mu$, and, as in the previous examples, it's now a relatively easy matter to check that this actually sets up an isometric isomorphism between Banach spaces if we endow $\mathcal{M}_{fa}(\mathbb{N})$ with the natural vectors space structure ($(\mu+\nu)(M) := \mu(M) + \nu(N)$ etc.) and the norm

$$\|\mu\| = \sup_{\|x\|=1} \left| R - \int_{\mathbb{N}} x_n \, d\mu(n) \right|.$$

I'll leave the details of these final steps to the reader.

Exercise 4.19. Show that $\|\mu\| = \sup \sum |\mu(M_j)|$, where the supremum is over all partitions of $\mathbb{N}$ into finitely many sets $M_1, \ldots, M_N$. Moreover,

$$\sum_{n=1}^{\infty} |\mu(\{n\})| \le \|\mu\|;$$

can you also show that strict inequality is possible? (Exercise 4.21 might be helpful here.)

Exercise 4.20. Show that $\ell^1 \subseteq \mathcal{M}_{fa}(\mathbb{N})$ in the sense that if $y \in \ell^1$, then $\mu(M) = \sum_{n\in M} y_n$ defines a bounded, finitely additive set function and $\|\mu\| = \|y\|_1$. Show that in fact these μ's are exactly the (complex) measures on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

Remark: Since X can be identified with a subspace of X^{**} for any Banach space X and since $\ell^\infty = (\ell^1)^*$, we knew right away that $\ell^1 \subseteq (\ell^\infty)^*$, provided this is suitably interpreted.

Exercise 4.21. The fact that $\ell^1 \subsetneqq (\ell^\infty)^*$ can also be seen more directly, without giving a description of $(\ell^\infty)^*$, as follows:
(a) Show that every $y \in \ell^1$ generates a functional $F_y \in (\ell^\infty)^*$ by letting

$$F_y(x) = \sum_{n=1}^{\infty} y_n x_n \qquad (x \in \ell^\infty). \tag{4.9}$$

(b) Show that not every $F \in (\ell^\infty)^*$ is of this form, by using the Hahn-Banach Theorem. More specifically, choose a subspace $Y \subseteq \ell^\infty$ and define a bounded functional F_0 on Y in such a way that no extension F of F_0 can be of the form (4.9). (This is an uncomplicated argument if done efficiently; it all depends on a smart choice of Y and F_0.)

Example 4.6. I'll quickly report on the spaces $L^p(X, \mu)$ here. The situation is similar to the discussion above; see Examples 4.3, 4.5. If $1 \le p < \infty$, then $(L^p)^* = L^q$, where $1/p + 1/q = 1$. This holds in complete generality for $1 < p < \infty$, but if $p = 1$, then we need the additional hypothesis that μ is σ-finite (which means that X can be written as a countable union of sets of finite measure). Again, this is an abbreviated way of stating the result; it really involves an identification of Banach spaces: the function $f \in L^q$ is identified with the functional $F_f \in (L^p)^*$ defined by $F_f(g) = \int_X fg\, d\mu$.

$(L^\infty)^*$ is again a complicated space that can be described as a space of finitely additive set functions, but this description is only moderately useful. In particular, except in special cases, $(L^\infty)^*$ is (much) bigger than L^1. In fact, for example $L^1(\mathbb{R}, m)$ is not the dual space of *any* Banach space: there is no Banach space X for which X^* is isometrically isomorphic to $L^1(\mathbb{R}, m)$!

Exercise 4.22. What is wrong with the following sketch of a "proof" that $(\ell^\infty)^* = \ell^1$:

Follow the strategy from Example 4.3. Obviously, if $y \in \ell^1$, then $F_y \in (\ell^\infty)^*$, if F_y is defined as in (4.7). Conversely, given an $F \in (\ell^\infty)^*$, let $y_n = F(e_n)$. Define $x \in \ell^\infty$ by

$$x_n = \begin{cases} \frac{|y_n|}{y_n} & n \le N, y_n \neq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Then $\|x\|_\infty \le 1$, so $F(x) = \sum_{n=1}^N |y_n| \le \|F\|$, and it follows that $y \in \ell^1$. By construction, $F = F_y$.

Exercise 4.23. Let X be a Banach space. Show that every weakly convergent sequence is bounded: If $x_n, x \in X$, $F(x_n) \to F(x)$ for all $F \in X^*$, then $\sup \|x_n\| < \infty$.

Hint: Think of the x_n as elements of the bidual $X^{**} \supseteq X$ and apply the uniform boundedness principle.

Exercise 4.24. (a) Show that $e_n \xrightarrow{w} 0$ in ℓ^2.
(b) Construct a sequence f_n with similar properties in $C[0,1]$: we want $\|f_n\| = 1$, $f_n \xrightarrow{w} 0$.