**MA3422 2016–17**

# Chapter 2: Main theorems on Banach spaces

# Contents


2.1 The Baire category theorem . . . . . . . . . . . . . . . . . . . . . . . . 2
2.2 Open mapping theorem . . . . . . . . . . . . . . . . . . . . . . . . . 2
2.3 Direct sum of two normed spaces . . . . . . . . . . . . . . . . . . . . 3
2.4 Closed graph theorem . . . . . . . . . . . . . . . . . . . . . . . . . . 4
2.5 Uniform boundedness theorem . . . . . . . . . . . . . . . . . . . . . 6
2.6 Zorn's lemma . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10
2.7 Hahn-Banach Theorem . . . . . . . . . . . . . . . . . . . . . . . . . 15
2.8 Definition of reflexivity . . . . . . . . . . . . . . . . . . . . . . . . . 21

A number of the major results rely on completeness via the Baire category theorem and we already saw that theorem and one of its consequences (the open mapping theorem for Banach spaces) in MA3421.

Here they stated are again.

## 2.1 The Baire category theorem

**2.1.1 Definition.** A subset $S \subset X$ of a metric space $(X, d)$ (or topological space $(X, \mathscr{T})$ is called *nowhere dense* if the interior of its closure is empty, $(\bar{S})^\circ = \emptyset$.

A subset $E \subset X$ is called of *first category* if it is a countable union of nowhere dense subsets, that is, the union $E = \bigcup_{n=1}^\infty S_n$ of a sequence of nowhere dense sets $S_n$) $((\bar{S}_n)^\circ = \emptyset \forall n)$.

A subset $Y \subset X$ is called of *second category* if it fails to be of first category.

**2.1.2 Theorem** (Baire Category)**.** *Let $(X, d)$ be a complete metric space which is not empty. Then the whole space $S = X$ is of second category in itself.*

## 2.2 Open mapping theorem

**2.2.1 Theorem** (Open Mapping Theorem)**.** *Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be Banach spaces (over $\mathbb{K}$) and $T\colon E \to F$ a surjective bounded linear operator. Then there exists $\delta > 0$ so that*

$$T(B_E) \supseteq \delta B_F$$

*where $B_E = \{x \in E : \|x\|_E < 1\}$ and $B_F = \{y \in F : \|y\|_F < 1\}$ are the open unit balls of $E$ and $F$.*

*Moreover, if $U \subseteq E$ is open, then $T(U)$ is open (in $F$).*

## 2.3 Direct sum of two normed spaces

If we start with normed spaces $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$, then we can make $E \times F$ into a vector space by defining vector space operations as follows:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \quad (\text{for } (x_1, y_1), (x_2, y_2) \in E \times F)$$

$$\lambda(x, y) = (\lambda x, \lambda y) \quad (\text{for } \lambda \in \mathbb{K}, (x, y) \in E \times F).$$

One can check in a straightforward way that this make $E \times F$ into a vector space, and the usual notation for this is $E \oplus F$ (called the *direct sum* of $E$ and $F$).

Inside $E \oplus F$ there is a subspace

$$E \oplus \{0\} = \{(x, 0) : x \in E\}$$

that behaves just like $E$ (is isomorphic as a vector space to $E$) and another subspace $\{0\} \oplus F = \{(0, y) : y \in F\}$ that is a copy of $F$. Every $(x, y) \in E \oplus F$ can be expressed in a unique way as

$$(x, y) = (x, 0) + (0, y)$$

as a sum of an element of $E \oplus \{0\}$ and an element of $\{0\} \oplus F$.

In general if $V$ is a vector space and if $V_1, V_2 \subseteq V$ are vector subspaces such that each $v \in V$ can be expressed in a unique way as $v = v_1 + v_2$ with $v_1 \in V_1$, $v_2 \in V_2$, then we say that $V$ is the [concrete or internal] direct sum of $V_1$ and $V_2$. It is equivalent to $V = V_1 + V_2$ and $V_1 \cap V_2 = \{0\}$.

So far this is about vector spaces and so belongs to linear algebra.

To make $E \oplus F$ a normed space, we have several choices for norms. For $1 \leq p \leq \infty$ we can define a norm on $E \oplus F$ by

$$\|(x, y)\|_p = \|(\|x\|_E, \|y\|_F)\|_p = \begin{cases} (\|x\|_E^p + \|y\|_F^p)^{1/p} & p < \infty \\ \max(\|x\|_E, \|y\|_F) & p = \infty \end{cases}$$

(in terms of the $p$-norm of $\mathbb{R}^2$). It is not hard to check that these norms are all equivalent and that the topology arising from any of these norms is the product topology on $E \times F$.

It is common to write $E \oplus_p F$ for $E \oplus F$ with the norm $\|\cdot\|_p$.

We recall the product topology for the case of a product of two topological spaces. (Later we will return to arbitrary (infinite) products.)

**2.3.1 Definition.** If $X$ and $Y$ are topological spaces, then the product topology on $X \times Y$ is defined as follows:

> A subset $W \subseteq X \times Y$ is open if for each $(x_0, y_0) \in W$ there exist $U \subset X$ and $V \subset Y$ open with $(x_0, y_0) \in U \times V \subseteq W$.

It is easy to check that this is indeed a topology on the set $X \times Y$ (the set of ordered pairs $(x, y)$ with $x \in X$, $y \in Y$) and that the 'projection maps' given by $\pi_X : X \times Y \to X$, $\pi_X(x, y) = x$ and $\pi_Y : X \times Y \to Y$, $\pi_X(x, y) = y$ are both continuous (with the product topology on $X \times Y$).

## 2.4 Closed graph theorem

**2.4.1 Theorem** (Closed graph theorem)**.** *If $E, F$ are Banach spaces and $T : E \to F$ is a linear transformation, then $T$ is bounded if and only if its 'graph'*

$$\{(x, y) \in E \times F : y = Tx\}$$

*is a closed subset of $E \times F$ in the product topology.*

*Proof.* It is quite easy to check that the graph must be closed if $T$ is bounded (continuous). The product topology is a metric space topology as $E$ and $F$ are metric spaces, and in fact arises from a norm on $E \oplus F$. (See §2.3.) For definiteness we take the norm

$$\|(x, y)\|_1 = \|x\| + \|y\|$$

on $E \oplus F$, or we could write $E \oplus_1 F$. So to show the graph is closed, suppose we have a sequence $(x_n, y_n)$ in the graph that converges in $E \times F$ to some limit $(x, y)$. We claim that $(x, y)$ must be in the graph, that is that $y = Tx$.

What we know is that $y_n = Tx_n$ for each $n$ (since $(x_n, y_n)$ is in the graph). As $(x_n, y_n) \to (x, y)$ in the product topology, it follows that $x_n \to x$ and $y_n \to y$ (as $n \to \infty$). (This is rather easy as $\|x_n - x\| \leq \|(x_n, y_n) - (x, y)\|_1 = \|(x_n - x, y_n - y)\|_n \to 0$ as $n \to \infty$.) But $T$ is continuous and so $T(x_n) \to T(x)$. Since $T(x_n) = y_n$, that means $y_n \to T(x)$. As $y_n \to y$ also (because $\|y_n - y\| \leq \|(x_n, y_n) - (x, y)\|_1$) and in metric spaces sequences can have at most one limit, we have $T(x) = y$ and $(x, y)$ is in the graph — as we wanted to show.

Now assume that the graph is closed. Our proof that $T$ must be continuous relies on the open mapping theorem. We introduce a new norm on $E$ by

$$|||x||| = \|x\|_E + \|Tx\|_F.$$

It is not at all difficult to check that this is indeed a norm:

$$\begin{aligned}
\|\!|x_1 + x_2|\!\| &= \|x_1 + x_2\|_E + \|T(x_1 + x_2)\|_F \\
&= \|x_1 + x_2\|_E + \|T(x_1) + T(x_2)\|_F \\
&\leq \|x_1\|_E + \|x_2\|_E + \|T(x_1)\|_F + \|T(x_2)\|_F \\
&= \|\!|x_1|\!\| + \|\!|x_2|\!\| \\
\|\!|\lambda x|\!\| &= \|\lambda x\|_E + \|T(\lambda x)\|_F \\
&= \|\lambda x\|_E + \|\lambda T x\|_F \\
&= |\lambda|\|x\|_E + |\lambda|\|Tx\|_F \\
&= |\lambda|\|\!|x|\!\| \\
\|\!|x|\!\| = 0 &\Rightarrow \|x\|_E = 0 \\
&\Rightarrow x = 0
\end{aligned}$$

We consider the identity mapping $\mathrm{id}\colon (E, \|\!|\cdot|\!\|) \to (E, \|\cdot\|_E)$ which is clearly linear, bijective and bounded since

$$\|\mathrm{id}(x)\|_E = \|x\|_E \leq \|\!|x|\!\|.$$

In order to apply the open mapping theorem we need to know that $(E, \|\!|\cdot|\!\|)$ is a Banach space (that is complete) and in doing that the hypothesis that the graph is closed will be used. Once we have verified completeness, the open mapping theorem guarantees that the inverse map $\mathrm{id}^{-1}\colon (E, \|\cdot\|_E) \to (E, \|\!|\cdot|\!\|)$ is continuous. Therefore its is bounded, or there exists $M \geq 0$ so that

$$\|\!|x|\!\| \leq M\|x\|_E$$

holds for all $x \in E$. That implies $\|Tx\|_F \leq \|\!|x|\!\| \leq M\|x\|_E$ and so $T$ bounded as we need.

It remains to verify the completeness of $(E, \|\!|\cdot|\!\|)$. Take a Cauchy sequence $(x_n)_{n=1}^\infty$ in $E$ with respect to the metric arising from $\|\!|\cdot|\!\|$. Then for $\varepsilon > 0$ given and $n, m$ large enough

$$\|\!|x_n - x_m|\!\| < \varepsilon.$$

But that implies both $\|x_n - x_m\|_E < \varepsilon$ and $\|T(x_n) - T(x_m)\|_F < \varepsilon$. This means $(x_n)_{n=1}^\infty$ is a Cauchy sequence in $(E, \|\cdot\|_E)$ and $(T(x_n))_{n=1}^\infty$ is a Cauchy sequence in $(F, \|\cdot\|_F)$. As both $E$ and $F$ are complete there exists $x = \lim_{n\to\infty} x_n \in E$ and $y = \lim_{n\to\infty} T(x_n) \in F$. We have then that

$$(x_n, T(x_n)) \to (x, y)$$

in the norm $\|\cdot\|_1$ and so in the product topology of $E \times F$. But each $(x_n, T(x_n))$ is in the graph of $T$ and we are assuming that the graph is closed. So the limit $(x, y)$ is also in the graph, or $y = Tx$. It follows that

$$\lim_{n\to\infty} |||x_n - x||| = \lim_{n\to\infty} \|x_n - x\|_E + \|T(x_n) - Tx\|_F = \lim_{n\to\infty} \|x_n - x\|_E + \|y_n - y\|_F = 0$$

and that means $x_n \to x$ in $(E, |||\cdot|||)$. We have shown that each Cauchy sequence in $(E, |||\cdot|||)$ converges. So $(E, |||\cdot|||)$ is complete.

As noted above, the result follows then from the open mapping theorem. □

## 2.5 Uniform boundedness theorem

**2.5.1 Theorem** (Uniform boundedness principle)**.** *Let $E$ be a Banach space and $F$ a normed space. Let $T_i \colon E \to F$ be bounded linear operators for $i \in I =$ some index set. Assume*

$$\sup_{i\in I} \|T_i(x)\| < \infty \text{ for each } x \in E$$

*(which can be stated as the family of operators being pointwise uniformly bounded).*

*Then*

$$\sup_{i\in I} \|T_i\| < \infty$$

*(uniform boundedness in norm).*

*Proof.* The proof relies on the Baire category theorem in $E$.

Let

$$\begin{aligned} W_n &= \{x \in E : \sup_{i\in I} \|T_i x\| \leq n\} \\ &= \bigcap_{i\in I} \{x \in E : \|T_i x\| \leq n\}. \end{aligned}$$

Then $W_n$ is closed as it is the intersection of closed sets. ($\{x \in E : \|T_i x\| \leq n\}$ is the inverse image under $T_i$ of the closed ball $\{y \in F : \|y\| \leq n\}$, and $T_i$ is continuous. So the inverse image is closed.)

The hypothesis implies that each $x \in E$ belongs in $W_n$ for some $n$, or that $\bigcup_{n=1}^\infty W_n = E$. By the Baire category theorem, there is some $n$ where $W_n$ is not nowhere dense. Fix this $n$. As $W_n$ is closed, $W_n$ not nowhere dense means that $W_n$ has nonempty interior. So there is $x_0$ and $r > 0$ so that

$$B(x_0, r) = x_0 + rB_E \subseteq W_n.$$

Thus

$$\|T_i(x_0 + rx)\| = \|T_i(x_0) + rT_i(x)\| \leq n$$

holds for all $x \in B_E$. It follows that

$$r\|T_i(x)\| - \|T_i(x_0)\| \leq n \quad (\forall x \in B_E)$$

and so

$$\|T_i(x)\| \leq \frac{1}{r}\|T_i(x_0)\| + \frac{n}{r} \leq \frac{2n}{r} \quad (\forall x \in B_E)$$

since $x_0 \in W_n \Rightarrow \|T_i(x_0)\| \leq n$ for all $i \in I$. Take $\sup_{x \in B_E}$ to get

$$\|T_i\| \leq \frac{2n}{r}$$

and as this is true for each $i \in I$, we have $\sup_{i \in I} \|T_i\| \leq 2n/r < \infty$ as required. $\square$

*2.5.2 Remark.* Here is an application of the uniform boundedness principle.

**2.5.3 Theorem.** *There exists a continuous* $1$*-periodic function* $f$ *on* $[0,1]$ *such that the partial sums of its Fourier series do not converge at* $t = 0$*, that is such that*

$$\lim_{n \to \infty} (S_n f)(0)$$

*does not exist, where*

$$S_n f(t) = \sum_{j=-n}^{n} \hat{f}(j) e^{2\pi i j t}.$$

*is the* $n^{\text{th}}$ *(symmetric) partial sum of the Fourier series*

$$\sum_{n=-\infty}^{\infty} \hat{f}(n) e^{2\pi i n t}$$

*of* $f$ *and the Fourier coefficients* $\hat{f}(n)$ *are given by*

$$\hat{f}(n) = \langle f, e^{2\pi i n t} \rangle_{L^2[0,1]} = \int_0^1 f(t) e^{-2\pi i n t} \, dt$$

*2.5.4 Remark.* In MA3421 we discussed the Fourier series of functions $f \in L^1[0,1]$ and discussed the fact that $\lim_{n\to\infty} |\hat{f}(n)| = 0$, but that there are further restrictions on the Fourier coefficients of $L^1[0,1]$ functions (because the map sending $f \in L^1[0,1]$ to the sequence

$$(\hat{f}(0), \hat{f}(1), \hat{f}(-1), \hat{f}(2), \ldots) \in c_0$$

cannot be surjective).

More recently in Chapter 1, we discussed $f \in L^2[0,1]$ and deduced from the fact that the functions $e^{2\pi int}$ ($n \in \mathbb{Z}$) form an orthonormal basis for $L^2[0,1]$ that

$$\lim_{n\to\infty} \|f - S_n f\|_2 = 0$$

holds for $f \in L^2[0,1]$ (root mean square convergence).

We are restricting now to continuous $f$ (which are certainly in $L^2[0,1]$) but discussing poiuntwise convergence.

*Proof.* We consider the Banach space $CP[0,1] = \{f \in C[0,1] : f(0) = f(1)\}$. (Functions $f \in CP[0,1]$ are the restrictions to $[0,1]$ of continuous 1-periodic functions $f\colon \mathbb{R} \to \mathbb{C}$.) (It is quite easy to see that $CP[0,1]$ is a closed linear subspace of $C[0,1]$ and so a Banach space in the supremum norm $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$. It is the kernel of the bounded linear functional $\alpha\colon C[0,1] \to \mathbb{C}$ given by $\alpha(f) = f(0) - f(1)$.)

The operator $s_n\colon CP[0,1] \to \mathbb{C}$ given by $s_n(f) = S_n f(0)$ can be written

$$s_n f = \int_0^1 f(t) D_n(t)\, dt$$

where $D_n(t)$ is the Dirichlet kernel

$$D_n(t) = \sum_{j=-n}^{n} e^{2\pi ijt}$$

as one can easily check by writing the integral of the sum as the sum of the integrals.

We can easily check that $s_n$ is a linear functional (easy manipulation of integal of a sum and integral of a constant multiple) and bounded because

$$\begin{aligned} |s_n f| &= \left| \int_0^1 f(t) D_n(t)\, dt \right| \\ &\leq \int_0^1 |f(t)||D_n(t)|\, dt \\ &\leq \int_0^1 \|f\|_\infty |D_n(t)|\, dt \\ &= \|f\|_\infty \int_0^1 |D_n(t)|\, dt = \|D_n\|_1 \|f\|_\infty \end{aligned}$$

If $\lim_{n\to\infty}(S_n f)(0)$ did exist for each $f \in CP[0,1]$, then it would also be true that

$$\sup_{n\geq 0} |(S_n f)(0)| = \sup_{n \geq 0} |s_n f| < \infty$$

for each $f \in CP[0, 2\pi]$.

But then the uniform boundedness principle (Theorem 2.5.1) would imply that $\sup_n \|s_n\|_{\mathrm{op}} < \infty$. To show that this is not so (and hence that there is $f \in CP[0,1]$ where $\lim_{n\to\infty} s_n f$ does not exist) we would need two more facts

- $\|s_n\|_{\mathrm{op}} = \|D_n\|_1$ (we only proved $\leq$ above)

  A rough idea is to take $f(t) = \overline{D_n(t)}/|D_n(t)|$ so that $\|f\|_\infty = 1$ while $s_n f = \int_0^1 |D_n(t)|\, dt = \|D_n\|_1$ but that has a problem because $D_n(t)$ is zero for some $t$, and so we will not be able to do this and get $f$ continuous. We can define $f \in L^\infty$ this way and then rely on Lusin's theorem to modify $f$ to keep it bounded by 1 and unchanged for a majority of $t \in [0,1]$.

  In fact since $D_n(t)$ is always real, the $f(t)$ we want will have values $\pm 1$ (and there will be some jump points where $D_n(t) = 0$). We can make $f$ go linearly between 1 and $-1$ (or vice versa) near those jump points. If the transitions are very steep, then this will have only a small affect on $s_n f = \int_0^1 f(t) D_n(t)\, dt$ as compared to integrating $\overline{D_n(t)}/|D_n(t)|$ times $D_n(t)$. So $s_n f$ will be almost $\|D_n\|_1$ with $\|f\|_\infty = 1$, $f \in CP[0,1]$.

- $\|D_n\|_1 \to \infty$ as $n \to \infty$

  We skipped this already once in MA3421, Lemma 3.3.2.

  It can be proved by a fairly direct calculation to show that

  $$D_n(t) = \frac{\sin(2n+1)\pi t}{\sin \pi t}$$

and then estimating the integral in a somewhat careful way.

□

*2.5.5 Remark.* We can easily show more generally

$$S_n f(t_0) = \int_0^1 f(t) D_n(t_0 - t)\, dt$$

(we used $t_0 = 0$ above) and using periodicity of $D_n$ that, ofr each fixed $t_0 \in [0, 1]$ the function $D_n(t - t_0)$ has the same $L^1[0, 1]$ norm $\|D_n\|_1$. So the above argument would apply to any other fixed $t_0$ and not just $t_0 = 0$. However if if you define $g(t) = f(t - t_0)$ then $\hat{g}(n) = \hat{f}(n)e^{2\pi i n t_0}$ and $S_n g(0) = S_n f(t_0)$. So if we want $\lim_{n\to\infty} S_n f(t_0)$ to not exist, then we can choose $g$ by appealing to Theorem 2.5.3 such that $\sup_n |S_n g(0)| = \infty$, and take $f(t) = g(t + t_0)$.

*2.5.6 Remark.* The open mapping, closed graph and uniform boundedness principle are all considered to be fundamental theorems in functional analysis.

There is at least one more theorem of central importance, the Hahn Banach theorem. It relies on Zorn's lemma and we will explain that first.

## 2.6 Zorn's lemma

**2.6.1 Definition.** A *partial order* $\leq$ on a set $S$ is a relation on $S$ satisfying the following properties

(i) $x \leq x$ is true for all $x \in S$

(ii) $x, y \in S$, $x \leq y$ and $y \leq x \Rightarrow x = y$

(iii) (transitivity) $x, y, z \in S, x \leq y, y \leq z \Rightarrow x \leq z$

The set $S$ together with some chosen partial order is called a *partially ordered set* $(S, \leq)$.

*2.6.2 Examples.* (i) The most familiar example of $\leq$ is the one for numbers (real numbers $S = \mathbb{R}$, rational numbers $S = \mathbb{Q}$, integers $S = \mathbb{Z}$, natural numbers $S = \mathbb{N}$). In fact we can take $S \subseteq \mathbb{R}$ to be any subset and keep the usual interpretation of $\leq$ to get a partially ordered set $(S, \leq)$.

These examples are rather too special to get a general feel for what a partial order might be like.

(ii) A fairly general kind of example comes by taking $\leq$ to be set inclusion $\subseteq$.

Start with any set $T$ and take $S = \mathcal{P}(T) =$ the power set of $T$, by which we mean the set of all subsets of $T$,

$$\mathcal{P}(T) = \{R : R \subseteq T\}.$$

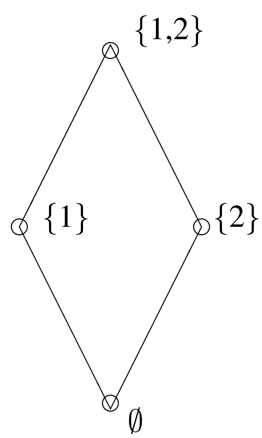
Declare $R_1 \leq R_2$ to mean $R_1 \subseteq R_2$.

It is not hard to see that the rules for a partial order are satisfied.

(iii) If we take small sets for $T$, we can picture what is going on.

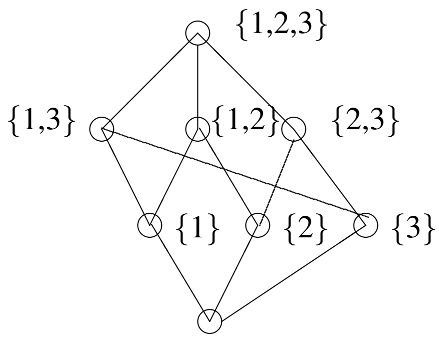
If $T = \{1, 2\}$ has 2 elements, then $\mathcal{P}(T)$ has $2^2 = 4$ elements

$$\mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$




What the picture represents is the inclusion ordering between the sets. $A \subseteq B$ means we can get from $A$ to $B$ traveling upwards always. The picture shows that $A = \{1\}$ and $B = \{2\}$ are not related by inclusion. Neither is contained in the other. This is why the word 'partial' is used.

If we go to a set with 3 elements and $2^3 = 8$ subsets, the picture will become a little trickier to draw. Some paths cross.




You can see now that the single element subsets $\{1\}$, $\{2\}$, $\{3\}$ are not comparable one to another, and neither are the 2 element subsets. There are sets like $\{1\}$ and $\{2, 3\}$ with different numbers of elements which are also not comparable.

(iv) One way to get new partially ordered sets from old ones is to take subsets. If $(S, \leq)$ is any partially ordered set and $S_0 \leq S$ is any subset we can use the same $\leq$ on $S_0$, or restricted to $S_0$, to get a partial order on $S_0$.

If we want to make it more formal, we define the relation $\leq_0$ on $S_0$ by the rule

$$x \leq_0 y \iff x \leq y.$$

So $x \leq_0 y$ for $x, y \in S_0$ means that $x \leq y$ holds in $S$. The difference between $\leq_0$ and $\leq$ is that $\leq_0$ forgets whatever relations there were involving elements of $S \setminus S_0$.

It is not hard to verify that $\leq_0$ is a partial order on $S_0$.

We call it the induced order on $S_0$. Looking at the picture above where $S = \mathcal{P}(\{1, 2, 3\})$ we could take

$$S_0 = \{\emptyset, \{1\}, \{2\}, \{1, 2, 3\}\}$$

and we would get the picture for $(S_0, \leq_0)$ by erasing all nodes in the picture corresponding to the other 4 subsets. In fact, if we do this, we will end up with a picture that is really similar to the picture for a power set of a set with 2 elements.

We can end up with other pictures by taking different $S_0$.

**2.6.3 Definition.** A *linear ordering* $\leq$ on a set $S$ is a partial order with the additional property

$$x, y \in S \Rightarrow x \leq y \text{ or } y \leq x.$$

(In other words, every pair of elements are comparable.)

A *linearly ordered set* is a set $S$ with a linear ordering $\leq$ specified.

**2.6.4 Definition.** If $(S, \leq)$ is a partially ordered set, then a *chain* in $S$ is a subset $C \subseteq S$ that becomes linearly ordered in the induced order.

*2.6.5 Example.* In the case $S = \mathcal{P}(\{1, 2, 3\})$, one possible chain is

$$C = \{\emptyset, \{1, 3\}, \{1, 2, 3\}\}.$$

**2.6.6 Definition.** If $(S, \leq)$ is a partially ordered set, and $R \subseteq S$ is any subset, then an element $u \in S$ is called an *upper bound* for $R$ if it is true that

$$x \leq u \forall x \in R.$$

*2.6.7 Example.* In $\mathbb{R}$, with the usual order, there is no upper bound for the subset $\mathbb{N}$.

In $S = \mathcal{P}(\{1, 2, 3\})$ (with set inclusion $\subseteq$ as our $\leq$), if we take

$$R = \{\emptyset, \{1, 3\}, \{2, 3\}\}$$

then we do have an upper bound $u = \{1, 2, 3\}$ in $S$ for $R$.

Notice however, in the same setting, that $\{1, 3\}$ is *not* an upper bound for $R$. For example $\{2, 3\} \not\subseteq \{1, 3\}$ So $\{1, 3\}$ fails to be an upper bound because it is not larger than everything in $R$.

But, there is a little difference that comes because we have only a partial order and not a linear order. In the case of $\mathbb{R}$, if something is not an upper bound for a subset, then there is something larger in the subset. In our example, there is nothing in $R$ strictly bigger than $\{1, 3\}$. There are elements in $R$ that are not comparable to $\{1, 3\}$.

**2.6.8 Definition.** If $(S, \leq)$ is a partially ordered set, then $m \in S$ is called a *maximal element* of $S$ if

$$x \in S, m \leq x \Rightarrow m = x.$$

(There is nothing in $S$ strictly larger than $m$.)

*2.6.9 Example.* This is **not** the same as an element of $S$ that is an upper bound for $S$. If we take a subset of the example drawn above where $S_0 = \{\emptyset, \{1\}, \{2\}, \{3\}\} \subseteq \mathcal{P}(\{1, 2, 3\})$, we can see that each of $\{1\}$, $\{2\}$ and $\{3\}$ is maximal inside $S_0$. (There is nothing in $S_0$ strictly larger.)

**2.6.10 Theorem** (Zorn's lemma)**.** *If $(S, \leq)$ is a nonempty partially ordered set, with the property that every nonempty chain $C \subset S$ has an upper bound in $S$, then $S$ has a maximal element.*

*Proof (not given).* A 'proof' would have to use the Axiom of Choice, or something equivalent to it, because one can use Zorn's lemma to prove the Axiom of Choice and there is no way to prove the Axiom of Choice itself based on standard set theory.

It is perhaps not very satisfactory to assume Zorn's lemma, but that is what we are going to do. □

Here is an example of an application of Zorn's lemma. It is perhaps a fairly typical application.

**2.6.11 Proposition.** *A (nonzero) unital ring contains a maximal proper ideal.*

Recall that a *ring* $R$ is an abelian group under an operation $+$ and has a multiplication satisfying reasonable properties (distributivity $x(y+z) = xy + xz$ and $(x+y)z = xz + yz$). Associativity ($(xy)z = x(yz)$) is usually assumed but it is not really necessary for us. A unital ring is one where there is a multiplicative identity $1_R \in R$ (satisfying $1_R x = x = x 1_R$, $x \in R$). It is common to assume $1_R \neq 0$ — but if $1_R = 0$ then for each $x \in R$ we have $x = 1_R x = 0x = 0$, and so $R = \{0\}$. An ideal $I \subseteq R$ is a subgroup under addition that satisfies $rx, xr \in I \forall x \in I, r \in R$.

By a maximal (proper) ideal $M \subseteq R$ we mean an ideal with $M \neq R$ and the maximality property that if $M \subseteq I \subsetneq R$ for any (proper) ideal $I$, then $M = I$.

*Proof.* The idea is to apply Zorns lemma 2.6.10 where the the partially ordered set $S$ is the set of all proper ideals $I \subsetneq R$ and the relation $\leq$ is $\subseteq$. The proposition states exactly that $(S, \leq)$ must have a maximal element and this is also the conclusion of Zorns lemma. So what we need to do is establish the hypotheses of Zorns lemma (that every chain in $S$ has an upper bound).

To show $S \neq \emptyset$, we can take $\mathcal{I} = \{0\}$ as proper ideal. [Since $1 \neq 0$, $\mathcal{I} \in S$ (that is $\mathcal{I}$ is a proper ideal).]

If $C$ is a nonempty chain (in $S$), we put $\mathcal{I} = \bigcup_{I \in C} I$. Clearly $I \subseteq \mathcal{I} \forall I \in C$ and the issue in proving that $\mathcal{I}$ is an upper bound for $C$ (in $S$) is to show $\mathcal{I} \in S$. Now $0 \in \mathcal{I}$ since $C \neq \emptyset$, so $\exists I \in C$ and thus $0 \in I \Rightarrow 0 \in \mathcal{I}$. Next to establish that $I$ is an additive subgroup of $R$ we need to show that $x, y \in \mathcal{I} \Rightarrow x - y \in \mathcal{I}$.

Now $x, y \in \mathcal{I}$ implies $\exists I_1, I_2 \in C$ with $x \in I_1$ and $y \in I_2$. As $C$ is a chain we have either $I_1 \subseteq I_2$ or $I_2 \subseteq I_1$. Say $I_1 \subseteq I_2$ as the other case is similar. Then we have $x, y \in I_2$ and so $x - y \in I_2 \subseteq \mathcal{I}$ (because $I_2$ is an ideal).

Next to show $\mathcal{I}$ has the ideal property, pick $x \in \mathcal{I}$ and $r \in R$. Then $x \in I$ for some $I \in C$ and so $rx, xr \in I \subseteq \mathcal{I}$. Thus $rx, xr \in \mathcal{I}$.

Finally, we must show $\mathcal{I} \neq R$ (that is that $\mathcal{I}$ is a proper ideal and so $\mathcal{I} \in S$). To do this we note that $1_R \notin \mathcal{I}$. If $1_R \in \mathcal{I}$, then $1_R \in I$ for some $I \in C$ and so $r = r1_R \in I \forall r \in R$ (by the ideal property of $I$) and this would mean $R = I$, $I \notin S$. □

The above is a fairly typical application of Zorns lemma, but other applications can require more technicalities.

## 2.7 Hahn-Banach Theorem

Please refer to section 1.4 for the notion of the dual space of a Banach space.

*2.7.1 Examples.* (i) If $E$ is a normed space with $\dim E = n < \infty$, then $\dim(E^*) = \dim(E)$.

*Proof.* We know that $E^*$ coincides with the vector space dual (all linear transformations : $E \to \mathbb{K}$ are continuous since $\dim E < \infty$). The vector space dual has the same dimension $n$ as $E$ does.

Why is that? Recall that if $e_1, e_2, \ldots, e_n$ is a basis for $E$, then we can make a basis for $E^*$ by taking the linear functionals $e_j^* \colon E \to \mathbb{K}$ given by

$$e_j^* \left( \sum_{k=1}^{n} x_k e_k \right) = x_j \quad (1 \leq j \leq n).$$

□

(ii) But (unlike the Hilbert space case) if $E$ is a finite dimensional normed space, then there is no reason to believe that $E^*$ is the 'same normed space' as $E$

All $n$-dimensional normed spaces are isomorphic and so there is an isomorphism between $E$ and $E^*$ if $\dim E = n < \infty$. (For instance, fix a basis $e_1, e_2, \ldots, e_n$ of $E$ and have the isomorphism map $e_j$ to $e_j^*$.) But the norms will usually not match up, that is there is no isometric linear isomorphism in general.

As an example, take $E = \mathbb{K}^n$ with the norm $\|\cdot\|_p$. Using Hölder's inequality in place of the Cauchy-Schwarz inequality we used when $p = 2$ (in Proposition 1.5.1) it is possible to show that $E^*$ is $\mathbb{K}^n$ with the $\ell^q$ norm, where $q$ is the conjugate exponent to $p$, that is $1/p + 1/q = 1$.

That is given $y = (y_1, y_2, \ldots, y_n) \in \ell^q_n = (\mathbb{K}^n, \|\cdot\|_q)$ we define a linear functional $\alpha_y \colon \ell^p_n \to \mathbb{K}$ by

$$\alpha_y(x) = \alpha_y(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} x_j y_j$$

and then $y \mapsto \alpha_y$ is a linear isomorphism from $\ell^q_n$ to $(\ell^p_n)^*$ with $\|\alpha_y\| = \|y\|_q$ (so that the isomorphism is isometric).

This works for $1 \leq p \leq \infty$.

(iii) In a similar way to the proof of Proposition 1.5.2, we could also show that $(\ell^p)^* = \ell^q$ in the infinite dimensional case, but we need to rule out $p = \infty$. For the $\|\cdot\|_\infty$ norm on sequence spaces, there is a substitute that $c_0^* = \ell^1$.

(We should not really say 'equals'. There is a similar thing to $y \mapsto \alpha_y$ where now $y$ and $x$ are infinite sequences.)

(iv) It is also true that $(L^p)^* = L^q$ for $1 \leq p < \infty$ (and $q$ the conjugate exponent) but the proof for this is rather harder.

Again the correct statement is not equality, but an isomorphism sending $g \in L^q$ to a linear functional $\alpha_g \colon L^p \to \mathbb{K}$ given by $\alpha_g(f) = \int gf$. And we need to be specific about which $L^p$ space we mean. But is is true for $L^p[0,1]$, $L^p[a,b]$ and $L^p(\mathbb{R})$ (when we take $L^q$ over the same domain and integrals over the appropriate domain).

**2.7.2 Theorem** (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*

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

*Then there exists an extension $\beta \colon E \to \mathbb{R}$ of $\alpha$ which is linear and satisfies* *(i)* $\beta(x) = \alpha(x)$ *for all* $x \in M$ *(i.e.* $\beta$ *extends* $\alpha$*)*

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

[The most frequently used case is where $E$ is a normed space and $\alpha$ 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 $\alpha$ 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 $\beta$ cannot have a smaller norm than $\alpha$ 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 $\alpha$ satisfying the inequality. What we do is consider all possible linear extensions $\gamma \colon M_\gamma \to \mathbb{R}$ of $\alpha$ to linear subspaces $M_\gamma$ of $E$ containing $M$. To be more precise we demand that $M \subseteq M_\gamma \subseteq E$, that $\gamma$ 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 $\delta$ extends $\gamma$). 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 $S \neq \emptyset$ as $\alpha \colon M \to \mathbb{R}$ belongs to $S$. Take a nonempty chain $C$ in $S$. We can take $M_\delta = \bigcup_{\gamma \in C} M_\gamma$ and define $\delta \colon M_\delta \to \mathbb{R}$ 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 $\delta$ 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 have $\gamma_1 \leq \gamma_2$ 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 $\gamma_1$ and $\gamma_2$). 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 $M_\delta$ is a subspace, then that $\delta$ 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 $M_\gamma$ is linear, we have $\lambda x + y \in M_\gamma \subset M_\delta$. So $M_\delta$ 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 $\gamma$ is linear). So $\delta$ 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 $\alpha$. 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 $\delta$ 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 $\gamma$ 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 $\gamma$ to the space spanned by $x_0$ and $\gamma$ while keeping the extension linear and still obeying the inequality. In other words, we can get a contradiction to the maximality of $\gamma$ if $M_\gamma \neq E$.

The linear span of $M_\gamma$ and $x_0$ is

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

Also each element $y \in N$ 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 $x - x' = (\lambda' - \lambda)x_0$. If $\lambda' - \lambda \neq 0$ the $x_0 = (1/(\lambda' - \lambda))(x - x') \in M_\gamma$, which is false. Thus $\lambda' - \lambda = 0$, $\lambda = \lambda'$ 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 $c$, $\delta$ will be linear on $N$ and will extend $\gamma$, 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

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

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

$$c \leq p(x+x_0) - \gamma(x) \text{ and } -p(x+x_0) - \gamma(x) \leq c \quad (\forall x \in M_\gamma) \qquad (2.7.1)$$

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 (2.7.1).

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 $\gamma$. So we must have $M_\gamma = E$. □

**2.7.3 Theorem** (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*

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

*Then there exists an extension $\beta \colon E \to \mathbb{C}$ of $\alpha$ 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 $\alpha$.

Note that $\alpha_1$ is $\mathbb{R}$-linear and $|\alpha_1(x)| \leq |\alpha(x)| \leq p(x)$. So, by Theorem 2.7.2, we can extend $\alpha_1$ to get an $\mathbb{R}$-linear $\beta_1\colon E \to \mathbb{R}$ 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

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

Then $\beta$ is easily seen to be $\mathbb{R}$-linear on $E$ and, by our observation, it extends $\alpha$. To show $\beta$ 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

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

**2.7.4 Corollary.** *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 $\alpha$ 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 $\alpha$ to a linear functional on the whole space $E$ of norm 1. $\square$

**2.7.5 Corollary.** *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 2.7.4 to $x - y$ and observe that $\alpha(x - y) \neq 0 \Rightarrow \alpha(x) \neq \alpha(y)$. $\square$

## 2.8 Definition of reflexivity

**2.8.1 Corollary.** *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 2.7.4, given $x \in E \setminus \{0\}$, 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. □

**2.8.2 Definition.** 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 2.8.1).

*2.8.3 Examples.* i) Finite dimensional normed spaces are reflexive.

*Proof.* We know from Examples 2.7.1 (i) that if $\dim E = n < \infty$, then $\dim E^* = \dim E$.

Applying that to $E^*$ we see then that $\dim E^{**} = \dim(E^*)^* = \dim E^* = \dim E = n$.

Since the map $J\colon E \to E^{**}$ is an injective linear operator, we have

$$n = \dim E = \dim J(E) = \dim(E^{**})$$

and so $J(E) = E^{**}$. Thus $E$ is reflexive. □

ii) Hilbert spaces are reflexive (if separable).

*Proof.* By the Riesz representation theorem we know that for $H$ a (separable in our proof) Hilbert space, each $\alpha \in H^*$ is of the form $\alpha(x) = \alpha_\eta(x) = \langle x, \eta \rangle$ and the map $\eta \mapsto \alpha_\eta\colon H \to H^*$ is a conjugate linear isometric isomorphism.

We can consider $H^*$ as being the space $\bar{H}$ which is the same set as $H$, with the same addition, multiplication modified by a complex conjugation and inner product also conjugated.

A convenient formal way to keep track of this is to write $\bar{\eta}$ for elements of $\bar{H}$, though $\bar{\eta}$ is really the element $\eta$ of the set $H$. The vector space and inner product rules on $\bar{H}$ are then given by

$$\begin{aligned} \bar{\eta} + \bar{\theta} &= \overline{\eta + \theta} \\ \lambda\bar{\eta} &= \overline{\bar{\lambda}\eta} \\ \langle \bar{\eta}, \bar{\theta} \rangle &= \overline{\langle \eta, \theta \rangle} = \langle \theta, \eta \rangle. \end{aligned}$$

By the Riesz representation theorem again, each element of the dual of $\bar{H}$ is of the form

$$\bar{\eta} \mapsto \langle \bar{\eta}, \bar{\theta} \rangle$$

for some $\bar{\theta} \in \bar{H}$. But this means that the dual of $H^*$ consists entirely of maps of the form $J(\theta) \colon H^* \to \mathbb{K}$ with

$$J(\theta)(\alpha_\eta) = \alpha_\eta(\theta) = \langle \theta, \eta \rangle.$$

So $J(H) = H^{**}$ and $H$ is reflexive. □

iii) $c_0$ is not reflexive.

A proof of this would be a long verification, based on an explicit identification of $c_0^*$ and $c_0^{**}$.

One usually relies upon the following identifications:

$$c_0^* = \ell^1, \quad (\ell^1)^* = \ell^\infty, \tag{2.8.1}$$

outlined in Examples 2.7.1 (iii) where we need some explanations about what they mean.

When we consider the map $J \colon c_0 \to c_0^{**}$, it turns out that it corresponds under the identification of $c_0^{**}$ with $\ell^\infty$ to the inclusion map $c_0 \subset \ell^\infty$. The non-reflexivity of $c_0$ comes down then to the fact that $c_0 \neq \ell^\infty$.

iv) To give more examples of reflexive space (of infinite dimensions) we consider $\ell^p$ for $1 < p < \infty$. Let $q$ be given by

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

(so that when $p = 2$ then $q = 2$ also).

We usually cite the result $(\ell^p)^* = \ell^q$ and another case of the same result says $(\ell^q)^* = \ell^p$, which combine to give $(\ell^p)^{**} = \ell^p$.

More precisely, the statement $(\ell^p)^* = \ell^q$ stated formally is the result that there is a linear bijection $T\colon \ell^q \to (\ell^p)^*$ given by

$$T(y)(x) = \sum_{n=1}^{\infty} x_n y_n$$

for $x = (x_n)_{n=1}^{\infty} \in \ell^p$, $y = (y_n)_{n=1}^{\infty} \in \ell^q$.

The statement $(\ell^q)^* = \ell^p$ is then the same idea, but we could use a different letter than $T$. So there is a linear bijection $S\colon \ell^p \to (\ell^q)^*$ given by

$$S(x)(y) = \sum_{n=1}^{\infty} x_n y_n$$

for $x = (x_n)_{n=1}^{\infty} \in \ell^p$, $y = (y_n)_{n=1}^{\infty} \in \ell^q$.

March 1: Fix small typo in proof of Theorem 2.7.2.
March 9: Fix small typo in proof of Corollary 2.8.1.
Richard M. Timoney (March 9, 2017)