2. Banach spaces

Let X be a complex vector space. So the elements of X ("vectors") can be added and multiplied by complex numbers ("scalars"), and these operations obey the usual algebraic rules.

Definition 2.1. A map $\|\cdot\| : X \to [0,\infty)$ is called a norm (on X) if it has the following properties for arbitrary $x, y \in X$, $c \in \mathbb{C}$:

(1) $\|x\| = 0 \iff x = 0$
(2) $\|cx\| = |c|\,\|x\|$
(3) $\|x + y\| \le \|x\| + \|y\|$

We may interpret a given norm as assigning a length to a vector. Property (3) is again called the *triangle inequality.* It has a similar interpretation as in the case of a metric space. A vector space with a norm defined on it is called a *normed space.*

If $(X, \|\cdot\|)$ is a normed space, then $d(x, y) := \|x-y\|$ defines a metric on X.

Exercise 2.1. Prove this remark.

Therefore, all concepts and results from Chapter 1 apply to normed spaces also. In particular, a norm generates a topology on X. We repeat here some of the most basic notions: A sequence $x_n \in X$ is said to converge to $x \in X$ if $\|x_n - x\| \to 0$ (note that these norms form a sequence of *numbers*, so it's clear how to interpret this latter convergence). We call x_n a Cauchy sequence if $\|x_m - x_n\| \to 0$ as $m, n \to \infty$. The open ball of radius $r > 0$ about $x \in X$ is defined as

$$B_r(x) = \{y \in X : \|y - x\| < r\}.$$

This set is indeed open in the topology mentioned above; more generally, an arbitrary set $U \subseteq X$ is open precisely if for every $x \in U$, there exists an $r = r(x) > 0$ so that $B_r(x) \subseteq U$. Finally, recall that a space is called complete if every Cauchy sequence converges. Complete normed spaces are particularly important; for easier reference, they get a special name:

Definition 2.2. A *Banach space* is a complete normed space.

The following basic properties of norms are relatively direct consequences of the definition, but they are extremely important when working on normed spaces.

Exercise 2.2. (a) Prove the *second triangle inequality:*

$$\big|\, \|x\| - \|y\| \,\big| \le \|x - y\|$$

(b) Prove that the norm is a continuous map $X \to \mathbb{R}$; put differently, if $x_n \to x$, then also $\|x_n\| \to \|x\|$.

Exercise 2.3. Prove that the vector space operations are continuous. In other words, if $x_n \to x$ and $y_n \to y$ (and $c \in \mathbb{C}$), then also $x_n + y_n \to x + y$ and $cx_n \to cx$.

Let's now collect some examples of Banach spaces. It turns out that most of the examples for metric spaces that we considered in Chapter 1 actually have a natural vector space structure and the metric comes from a norm.

Example 2.1. The simplest vector spaces are the finite-dimensional spaces. Every n-dimensional (complex) vector space is isomorphic to $\mathbb{C}^n$, so it will suffice to consider $X = \mathbb{C}^n$. We would like to define norms on this space, and we can in fact turn to Example 1.3 for inspiration. For $x = (x_1, \ldots, x_n) \in X$, let

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p}, \tag{2.1}$$

for $1 \leq p < \infty$, and

$$\|x\|_\infty = \max_{j=1,\ldots,n} |x_j|. \tag{2.2}$$

I claim that this defines a family of norms (one for each p, $1 \leq p \leq \infty$), but we will not prove this in this setting. Rather, we will right away prove a more general statement in Example 2.2 below. (Only the triangle inequality for $1 < p < \infty$ needs serious proof; everything else is fairly easy to check here anyway.)

Example 2.2. We now consider infinite-dimensional versions of the Banach spaces from the previous example. Instead of finite-dimensional vectors $(x_1, \ldots, x_n)$, we now want to work with infinite sequences $x = (x_1, x_2, \ldots)$, and we want to continue to use (2.1), (2.2), or at least something similar. We first of all introduce the maximal spaces on which these formulae seem to make sense. Let

$$\ell^p = \left\{ x = (x_n)_{n \geq 1} : \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}$$

(for $1 \leq p < \infty$) and

$$\ell^\infty = \left\{ x = (x_n)_{n \geq 1} : \sup_{n \geq 1} |x_n| < \infty \right\}.$$

Then, as expected, for $x \in \ell^p$, define

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} \qquad (p < \infty),$$
$$\|x\|_\infty = \sup_{n \geq 1} |x_n|.$$

Proposition 2.3. *ℓ^p is a Banach space for $1 \leq p \leq \infty$.*

Here, the algebraic operations on ℓ^p are defined in the obvious way: we perform them componentwise; for example, $x + y$ is the sequence whose nth element is $x_n + y_n$.

Proof. We will explicitly prove this only for $1 < p < \infty$; the cases $p = 1$, $p = \infty$ are easier and can be handled by direct arguments. First of all, we must check that ℓ^p is a vector space. Clearly, if $x \in \ell^p$ and $c \in \mathbb{C}$, then also $cx \in \ell^p$. Moreover, if $x, y \in \ell^p$, then, since $|x_n + y_n|^p \leq (2|x_n|)^p + (2|y_n|)^p$, we also have $x + y \in \ell^p$. So addition and multiplication by scalars can be defined on all of ℓ^p, and it is clear that the required algebraic laws hold because all calculations are performed on individual components, so we just inherit the usual rules from $\mathbb{C}$.

Next, we verify that $\| \cdot \|_p$ is a norm on ℓ^p. Properties (1), (2) from Definition 2.1 are obvious. The proof of the triangle inequality will depend on the following very important inequality:

Theorem 2.4 (Hölder's inequality)**.** *Let $x \in \ell^p$, $y \in \ell^q$, where p, q satisfy*

$$\frac{1}{p} + \frac{1}{q} = 1$$

($1/0 := \infty$, $1/\infty := 0$ in this context). Then $xy \in \ell^1$ and

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

Proof of Theorem 2.4. Again, we focus on the case $1 < p < \infty$; if $p = 1$ or $p = \infty$, an uncomplicated direct argument is available.

The function $\ln x$ is concave, that is, the graph lies above line segments connecting any two of its points (formally, this follows from the fact that $(\ln x)'' = -1/x^2 < 0$). In other words, if $a, b > 0$ and $0 \leq \alpha \leq 1$, then

$$\alpha \ln a + (1 - \alpha) \ln b \leq \ln \left(\alpha a + (1 - \alpha) b \right).$$

We apply the exponential function on both sides and obtain $a^\alpha b^{1-\alpha} \leq \alpha a + (1 - \alpha) b$. If we introduce the new variables c, d by writing $a = c^p$,

$b = d^q$, with $1/p = \alpha$ (so $1/q = 1-\alpha$), then this becomes

$$cd \le \frac{c^p}{p} + \frac{d^q}{q}. \tag{2.3}$$

This holds for all $c, d \ge 0$ (the original argument is valid only if $c, d > 0$, but of course (2.3) is trivially true if $c = 0$ or $d = 0$). In particular, we can use (2.3) with $c = |x_n|/\|x\|_p$, $d = |y_n|/\|y\|_q$ (at least if $\|x\|_p, \|y\|_q \ne 0$, but if that fails, then the claim is trivial anyway) and then sum over $n \ge 1$. This shows that

$$\sum_{n=1}^{\infty} \frac{|x_n y_n|}{\|x\|_p \|y\|_q} \le \sum_{n=1}^{\infty} \frac{|x_n|^p}{p\|x\|_p^p} + \sum_{n=1}^{\infty} \frac{|y_n|^q}{q\|y\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1,$$

so $xy \in \ell^1$, as claimed, and we obtain Hölder's inequality. □

We are now in a position to establish the triangle inequality on ℓ^p:

Theorem 2.5 (Minkowski's inequality = triangle inequality on ℓ^p)**.** *Let $x, y \in \ell^p$. Then $x + y \in \ell^p$ and*

$$\|x + y\|_p \le \|x\|_p + \|y\|_p.$$

Proof of Theorem 2.5. Again, we will discuss explicitly only the case $1 < p < \infty$. We already know that $x+y \in \ell^p$. Hölder's inequality with the given p (and thus $q = p/(p-1)$) shows that

$$\begin{aligned} \|x+y\|_p^p = \sum |x_n + y_n|^p &= \sum |x_n + y_n|\,|x_n + y_n|^{p-1} \\ &\le \sum |x_n|\,|x_n + y_n|^{p-1} + \sum |y_n|\,|x_n + y_n|^{p-1} \\ &\le (\|x\|_p + \|y\|_p)\,\|x+y\|_p^{p-1}. \end{aligned}$$

If $x+y \ne 0$, we can divide by $\|x+y\|_p^{p-1}$ to obtain the desired inequality, and if $x + y = 0$, then the claim is trivial. □

It remains to show that ℓ^p is complete. So let $x^{(n)} \in \ell^p$ be a Cauchy sequence (since the elements of ℓ^p are themselves sequences, we really have a sequence whose members are sequences; we use a *super*script to label the elements of the Cauchy sequence from $X = \ell^p$ to avoid confusion with the index labeling the components of a fixed element of ℓ^p). Clearly,

$$\left| x_j^{(m)} - x_j^{(m)} \right|^p \le \|x^{(m)} - x^{(n)}\|_p^p$$

for each fixed $j \ge 1$, so $\left(x_j^{(n)}\right)_{n\ge 1}$ is a Cauchy sequence of complex numbers. Now $\mathbb{C}$ is complete, so these sequences have limits in $\mathbb{C}$. Define

$$x_j = \lim_{n\to\infty} x_j^{(n)}.$$

I claim that $x = (x_j) \in \ell^p$ and $x^{(n)} \to x$ in the norm of ℓ^p. To verify that $x \in \ell^p$, we observe that for arbitrary $N \in \mathbb{N}$,

$$\sum_{j=1}^{N} |x_j|^p = \lim_{n\to\infty} \sum_{j=1}^{N} \left| x_j^{(n)} \right|^p \le \limsup_{n\to\infty} \|x^{(n)}\|^p.$$

Exercise 2.4. Let $x_n \in X$ be Cauchy sequence in a normed space X. Prove that x_n is bounded in the following sense: There exists $C > 0$ such that $\|x_n\| \le C$ for all $n \ge 1$.

Exercise 2.4 now shows that

$$\sum_{j=1}^{N} |x_j|^p \le C$$

for some fixed, N independent constant C, so $x \in \ell^p$, as required.

It remains to show that $\|x^{(n)} - x\|_p \to 0$. Let $\epsilon > 0$ be given and pick $N_0 \in \mathbb{N}$ so large that $\|x^{(n)} - x^{(m)}\| < \epsilon$ if $m, n \ge N_0$ (this is possible because $x^{(n)}$ is a Cauchy sequence). Then, for fixed $N \in \mathbb{N}$, we have

$$\sum_{j=1}^{N} \left| x_j^{(n)} - x_j \right|^p = \lim_{m\to\infty} \sum_{j=1}^{N} \left| x_j^{(n)} - x_j^{(m)} \right|^p \le \epsilon$$

if $n \ge N_0$. Since $N \in \mathbb{N}$ was arbitrary, it also follows that $\|x^{(n)} - x\|_p^p \le \epsilon$ for $n \ge N_0$. $\square$

Similar spaces can be defined for arbitrary index sets I instead of $\mathbb{N}$. For example, by definition, the elements of $\ell^p(I)$ are complex valued functions $x : I \to \mathbb{C}$ with

$$\sum_{j\in I} |x_j|^p < \infty. \tag{2.4}$$

If I is uncountable, this sum needs interpretation. We can do this by hand, as follows: (2.4) means that $x_j \ne 0$ only for countably many $j \in I$, and the corresponding sum is finite. Equivalently, but more elegantly, we can also use the counting measure on I and interpret the sum as an integral.

If we want to emphasize the fact that we're using $\mathbb{N}$ as the index set, we can also denote the spaces discussed above by $\ell^p(\mathbb{N})$. When no confusion has to be feared, we will usually prefer the shorter notation ℓ^p. We can also consider finite index sets $I = \{1, 2, \ldots, n\}$. We have $\ell^p(\{1, 2, \ldots, n\}) = \mathbb{C}^n$ as a set, and the norms on these spaces are the ones that were already introduced in Example 2.1 above.

Example 2.3. Two more spaces of sequences are in common use. In both cases, the index set is usually $\mathbb{N}$ (or sometimes $\mathbb{Z}$). Put

$$c = \left\{ x : \lim_{n\to\infty} x_n \text{ exists} \right\},$$
$$c_0 = \left\{ x : \lim_{n\to\infty} x_n = 0 \right\}.$$

It is clear that $c_0 \subseteq c \subseteq \ell^\infty$. In fact, more is true: the smaller spaces are (algebraic linear) subspaces of the bigger spaces. On c and c_0, we also use the norm $\|\cdot\|_\infty$ (as on the big space ℓ^∞).

Proposition 2.6. *c and c_0 are Banach spaces.*

Proof. We can make use of the observation made above, that $c_0 \subseteq c \subseteq \ell^\infty$ and then refer to the following fact:

Proposition 2.7. *Let $(X, \|\cdot\|)$ be a Banach space, and let $Y \subseteq X$. Then $(Y, \|\cdot\|)$ is a Banach space if and only if Y is a closed (linear) subspace of X.*

Exercise 2.5. Prove Proposition 2.7. Recall that on metric (and thus also normed and Banach) spaces, you can use sequences to characterize topological notions. So a subset is closed precisely if all limits of convergent sequences from the set lie in the set again.

Since c and c_0 are obviously (linear) subspaces of ℓ^∞, we now only need to show that these spaces are closed subsets of ℓ^∞.

Exercise 2.6. Complete the proof of Proposition 2.6 along these lines.

□

Example 2.4. *Function spaces* provide another very important class of Banach spaces. The discussion is in large parts analogous to our treatment of sequence spaces (Examples 2.2, 2.3); sometimes, sequence spaces are somewhat more convenient to deal with and, as we will see in a moment, they can actually be interpreted as function spaces of a particular type.

Let $(X, \mathcal{M}, \mu)$ be a measure space (with a positive measure μ). The discussion is most conveniently done in this completely general setting, but if you prefer a more concrete example, you could think of $X = \mathbb{R}^n$ with Lebesgue measure, as what is probably the most important special case.

If we recall what we did above, then it seems natural to introduce (for $1 \le p < \infty$)

$$\mathcal{L}^p(X, \mu) = \left\{ f : X \to \mathbb{C} : f \text{ measurable}, \int_X |f(x)|^p \, d\mu(x) < \infty \right\}.$$

Note that this set also depends on the σ-algebra $\mathcal{M}$, but this dependence is not made explicit in the notation. We would then like to define

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

This, however, does not give a norm in general because $\|f\|_p = 0$ precisely if $f = 0$ almost everywhere, so usually there will be functions of zero "norm" that are not identically equal to zero. Fortunately, there is an easy fix for this problem: we simply identify functions that agree almost everywhere. More formally, we introduce an equivalence relation on $\mathcal{L}^p$, as follows:

$$f \sim g \iff f(x) = g(x) \quad \text{for } \mu\text{-almost every } x \in X$$

We then let L^p be the set of equivalence classes:

$$L^p(X, \mu) = \{(f) : f \in \mathcal{L}^p(X, \mu)\} ,$$

where $(f) = \{g \in \mathcal{L}^p : g \sim f\}$. We obtain a vector space structure on L^p in the obvious way; for example, $(f) + (g) := (f + g)$ (it needs to be checked here that the equivalence class on the right-hand side is independent of the choice of representatives f, g, but this is obvious from the definitions). Moreover, we can put

$$\|(f)\|_p := \|f\|_p;$$

again, it doesn't matter which function from (f) we take on the right-hand side, so this is well defined.

In the same spirit ("ignore what happens on null sets"), we define

$$\mathcal{L}^\infty(X, \mu) = \{f : X \to \mathbb{C} : f \text{ essentially bounded}\}.$$

A function f is called *essentially bounded* if there is a null set $N \subseteq X$ such that $|f(x)| \le C$ for $x \in X \setminus N$. Such a C is called an *essential bound.* If f is essentially bounded, its *essential supremum* is defined as the best essential bound:

$$\begin{aligned} \text{ess sup } |f(x)| &= \inf_{N:\mu(N)=0} \sup_{x \in X \setminus N} |f(x)| \\ &= \inf\{C \ge 0 : \mu(\{x \in X : |f(x)| > C\}) = 0\} \end{aligned}$$

Exercise 2.7. (a) Prove that both formulae give the same result.
(b) Prove that ess sup $|f|$ is itself an essential bound: $|f| \le$ ess sup $|f|$ almost everywhere.

Finally, we again let

$$L^\infty = \{(f) : f \in \mathcal{L}^\infty\} ,$$

and we put

$$\|(f)\|_\infty = \text{ess sup } |f(x)|.$$

Strictly speaking, the elements of the spaces L^p are not functions, but equivalence classes of functions. Sometimes, it is important to keep this distinction in mind; for example, it doesn't make sense to talk about $f(0)$ for an $(f) \in L^1(\mathbb{R}, m)$, say, because $m(\{0\}) = 0$, so we can change f at $x = 0$ without leaving the equivalence class (f). However, for most purposes, no damage is done if, for convenience and as a figure of speech, we simply refer to the elements of L^p as "functions" anyway (as in "let f be a function from L^1", rather than the pedantic and clumsy "let F be an element of L^1 and pick a function $f \in \mathcal{L}^1$ that represents the equivalence class F"). This convention is in universal use (it is similar to, say, "right lane must exit").

Proposition 2.8. *$L^p(X, \mu)$ is a Banach space for $1 \le p \le \infty$.*

We will not give the complete proof of this because the discussion is reasonably close to our previous treatment of ℓ^p. Again, the two main issues are the triangle inequality and completeness. The proof of the triangle inequality follows the pattern of the above proof very closely. To establish completeness, we (unsurprisingly) need facts from the theory of the Lebesgue integral, so this gives us a good opportunity to review some of these tools. We will give this proof only for $p = 1$ ($1 < p < \infty$ is similar, and $p = \infty$ can again be handled by a rather direct argument).

So let $f_n \in L^1$ be a Cauchy sequence. Pick a subsequence $n_k \to \infty$ so that $\|f_{n_{k+1}} - f_{n_k}\| < 2^{-k}$.

Exercise 2.8. Prove that n_k's with these properties can indeed be found.

Let

$$S_j(x) = \sum_{k=1}^{j} \left| f_{n_{k+1}}(x) - f_{n_k}(x) \right| .$$

Then S_j is measurable, non-negative, and $S_{j+1} \ge S_j$. So, if we let $S(x) = \lim_{j\to\infty} S_j(x) \in [0, \infty]$, then the Monotone Convergence Theorem shows that

$$\int_X S \, d\mu = \lim_{j\to\infty} \int_X S_j \, d\mu = \lim_{j\to\infty} \sum_{k=1}^{j} \int_X \left| f_{n_{k+1}} - f_{n_k} \right| \, d\mu$$

$$= \lim_{j\to\infty} \sum_{k=1}^{j} \left\| f_{n_{k+1}} - f_{n_k} \right\| < \sum_{k=1}^{\infty} 2^{-k} = 1.$$

In particular, $S \in L^1$, and this implies that $S < \infty$ almost everywhere.

The same conclusion can be obtained from Fatou's Lemma; let us do this too, for review purposes:

$$\int_X S\, d\mu = \int_X \lim_{j\to\infty} S_j\, d\mu = \int_X \liminf_{j\to\infty} S_j\, d\mu \le \liminf_{j\to\infty} \int_X S_j\, d\mu$$

We can conclude the argument as in the preceding paragraph, and we again see that $\int S < 1$, so $S < \infty$ almost everywhere.

For almost every $x \in X$, we can define

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

indeed, we just verified that this series actually converges *absolutely* for almost every $x \in X$. Moreover, the sum is telescoping, so in fact

$$f(x) = \lim_{j\to\infty} f_{n_j}(x)$$

for a.e. x. Also,

$$\left| f(x) - f_{n_j}(x) \right| \le \sum_{k=j}^{\infty} \left| f_{n_{k+1}}(x) - f_{n_k}(x) \right|.$$

Since this latter sum is dominated by $S \in L^1$, this shows, first of all, that $|f - f_{n_j}| \in L^1$ and thus also $f \in L^1$ (because $|f| \le |f_{n_j}| + |f - f_{n_j}|$). Moreover, the functions $|f - f_{n_j}|$ satisfy the hypotheses of Dominated Convergence, and thus

$$\lim_{j\to\infty} \int_X \left| f - f_{n_j} \right| d\mu = 0.$$

To summarize: given the Cauchy sequence $f_n \in L^1$, we have constructed a function $f \in L^1$, and $\|f_{n_j} - f\| \to 0$. This is almost what we set out to prove. For the final step, we can refer to the following general fact.

Exercise 2.9. Let x_n be a Cauchy sequence from a metric space Y. Suppose that $x_{n_j} \to x$ for some subsequence (and some $x \in Y$). Prove that then in fact $x_n \to x$.

We also saw in this proof that $f_{n_j} \to f$ pointwise almost everywhere. This is an extremely useful fact, so it's a good idea to state it again (for general p).

Corollary 2.9. *If $\|f_n - f\|_p \to 0$, then there exists a subsequence f_{n_j} that converges to f pointwise almost everywhere.*

Exercise 2.10. Give a (short) direct argument for the case $p = \infty$. Show that in this case, it is not necessary to pass to a subsequence.

If I is an arbitrary set (the case $I = \mathbb{N}$ is of particular interest here), $\mathcal{M} = \mathcal{P}(I)$ and μ is the counting measure on I (so $\mu(A)$ equals the number of elements of A), then $L^p(I, \mu)$ is the space $\ell^p(I)$ that was discussed earlier, in Example 2.2. Note that on this measure space, the only null set is the empty set, so there's no difference between $\mathcal{L}^p$ and L^p here.

Example 2.5. Our final example can perhaps be viewed as a mere variant of L^∞, but this space will become very important for us later on. We start out with a compact Hausdorff space K. A popular choice would be $K = [a, b]$, with the usual topology, but the general case will also be needed. We now consider

$$C(K) = \{f : K \to \mathbb{C} : f \text{ continuous } \},$$

with the norm

$$\|f\| = \|f\|_\infty = \max_{x \in K} |f(x)|.$$

The maximum exists because $|f|(K)$, being a continuous image of a compact space, is a compact subset of $\mathbb{R}$. As anticipated, we then have the following:

Proposition 2.10. *$\|\cdot\|_\infty$ is a norm on $C(K)$, and $C(K)$ with this norm is a Banach space.*

The proof is very similar to the corresponding discussion of L^∞; I don't want to discuss it in detail here. In fact, if there is a measure on K that gives positive weight to all non-empty open sets (such as Lebesgue measure on $[a, b]$), then $C(K)$ can be thought of as a subspace of L^∞.

Exercise 2.11. Can you imagine why we want the measure to give positive weight to open sets?

Hint: Note that the elements of $C(K)$ are genuine functions, while the elements of $L^\infty(K, \mu)$ were defined as equivalence classes of functions, so if we want to think of $C(K)$ as a subset of L^∞, we need a way to identify continuous functions with equivalence classes.

Exercise 2.12. Prove that $C(K)$ is complete.

In the sequel, we will be interested mainly in linear maps between Banach spaces (and not so much in the spaces themselves). More generally, let X, Y be normed spaces. Recall that a map $A : X \to Y$ is called *linear* if $A(x_1 + x_2) = Ax_1 + Ax_2$ and $A(cx) = cAx$. In functional analysis, we usually refer to linear maps as (linear) *operators*. The *null* *space* (or *kernel*) and the *range* (or *image*) of an operator A are defined as follows:

$$N(A) = \{x \in X : Ax = 0\},$$
$$R(A) = \{Ax : x \in X\}$$

Theorem 2.11. *Let $A : X \to Y$ be a linear operator. Then the following are equivalent:*
(a) A is continuous (everywhere);
(b) A is continuous at $x = 0$;
(c) A is bounded: *There exists a constant $C \geq 0$ such that $\|Ax\| \leq C\|x\|$ for all $x \in X$.*

Proof. $(a) \implies (b)$: This is trivial.

$(b) \implies (c)$: Suppose that A was not bounded. Then we can find, for every $n \in \mathbb{N}$, a vector $x_n \in X$ with $\|Ax_n\| > n\|x_n\|$. Let $y_n = (1/(n\|x_n\|))x_n$. Then $\|y_n\| = 1/n$, so $y_n \to 0$, but $\|Ay_n\| > 1$, so Ay_n can not go to the zero vector, contradicting (b).

$(c) \implies (a)$: Suppose that $x_n \to x$. We want to show that then also $Ax_n \to Ax$, and indeed this follows immediately from the linearity and boundedness of A:

$$\|Ax_n - Ax\| = \|A(x_n - x)\| \leq C\|x_n - x\| \to 0$$

□

Given two normed spaces X, Y, we introduce the space $B(X, Y)$ of bounded (or continuous) linear operators from X to Y. The special case $X = Y$ is of particular interest; in this case, we usually write $B(X)$ instead of $B(X, X)$.

$B(X, Y)$ becomes a vector space if addition and multiplication by scalars are defined in the obvious way (for example, $(A+B)x := Ax + Bx$). We can go further and also introduce a norm on $B(X, Y)$, as follows:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Since A is assumed to be bounded here, the supremum will be finite. We call $\|A\|$ the *operator norm* of A (that this is a norm will be seen in Theorem 2.12 below).

There are a number of ways to compute $\|A\|$.

Exercise 2.13. Prove the following formulae for $\|A\|$ (for $A \in B(X,Y)$):

$$\begin{aligned}\|A\| &= \inf\{C \geq 0 : \|Ax\| \leq C\|x\| \text{ for all } x \in X\}\\ &= \min\{C \geq 0 : \|Ax\| \leq C\|x\| \text{ for all } x \in X\}\\ \|A\| &= \sup_{\|x\|=1} \|Ax\|\end{aligned}$$

In particular, this shows that $\|Ax\| \leq \|A\|\,\|x\|$, and $\|A\|$ is the smallest constant for which this inequality holds.

Exercise 2.14. However, it is not necessarily true that $\|A\| = \max_{\|x\|=1} \|Ax\|$. Provide an example of such an operator A.

Suggestion: $X = Y = c_0$ (or ℓ^1 if you prefer, this also works very well), and define $(Ax)_n = a_n x_n$, where a_n is a suitably chosen bounded sequence.

Theorem 2.12. *Let X, Y be normed spaces.*

(a) $B(X,Y)$ with the operator norm is a normed space.

(b) If Y is a Banach space, then $B(X,Y)$ (with the operator norm) is a Banach space.

The special case $Y = \mathbb{C}$ (recall that this is a Banach space if we use the absolute value as the norm) is particularly important. We use the alternative notation $X^* = B(X, \mathbb{C})$, and we call the elements of X^* (continuous, linear) *functionals*. X^* itself is called the *dual space* (or just the dual) of X.

This must not be confused with the dual space from linear algebra, which is defined as the set of all linear maps from the original vector space back to its base field (considered as a vector space also). This is of limited use in functional analysis. The (topological) dual X^* contains only the *continuous* linear maps back to the base field; it is usually much smaller than the algebraic dual described above.

Proof. (a) We observed earlier that $B(X,Y)$ is a vector space, so we need to check that the operator norm satisfies the properties from Definition 2.1. First of all, we will have $\|A\| = 0$ precisely if $Ax = 0$ for all $x \in X$, that is, precisely if A is the zero map or, put differently, $A = 0$ in $B(X,Y)$. Next, if $c \in \mathbb{C}$ and $A \in B(X,Y)$, then

$$\|cA\| = \sup_{\|x\|=1} \|cAx\| = \sup_{\|x\|=1} |c|\|Ax\| = |c|\|A\|.$$

A similar calculation establishes the third property from Definition 2.1:

$$\|A + B\| = \sup_{\|x\|=1} \|(A+B)x\| \leq \sup_{\|x\|=1} (\|Ax\| + \|Bx\|) \leq \|A\| + \|B\|$$

(b) Let A_n be a Cauchy sequence from $B(X,Y)$. We must show that A_n converges. Observe that for fixed x, $A_n x$ will be a Cauchy sequence in Y. Indeed,

$$\|A_m x - A_n x\| \le \|A_m - A_n\| \|x\|$$

can be made arbitrarily small by taking both m and n large enough. Since Y is now assumed to be complete, the limits $Ax := \lim_{n\to\infty} A_n x$ exist, and we can define a map A on X in this way. We first check that A is linear:

$$\begin{aligned} A(x_1 + x_2) &= \lim_{n\to\infty} A_n(x_1 + x_2) = \lim_{n\to\infty} (A_n x_1 + A_n x_2) \\ &= \lim_{n\to\infty} A_n x_1 + \lim_{n\to\infty} A_n x_1 = Ax_1 + Ax_2, \end{aligned}$$

and a similar (if anything, this is easier) argument shows that $A(cx) = cAx$.

A is also bounded because

$$\|Ax\| = \|\lim A_n x\| = \lim \|A_n x\| \le (\sup \|A_n\|) \|x\|;$$

the supremum is finite because $|\|A_m\| - \|A_n\|| \le \|A_m - A_n\|$, so $\|A_n\|$ forms a Cauchy sequence of real numbers and thus is convergent and, in particular, bounded. Notice also that we used the continuity of the norm for the second equality (see Exercise 2.2(b)).

Summing up: we have constructed a map A and confirmed that in fact $A \in B(X,Y)$. The final step will be to show that $A_n \to A$, with respect to the operator norm in $B(X,Y)$. Let $x \in X$, $\|x\| = 1$. Then, by the continuity of the norm again,

$$\|(A - A_n)x\| = \lim_{m\to\infty} \|(A_m - A_n)x\| \le \limsup_{m\to\infty} \|A_m - A_n\|.$$

Since x was arbitrary, it also follows that

$$\|A - A_n\| \le \limsup_{m\to\infty} \|A_m - A_n\|.$$

Since A_n is a Cauchy sequence, the $\limsup$ can be made arbitrarily small by taking n large enough. $\square$

There are discontinuous linear maps if the first space, X, is infinite-dimensional. We can then even take $Y = \mathbb{C}$. An abstract construction can be done as follows: Let $\{e_\alpha\}$ be an algebraic basis of X (that is, every $x \in X$ can be written in a unique way as a linear combination of (finitely many) e_α's). For arbitrary complex numbers c_α, there exists a linear map $A : X \to \mathbb{C}$ with $Ae_\alpha = c_\alpha \|e_\alpha\|$.

Exercise 2.15. This problem reviews the linear algebra fact needed here. Let V, W be vector spaces (over $\mathbb{C}$, say), and let $\{e_\alpha\}$ be a basis of V.

Show that for every collection of vectors $w_\alpha \in W$, there exists a unique linear map $A : V \to W$ with $Ae_\alpha = w_\alpha$ for all α.

Since $\|Ae_\alpha\|/\|e_\alpha\| = |c_\alpha|$, we see that A can not be bounded if $\sup_\alpha |c_\alpha| = \infty$.

On the other hand, if $\dim X < \infty$, then linear operators $A : X \to Y$ are always bounded. We will see this in a moment; before we do this, we introduce a new concept and prove a related result.

Definition 2.13. Two norms on a common space X are called *equivalent* if they generate the same topology.

This can be restated in a less abstract way:

Proposition 2.14. *The norms $\|\cdot\|_1$, $\|\cdot\|_2$ are equivalent if and only if there are constants $C_1, C_2 > 0$ such that*

$$C_1\|x\|_1 \le \|x\|_2 \le C_2\|x\|_1 \qquad \text{for all } x \in X. \tag{2.5}$$

Proof. Consider the identity as a map from $(X, \|\cdot\|_1)$ to $(X, \|\cdot\|_2)$. Clearly, this is bijective, and, by Theorem 2.11 this map and its inverse are continuous precisely if (2.5) holds. Put differently, (2.5) is equivalent to the identity map being a homeomorphism (a bijective continuous map with continuous inverse), and this holds if and only if $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ have the same topology. □

Exercise 2.16. (a) Let $\|\cdot\|_1$, $\|\cdot\|_2$ be equivalent norms on X. Show that then $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are either both complete or both not complete.

(b) Construct a metric d on $\mathbb{R}$ that produces the usual topology, but $(\mathbb{R}, d)$ is not complete. (Since $(\mathbb{R}, |\cdot|)$ has the same topology and is complete, this shows that the analog of (a) for metric spaces is false.)

Theorem 2.15. *Let X be a (complex) vector space with $\dim X < \infty$. Then all norms on X are equivalent.*

In particular, by combining Example 2.1 with Exercise 2.16, we see that finite-dimensional normed spaces are automatically complete and thus Banach spaces.

Proof. By fixing a basis on X, we may assume that $X = \mathbb{C}^n$. We will show that every norm on $\mathbb{C}^n$ is equivalent to $\|\cdot\|_1$. We will do this by verifying (2.5). So let $\|\cdot\|$ be a norm. Then, first of all,

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \le \sum_{j=1}^n |x_j| \, \|e_j\| \le \left(\max_{j=1,\ldots,n} \|e_j\| \right) \|x\|_1. \tag{2.6}$$

To obtain the other inequality, consider again the identity as a map from $(\mathbb{C}^n, \|\cdot\|_1)$ to $(\mathbb{C}^n, \|\cdot\|)$. As we have just seen in (2.6), this map is bounded, thus continuous. Since a norm always defines a continuous map, the composite map from $(\mathbb{C}^n, \|\cdot\|_1)$ to $\mathbb{R}$, $x \mapsto \|x\|$ is also continuous. Now $\{x \in \mathbb{C}^n : \|x\|_1 = 1\}$ is a compact subset of $\mathbb{C}^n$, with respect to the topology generated by $\|\cdot\|_1$ (which is just the usual topology on $\mathbb{C}^n$). Therefore, the image under our map, which is given by $\{\|x\| : \|x\|_1 = 1\}$ is a compact subset of $\mathbb{R}$, and it doesn't contain zero, so

$$\inf_{\|x\|_1=1} \|x\| = \min_{\|x\|_1=1} \|x\| =: c > 0,$$

and the homogeneity of norms now implies that $\|x\| \geq c\|x\|_1$ for all $x \in \mathbb{C}^n$, as required. □

Corollary 2.16. *Suppose that* $\dim X < \infty$*, and let* $A : X \to Y$ *be a linear operator. Then* A *is bounded.*

Proof. By Theorem 2.15, it suffices to discuss the case $X = \mathbb{C}^n$, equipped with the norm $\|\cdot\|_1$. As above, we estimate

$$\|Ax\| = \left\| A\left(\sum_{j=1}^n x_j e_j\right)\right\| \leq \sum_{j=1}^n |x_j|\,\|Ae_j\| \leq \left(\max_{j=1,\ldots,n} \|Ae_j\|\right)\|x\|_1.$$

□

We conclude this chapter by discussing sums and quotients of Banach spaces. Let $X_1, \ldots, X_n$ be Banach spaces. We form their direct sum (as vector spaces). More precisely, we introduce

$$X = \{(x_1, \ldots, x_n) : x_j \in X_j\};$$

this becomes a vector space in the obvious way: the algebraic operations are defined componentwise. Of course, we want more: We want to introduce a norm on X that makes X a Banach space, too. This can be done in several ways; for example, the following works.

Theorem 2.17. $\|x\| = \sum_{j=1}^n \|x_j\|_j$ *defines a norm on* X*, and with this norm,* X *is a Banach space.*

Exercise 2.17. Prove Theorem 2.17.

We will denote this new Banach space by $X = \bigoplus_{j=1}^n X_j$.

Moving on to quotients now, we consider a Banach space X and a closed subspace $M \subseteq X$.

Exercise 2.18. (a) In general, subspaces need not be closed. Give an example of a dense subspace $M \subseteq \ell^1$, $M \neq \ell^1$ (in other words, we want $\overline{M} = \ell^1$, $M \neq \ell^1$; in particular, such an M is definitely not closed).

(b) What can you say about open subspaces of a normed space?

Exercise 2.19. However, show that *finite-dimensional* subspaces of a normed space are always closed.

Suggestion: Use Theorem 2.15.

As a vector space, we define the quotient X/M as the set of equivalence classes (x), $x \in X$, where $x, y \in X$ are equivalent if $x - y \in M$. So $(x) = x + M = \{x + m : m \in M\}$, and to obtain a vector space structure on X/M, we do all calculations with representatives. In other words, $(x) + (y) := (x + y)$, $c(x) := (cx)$, and this is well defined, because the right-hand sides are independent of the choice of representatives x, y.

Theorem 2.18. *Let X be a Banach space, and let $M \subseteq X$ be a closed subspace. Then $\|(x)\| := \inf_{y \in (x)} \|y\|$ defines a norm on X/M, and X/M with this norm is a Banach space.*

Proof. First of all, we must check the conditions from Definition 2.1. We have $\|(x)\| = 0$ precisely if there are $m_n \in M$ such that $\|x - m_n\| \to 0$. This holds if and only if $x \in \overline{M}$, but M is assumed to be closed, so $\|(x)\| = 0$ if and only if $x \in M$, that is, if and only if x represents the zero vector from X/M (equivalently, $(x) = (0)$).

If $c \in \mathbb{C}$, $c \neq 0$, then

$$\begin{aligned} \|c(x)\| &= \|(cx)\| = \inf_{m \in M} \|cx - m\| = \inf_{m \in M} \|cx - cm\| \\ &= |c| \inf_{m \in M} \|x - m\| = |c| \, \|(x)\|. \end{aligned}$$

If $c = 0$, then this identity ($\|0(x)\| = 0\|(x)\|$) is also true and in fact trivial.

The triangle inequality follows from a similar calculation:

$$\begin{aligned} \|(x) + (y)\| &= \|(x + y)\| = \inf_{m \in M} \|x + y - m\| = \inf_{m,n \in M} \|x + y - m - n\| \\ &\leq \inf_{m,n \in M} (\|x - m\| + \|y - n\|) = \|(x)\| + \|(y)\| \end{aligned}$$

Finally, we show that X/M is complete. Let (x_n) be a Cauchy sequence. Pass to a subsequence such that $\|(x_{n_{j+1}}) - (x_{n_j})\| < 2^{-j}$ (compare Exercise 2.8). Since the quotient norm was defined as the infimum of the norms of the representatives, we can now also (inductively) find representatives (we may assume that these are the x_n's themselves) such that $\|x_{n_{j+1}} - x_{n_j}\| < 2^{-j}$. Since $\sum 2^{-j} < \infty$, it follows that x_{n_j} is a Cauchy sequence in X, so $x = \lim_{j \to \infty} x_{n_j}$ exists. But then we also have

$$\|(x) - (x_{n_j})\| \leq \|x - x_{n_j}\| \to 0,$$

so a subsequence of the original Cauchy sequence (x_n) converges, and this forces the whole sequence to converge; see Exercise 2.9. $\square$

Exercise 2.20. Let X be a normed space, and define

$$\overline{B}_r(x) = \{y \in X : \|x - y\| \le r\}.$$

Show that $\overline{B}_r(x) = \overline{B_r(x)}$, where the right-hand side is the closure of the (open) ball $B_r(x)$. (Compare Exercise 1.16, which discussed the analogous problem on metric spaces.)

Exercise 2.21. Call a subset B of a Banach space X *bounded* if there exists $C \ge 0$ such that $\|x\| \le C$ for all $x \in B$.
(a) Show that if $K \subseteq X$ is compact, then K is closed and bounded.
(b) Consider $X = \ell^\infty$, $B = \overline{B}_1(0) = \{x \in \ell^\infty : \|x\| \le 1\}$. Show that B is closed and bounded, but not compact (in fact, the closed unit ball of an infinite-dimensional Banach space is never compact).

Exercise 2.22. If x_n are elements of a normed space X, we define, as usual, the *series* $\sum_{n=1}^\infty x_n$ as the limit as $N \to \infty$ of the partial sums $S_N = \sum_{n=1}^N x_n$, if this limit exists (of course, this limit needs to be taken with respect to the norm, so $S = \sum_{n=1}^\infty x_j$ means that $\|S - S_N\| \to 0$). Otherwise, the series is said to be divergent. Call a series *absolutely convergent* if $\sum_{n=1}^\infty \|x_n\| < \infty$.

Prove that a normed space is complete if and only if every absolutely convergent series converges.

Exercise 2.23. Find the operator norm of the identity map $(x \mapsto x)$ as an operator
(a) from $(\mathbb{C}^n, \|\cdot\|_1)$ to $(\mathbb{C}^n, \|\cdot\|_2)$;
(b) from $(\mathbb{C}^n, \|\cdot\|_2)$ to $(\mathbb{C}^n, \|\cdot\|_1)$.

Exercise 2.24. Find the operator norms of the following operators on $\ell^2(\mathbb{Z})$. In particular, prove that these operators are bounded.

$$(Ax)_n = x_{n+1} + x_{n-1}, \qquad (Bx)_n = \frac{n^2}{n^2+1}\, x_n$$

Exercise 2.25. Let X, Y, Z be Banach spaces, and let $S \in B(X, Y)$, $T \in B(Y, Z)$. Show that the composition TS lies in $B(X, Z)$ and $\|TS\| \le \|T\|\,\|S\|$. Show also that strict inequality is possible here. Give an example; as always, it's sound strategy to try to keep this as simple as possible. Here, finite-dimensional spaces X, Y, Z should suffice.

Exercise 2.26. Let X, Y be Banach spaces and let M be a dense subspace of X (there is nothing unusual about that on infinite-dimensional spaces; compare Exercise 2.18). Prove the following: Every $A_0 \in B(M,Y)$ has a unique continuous extension to X. Moreover, if we call this extension A, then $A \in B(X,Y)$ (by construction, A is continuous, so we're now claiming that A is also linear), and $\|A\| = \|A_0\|$.

Exercise 2.27. (a) Let $A \in B(X,Y)$. Prove that $N(A)$ is a closed subspace of X.
(b) Now assume that F is a linear functional on X, that is, a linear map $F : X \to \mathbb{C}$. Show that F is continuous if $N(F)$ is closed (so, for linear functionals, continuity is equivalent to $N(F)$ being closed).

Suggestion: Suppose F is not continuous, so that we can find $x_n \in X$ with $\|x_n\| = 1$ and $|F(x_n)| \geq n$, say. Also, fix another vector $z \notin N(F)$ (what if $N(F) = X$?). Use these data to construct a sequence $y_n \in N(F)$ that converges to a vector not from $N(F)$. (If this doesn't seem helpful, don't give up just yet, but try something else; the proof is quite short.)