

1. Banach Spaces

Definition 1.1. A (real) complex normed space is a (real) complex vector space X together with a map $\|\cdot\| : X \rightarrow \mathbb{R}$, called the norm and denoted $\|\cdot\|$, such that

- (i) $\|x\| \geq 0$, for all $x \in X$, and $\|x\| = 0$ if and only if $x = 0$.
- (ii) $\|\alpha x\| = |\alpha|\|x\|$, for all $x \in X$ and all $\alpha \in \mathbb{C}$ (or \mathbb{R}).
- (iii) $\|x + y\| \leq \|x\| + \|y\|$, for all $x, y \in X$.

Remark 1.2. If in (i) we only require that $\|x\| \geq 0$, for all $x \in X$, then $\|\cdot\|$ is called a seminorm.

Remark 1.3. If X is a normed space with norm $\|\cdot\|$, it is readily checked that the formula $d(x, y) = \|x - y\|$, for $x, y \in X$, defines a metric d on X . Thus a normed space is naturally a metric space and all metric space concepts are meaningful. For example, convergence of sequences in X means convergence with respect to the above metric.

Definition 1.4. A complete normed space is called a Banach space.

Thus, a normed space X is a Banach space if every Cauchy sequence in X converges (where X is given the metric space structure as outlined above). One may consider real or complex Banach spaces depending, of course, on whether X is a real or complex linear space.

Examples 1.5.

1. If \mathbb{R} is equipped with the norm $\|\lambda\| = |\lambda|$, $\lambda \in \mathbb{R}$, then it becomes a real normed space. More generally, for $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ define

$$\|x\| = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$$

Then \mathbb{R}^n becomes a real Banach space (with the obvious component-wise linear structure).

In a similar way, one sees that \mathbb{C}^n , equipped with the similar norm, is a (complex) Banach space.

2. Equip $\mathcal{C}([0, 1])$, the linear space of continuous complex-valued functions on the interval $[0, 1]$, with the norm

$$\|f\| = \sup\{|f(x)| : x \in [0, 1]\}.$$

Then $\mathcal{C}([0, 1])$ becomes a Banach space. This norm is called the supremum (or uniform) norm and is often denoted $\|\cdot\|_\infty$. Notice that convergence with respect to this norm is precisely that of uniform convergence of the functions on $[0, 1]$.

Suppose that we now equip $\mathcal{C}([0, 1])$ with the norm

$$\|f\|_1 = \int_0^1 |f(x)| dx.$$

One can check that this is indeed a norm but $\mathcal{C}([0, 1])$ is no longer complete (so is not a Banach space). In fact, if h_n is the function given by

$$h_n(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2} \\ n(x - \frac{1}{2}), & \frac{1}{2} < x \leq \frac{1}{2} + \frac{1}{n} \\ 1, & \frac{1}{2} + \frac{1}{n} < x \leq 1 \end{cases}$$

then one sees that (h_n) is a Cauchy sequence with respect to the norm $\|\cdot\|_1$. Suppose that $h_n \rightarrow h$ in $(\mathcal{C}([0, 1]), \|\cdot\|_1)$ as $n \rightarrow \infty$. Then

$$\int_0^{1/2} |h(x)| dx = \int_0^{1/2} |h(x) - h_n(x)| dx \leq \|h - h_n\|_1 \rightarrow 0$$

and so we see that h vanishes on the interval $[0, \frac{1}{2}]$. Similarly, for any $0 < \varepsilon < \frac{1}{2}$, we have

$$\int_{\frac{1}{2}+\varepsilon}^1 |h(x) - 1| dx \leq \|h - h_n\|_1 \rightarrow 0$$

as $n \rightarrow \infty$. Therefore h is equal to 1 on any interval of the form $[\frac{1}{2} + \varepsilon, 1]$, $0 < \varepsilon < \frac{1}{2}$. This means that h is equal to 1 on the interval $[\frac{1}{2}, 1]$. But such a function h is not continuous, so we conclude that $\mathcal{C}([0, 1])$ is not complete with respect to the norm $\|\cdot\|_1$.

3. Let S be any (non-empty) set and let X denote the set of bounded complex-valued functions on S . Then X is a Banach space when equipped

with the supremum norm $\|f\| = \sup\{|f(s)| : s \in S\}$ (and the usual pointwise linear structure).

In particular, if we take $S = \mathbb{N}$, then X is the linear space of bounded complex sequences. This Banach space is denoted ℓ^∞ (or sometimes $\ell^\infty(\mathbb{N})$). With $S = \mathbb{Z}$, the resulting Banach space is denoted $\ell^\infty(\mathbb{Z})$.

4. The set of complex sequences, $x = (x_n)$, satisfying

$$\|x\|_1 = \sum_{n=1}^{\infty} |x_n| < \infty$$

is a linear space under componentwise operations (and $\|\cdot\|_1$ is a norm). Moreover, one can check that the resulting normed space is complete. This Banach space is denoted ℓ^1 .

5. The Banach space ℓ^2 is the linear space of complex sequences, $x = (x_n)$, satisfying

$$\|x\|_2 = \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2} < \infty.$$

In fact, ℓ^2 has the inner product

$$\langle x, y \rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

and so is a (complex) Hilbert space.

The following suggests that there is not a great deal of excitement to be got from finite-dimensional normed spaces.

Suppose that X is a finite-dimensional normed vector space with basis e_1, \dots, e_n . Define a map $T : \mathbb{C}^n \rightarrow \mathbb{R}$ by

$$T(\alpha_1, \dots, \alpha_n) = \|\alpha_1 e_1 + \dots + \alpha_n e_n\|.$$

Then the inequality

$$\left| \|\alpha_1 e_1 + \dots + \alpha_n e_n\| - \|\beta_1 e_1 + \dots + \beta_n e_n\| \right| \leq \|(\alpha_1 - \beta_1)e_1 + \dots + (\alpha_n - \beta_n)e_n\|$$

shows that T is continuous on \mathbb{C}^n . Now, $T(\alpha_1, \dots, \alpha_n) = 0$ only if every $\alpha_i = 0$. In particular, T does not vanish on the unit sphere, $\{z : \|z\| = 1\}$, in \mathbb{C}^n . By compactness, T attains its bounds on the unit sphere and is therefore strictly positive on this sphere. Hence there is $m > 0$ and $M > 0$ such that

$$m \sqrt{\sum_{k=1}^n |\alpha_k|^2} \leq \|\alpha_1 e_1 + \dots + \alpha_n e_n\| \leq M \sqrt{\sum_{k=1}^n |\alpha_k|^2}.$$

We have shown that the norm $\|\cdot\|$ on X is equivalent to the usual Euclidean norm on X determined by any particular basis. Consequently, any finite-dimensional linear space can be given a norm, and, moreover, all norms on a finite-dimensional linear space X are equivalent: for any pair of norms $\|\cdot\|_1$ and $\|\cdot\|_2$ there are positive constants μ, μ' such that

$$\mu\|x\|_1 \leq \|x\|_2 \leq \mu'\|x\|_1$$

for every $x \in X$. (We say that two norms $\|\cdot\|$ and $\|\cdot\|$ on X are equivalent if there are strictly positive constants m and M such that $m\|x\| \leq \|x\| \leq M\|x\|$ for every $x \in X$. This means that the norms induce the same open sets in X and equivalently (to anticipate results from the next section) that the identity map is continuous both from $(X, \|\cdot\|)$ to $(X, \|\cdot\|)$ and also from $(X, \|\cdot\|)$ to $(X, \|\cdot\|)$.

Proposition 1.6. *Suppose that X is a normed space. Then X is complete if and only if the series $\sum_{n=1}^{\infty} x_n$ converges, where (x_n) is any sequence in X satisfying $\sum_{n=1}^{\infty} \|x_n\| < \infty$.*

Proof. Suppose that X is complete, and let (x_n) be any sequence in X such that $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Let $\varepsilon > 0$ be given and put $y_n = \sum_{k=1}^n x_k$. Then, for $n > m$,

$$\begin{aligned} \|y_m - y_n\| &= \left\| \sum_{k=m+1}^n x_k \right\| \\ &\leq \sum_{k=m+1}^n \|x_k\| \\ &< \varepsilon \end{aligned}$$

for all sufficiently large m and n , since $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Hence (y_n) is a Cauchy sequence and so converges since X is complete, by hypothesis.

Conversely, suppose $\sum_{n=1}^{\infty} x_n$ converges in X whenever $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Let (y_n) be any Cauchy sequence in X . We must show that (y_n) converges. Now, since (y_n) is Cauchy, there is $n_1 \in \mathbb{N}$ such that $\|y_{n_1} - y_m\| < \frac{1}{2}$ whenever $m > n_1$. Furthermore, there is $n_2 > n_1$ such that $\|y_{n_2} - y_m\| < \frac{1}{4}$ whenever $m > n_2$. Continuing in this way, we see that there is $n_1 < n_2 < n_3 < \dots$ such that $\|y_{n_k} - y_m\| < \frac{1}{2^k}$ whenever $m > n_k$. In particular, we have

$$\|y_{n_{k+1}} - y_{n_k}\| < \frac{1}{2^k}$$

for $k \in \mathbb{N}$. Set $x_k = y_{n_{k+1}} - y_{n_k}$. Then

$$\begin{aligned} \sum_{k=1}^n \|x_k\| &= \sum_{k=1}^n \|y_{n_{k+1}} - y_{n_k}\| \\ &< \sum_{k=1}^n \frac{1}{2^k}. \end{aligned}$$

It follows that $\sum_{k=1}^{\infty} \|x_k\| < \infty$. By hypothesis, there is $x \in X$ such that $\sum_{k=1}^m x_k \rightarrow x$ as $m \rightarrow \infty$, that is,

$$\begin{aligned} \sum_{k=1}^m x_k &= \sum_{k=1}^m (y_{n_{k+1}} - y_{n_k}) \\ &= y_{n_{m+1}} - y_{n_1} \rightarrow x. \end{aligned}$$

Hence $y_{n_m} \rightarrow x + y_{n_1}$ in X as $m \rightarrow \infty$. Thus the Cauchy sequence (y_n) has a convergent subsequence and so must itself converge. ■

We shall apply this result to quotient spaces, to which we now turn. Let X be a vector space, and let M be a vector subspace of X . We define an equivalence relation \sim on X by $x \sim y$ if and only if $x - y \in M$. It is straightforward to check that this really is an equivalence relation on X . For $x \in X$, let $[x]$ denote the equivalence class containing the element x and denote the set of equivalence classes by X/M . The definitions $[x] + [y] = [x + y]$ and $\alpha[x] = [\alpha x]$, for $\alpha \in \mathbb{C}$ and $x, y \in X$, make X/M into a linear space. (These definitions are meaningful since M is a linear subspace of X . For example, if $x \sim x'$ and $y \sim y'$, then $x + y \sim x' + y'$, so that the definition is independent of the particular representatives taken from the various equivalence classes.) We consider the possibility of defining a norm on the quotient space X/M . Set

$$\|[x]\| = \inf\{\|y\| : y \in [x]\}.$$

Note that if $y \in [x]$, then $y \sim x$ so that $y - x \in M$; that is, $y = x + m$ for some $m \in M$. Hence

$$\begin{aligned} \|[x]\| &= \inf\{\|y\| : y \in [x]\} = \inf\{\|x + m\| : m \in M\} \\ &= \inf\{\|x - m\| : m \in M\}, \text{ since } M \text{ is a subspace,} \end{aligned}$$

and this is the distance between x and M in the usual metric space sense. The zero element of X/M is $[0] = M$, and so $\|[x]\|$ is the distance between x

and the zero in X/M . Now, in a normed space it is certainly true that the norm of an element is just the distance between it and zero; $\|z\| = \|z - 0\|$. So the definition of $\|[x]\|$ above is perhaps a reasonable choice.

To see whether this does give a norm or not we shall consider the various requirements. First, suppose that $\alpha \neq 0$, and consider

$$\begin{aligned}\|\alpha[x]\| &= \|[\alpha x]\| \\ &= \inf_{m \in M} \|\alpha x + m\| \\ &= \inf_{m \in M} \|\alpha x + \alpha m\|, \quad \text{since } \alpha \neq 0, \\ &= |\alpha| \inf_{m \in M} \|x + m\| \\ &= |\alpha| \|[x]\|.\end{aligned}$$

If $\alpha = 0$, this equality remains valid because $[0] = M$ and $\inf_{m \in M} \|m\| = 0$.

Next, we consider the triangle inequality;

$$\begin{aligned}\|[x] + [y]\| &= \|[x + y]\| \\ &= \inf_{m \in M} \|x + y + m\| \\ &= \inf_{m, m' \in M} \|x + m + y + m'\| \\ &\leq \inf_{m, m' \in M} (\|x + m\| + \|y + m'\|) \\ &= \|x\| + \|y\|.\end{aligned}$$

Clearly, $\|[x]\| \geq 0$ and, as noted already, $\|[0]\| = 0$, so $\|\cdot\|$ is a seminorm on the quotient space X/M . To see whether or not it is a norm, all that remains is the investigation of the implication of the equality $\|[x]\| = 0$. Does this imply that $[x] = 0$ in X/M ? We will see that the answer is no, in general, but yes if M is closed, as the following argument shows.

Proposition 1.7. *Suppose that M is a closed linear subspace of the normed space X . Then $\|\cdot\|$ as defined above is a norm on the quotient space X/M — called the quotient norm.*

Proof. According to the discussion above, all that we need to show is that if $x \in X$ satisfies $\|[x]\| = 0$, then $[x] = 0$ in X/M ; that is, $x \in M$.

So suppose that $x \in X$ and that $\|[x]\| = 0$. Then $\inf_{m \in M} \|x + m\| = 0$, and hence, for each $n \in \mathbb{N}$, there is $z_n \in M$ such that $\|x + z_n\| < \frac{1}{n}$. This means that $-z_n \rightarrow x$ in X as $n \rightarrow \infty$. Since M is a closed subspace, it follows that $x \in M$ and hence $[x] = 0$ in X/M , as required. ■

Proposition 1.8. *Let M be a closed linear subspace of a normed space X and let $\pi : X \rightarrow X/M$ be the canonical map $\pi(x) = [x]$, $x \in X$. Then π is continuous.*

Proof. Suppose that $x_n \rightarrow x$ in X . Then

$$\begin{aligned} \|\pi(x_n) - \pi(x)\| &= \|[x_n] - [x]\| \\ &= \|[x_n - x]\| \\ &= \inf_{m \in M} \|x_n - x + m\| \\ &\leq \|x_n - x\|, \quad \text{since } 0 \in M, \\ &\rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

■

Proposition 1.9. *For any closed linear subspace M of a Banach space X , the quotient space X/M is a Banach space under the quotient norm.*

Proof. We know that X/M is a normed space, so all that remains is to show that it is complete. We use the criterion established above. Suppose, then, that $([x_n])$ is any sequence in X/M such that $\sum_n \|[x_n]\| < \infty$. We show that there is $[y] \in X/M$ such that $\sum_{n=1}^k [x_n] \rightarrow [y]$ as $k \rightarrow \infty$.

For each n , $\|[x_n]\| = \inf_{m \in M} \|x_n + m\|$, and therefore there is $m_n \in M$ such that

$$\|x_n + m_n\| \leq \|[x_n]\| + \frac{1}{2^n},$$

by definition of the infimum. Hence

$$\begin{aligned} \sum_n \|x_n + m_n\| &\leq \sum_n \left(\|[x_n]\| + \frac{1}{2^n} \right) \\ &< \infty. \end{aligned}$$

But $(x_n + m_n)$ is a sequence in the Banach space X , and so the limit $\lim_{k \rightarrow \infty} \sum_{n=1}^k (x_n + m_n)$ exists in X . Denote this limit by y . Then we have

$$\begin{aligned}
\left\| \sum_{n=1}^k [x_n] - [y] \right\| &= \left\| \sum_{n=1}^k [x_n - y] \right\| \\
&= \inf_{m \in M} \left\| \sum_{n=1}^k x_n - y + m \right\| \\
&\leq \left\| \sum_{n=1}^k x_n - y + \sum_{n=1}^k m_n \right\| \\
&= \left\| \sum_{n=1}^k (x_n + m_n) - y \right\| \\
&\rightarrow 0, \quad \text{as } k \rightarrow \infty.
\end{aligned}$$

Hence $\sum_{n=1}^k [x_n] \rightarrow [y]$ as $k \rightarrow \infty$ and we conclude that X/M is complete. ■

Example 1.10. Let X be the linear space $\mathcal{C}([0, 1])$ and let M be the subset of X consisting of those functions which vanish at the point 0 in $[0, 1]$. Then M is a linear subspace of X and so X/M is a vector space.

Define the map $\phi : X/M \rightarrow \mathbb{C}$ by setting $\phi([f]) = f(0)$, for $[f] \in X/M$. Clearly, ϕ is well-defined (if $f \sim g$ then $f(0) = g(0)$) and we have

$$\begin{aligned}
\phi(\alpha[f] + \beta[g]) &= \phi([\alpha f + \beta g]) \\
&= \alpha f(0) + \beta g(0) \\
&= \alpha \phi([f]) + \beta \phi([g])
\end{aligned}$$

for any $\alpha, \beta \in \mathbb{C}$, and $f, g \in X$. Hence $\phi : X/M \rightarrow \mathbb{C}$ is linear. Furthermore,

$$\begin{aligned}
\phi([f]) = \phi([g]) &\iff f(0) = g(0) \\
&\iff f \sim g \\
&\iff [f] = [g]
\end{aligned}$$

and so we see that ϕ is one-one.

Given any $\beta \in \mathbb{C}$, there is $f \in X$ with $f(0) = \beta$ and so $\phi([f]) = \beta$. Thus ϕ is onto. Hence ϕ is a vector space isomorphism between X/M and \mathbb{C} ; i.e. $X/M \cong \mathbb{C}$ as vector spaces.

Now, it is easily seen that M is closed in X with respect to the $\|\cdot\|_\infty$ -norm and so X/M is a Banach space when given the quotient norm. We have

$$\begin{aligned}
\|[f]\| &= \inf\{\|g\|_\infty : g \in [f]\} \\
&= \inf\{\|g\|_\infty : g(0) = f(0)\} \\
&= |f(0)| \quad (\text{take } g(s) = f(0) \text{ for all } s \in [0, 1]).
\end{aligned}$$

That is, $\|[f]\| = |\phi([f])|$, for $[f] \in X/M$, and so ϕ preserves the norm. Hence $X/M \cong \mathbb{C}$ as Banach spaces.

Now consider X equipped with the norm $\|\cdot\|_1$. Then M is no longer closed in X . We can see this by considering, for example, the sequence (g_n) given by

$$g_n(s) = \begin{cases} ns, & 0 \leq s \leq 1/n \\ 1, & 1/n < s \leq 1. \end{cases}$$

Then $g_n \in M$, for each $n \in \mathbb{N}$, and $g_n \rightarrow 1$ with respect to $\|\cdot\|_1$, but $1 \notin M$. The “quotient norm” is *not* a norm in this case. Indeed, $\inf\{\|g\|_1 : g \in [f]\} = 0$ for all $[f] \in X/M$. To see this, let $f \in X$, and, for $n \in \mathbb{N}$, set $h_n(s) = f(0)(1 - g_n(s))$, with g_n defined as above. Then $h_n(0) = f(0)$ and $\|h_n\|_1 = |f(0)|/2n$. Hence

$$\inf\{\|g\|_1 : g(0) = f(0)\} \leq \|h_n\|_1 \leq |f(0)|/2n$$

which implies that

$$\inf\{\|g\|_1 : g \in [f]\} = 0.$$

The “quotient norm” on X/M assigns “norm” zero to *all* vectors.

2. Linear Operators

Definition 2.1. A linear operator T between normed spaces X and Y is a map $T : X \rightarrow Y$ such that

$$T(\alpha x + \beta x') = \alpha T x + \beta T x'$$

for all $\alpha, \beta \in \mathbb{C}$ (or \mathbb{R} in the real case) and all $x, x' \in X$.

Definition 2.2. The linear operator $T : X \rightarrow Y$ is said to be bounded if there is some $k > 0$ such that

$$\|Tx\| \leq k\|x\|$$

for all $x \in X$. If T is bounded, we define $\|T\|$ to be

$$\|T\| = \inf\{k : \|Tx\| \leq k\|x\|, x \in X\}.$$

We will see shortly that $\|\cdot\|$ really is a norm on the set of bounded linear operators from X into Y . The following result is a direct consequence of the definitions.

Proposition 2.3. *Suppose that $T : X \rightarrow Y$ is a bounded linear operator. Then we have*

$$\begin{aligned}\|T\| &= \sup\{\|Tx\| : \|x\| \leq 1\} \\ &= \sup\{\|Tx\| : \|x\| = 1\} \\ &= \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\}.\end{aligned}$$

Proof. This follows using the fact that if $\|x\| \leq 1$ (and $x \neq 0$), then we have $\|Tx\| \leq \|Tx\|/\|x\| = \|Tx\|/\|x\|$. ■

Note that if T is bounded, then, by the very definition of $\|T\|$, we have $\|Tx\| \leq \|T\|\|x\|$, for any $x \in X$. Thus, a bounded linear operator maps any bounded set in X into a bounded set in Y . In particular, the unit ball in X is mapped into (but not necessarily onto) the ball of radius $\|T\|$ in Y . This will be used repeatedly without further comment.

The next result is a basic consequence of the linear structure.

Theorem 2.4. *Given a linear operator $T : X \rightarrow Y$, for normed spaces X and Y , the following three statements are equivalent.*

- (i) T is continuous at some point in X .
- (ii) T is continuous at every point in X .
- (iii) T is bounded on X .

Proof. Clearly (ii) implies (i). We shall show that (i) implies (ii). Suppose that T is continuous at $x_0 \in X$. Let $x \in X$ and let $\varepsilon > 0$ be given. Then there is $\delta > 0$ such that if $\|z - x_0\| < \delta$ then $\|Tz - Tx_0\| < \varepsilon$. But then we have

$$\begin{aligned} \|Tx' - Tx\| &= \|Tx' - Tx + Tx_0 - Tx_0\| \\ &= \|T(x' - x + x_0) - Tx_0\|, \quad \text{by the linearity of } T, \\ &< \varepsilon \end{aligned}$$

whenever $\|(x' - x + x_0) - x_0\| < \delta$, i.e., $\|x' - x\| < \delta$. This shows that T is continuous at any $x \in X$.

(Alternatively, one could argue as follows. Suppose that (x_n) is a sequence in X such that $x_n \rightarrow x$. Then $x_n - x + x_0 \rightarrow x_0$ and so $T(x_n - x + x_0) \rightarrow Tx_0$, since T is assumed to be continuous at x_0 . Thus $Tx_n - Tx + Tx_0 \rightarrow Tx_0$, since T is linear. In other words $Tx_n - Tx \rightarrow 0$, or $Tx_n \rightarrow Tx$.)

Next we show that (iii) implies (ii). Let $\varepsilon > 0$ be given. By (iii), there is $k > 0$ such that

$$\|Tx' - Tx\| = \|T(x' - x)\| \leq k\|x' - x\|$$

for any $x, x' \in X$. Putting $\delta = \varepsilon/k$, we conclude that if $\|x' - x\| < \delta$, then $\|Tx' - Tx\| < \varepsilon$. Thus T is continuous at $x \in X$. In fact, this estimate shows that T is uniformly continuous on X , and that therefore, continuity and uniform continuity are equivalent in this context. In other words, the notion of uniform continuity can play no special rôle in the theory of linear operators, as it does, for example, in classical real analysis.

Finally, we show that (ii) implies (iii). If T is assumed to be continuous at every point of X , then, in particular, it is continuous at 0. Hence, for given

$\varepsilon > 0$, there is $\delta > 0$ such that $\|Tx\| < \varepsilon$ whenever $\|x - 0\| = \|x\| < \delta$. Now, for any $x \neq 0$, $z = \delta x / 2\|x\|$ has norm equal to $\delta/2 < \delta$. Hence $\|Tz\| < \varepsilon$. But $\|Tz\| = \delta\|Tx\|/2\|x\|$ and so we get the inequality

$$\|Tx\| < \frac{2\varepsilon\|x\|}{\delta}$$

which is valid for any $x \in X$ with $x \neq 0$. Therefore

$$\|Tx\| \leq \frac{2\varepsilon\|x\|}{\delta}$$

holds for any $x \in X$, and we conclude that T is bounded.

(Alternatively, suppose that T is continuous at 0, but is not bounded. Then for each $n \in \mathbb{N}$ there is $x_n \in X$ such that $\|Tx_n\| > n\|x_n\|$. Evidently $x_n \neq 0$. Put $z_n = x_n/n\|x_n\|$. Then $\|z_n\| = 1/n \rightarrow 0$, and so $z_n \rightarrow 0$ in X . However, $\|Tz_n\| = \|(Tx_n/n\|x_n\|)\| = \|Tx_n\|/n\|x_n\| > 1$ for all $n \in \mathbb{N}$, and so (Tz_n) does not converge to 0. This contradicts the assumed continuity of T at 0.)

■

Remark 2.5. To establish the continuity of a given linear operator, it is enough to show continuity at 0. However, it is often (marginally) easier to check boundedness than continuity.

Definition 2.6. The set of bounded linear operators from a normed space X into a normed space Y is denoted $\mathcal{B}(X, Y)$. If $X = Y$, one simply writes $\mathcal{B}(X)$ for $\mathcal{B}(X, X)$.

Proposition 2.7. *The space $\mathcal{B}(X, Y)$ is a normed space when equipped with its natural linear structure and the norm $\|\cdot\|$.*

Proof. For $S, T \in \mathcal{B}(X, Y)$ and any $\alpha, \beta \in \mathbb{C}$, the linear operator $\alpha S + \beta T$ is defined by $(\alpha S + \beta T)x = \alpha Sx + \beta Tx$ for $x \in X$. Furthermore, for any $x \in X$,

$$\|Sx + Tx\| \leq \|Sx\| + \|Tx\| \leq (\|S\| + \|T\|)\|x\|$$

and we see that $\mathcal{B}(X, Y)$ is a linear space. To see that $\|\cdot\|$ is a norm on $\mathcal{B}(X, Y)$, note first that $\|T\| \geq 0$ and $\|T\| = 0$ if $T = 0$. On the other hand, if $\|T\| = 0$, then

$$0 = \|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\}$$

which implies that $\|Tx\| = 0$ for every $x \in X$ (including, trivially, $x = 0$). That is, $T = 0$.

Now let $\alpha \in \mathbb{C}$ and $T \in \mathcal{B}(X, Y)$. Then

$$\begin{aligned}\|\alpha T\| &= \sup\{\|\alpha Tx\| : \|x\| \leq 1\} \\ &= \sup\{|\alpha| \|Tx\| : \|x\| \leq 1\} \\ &= |\alpha| \sup\{\|Tx\| : \|x\| \leq 1\} \\ &= |\alpha| \|T\|.\end{aligned}$$

Finally, we see from the above, that for any $S, T \in \mathcal{B}(X, Y)$,

$$\begin{aligned}\|S + T\| &= \sup\{\|Sx + Tx\| : \|x\| \leq 1\} \\ &\leq \sup\{(\|S\| + \|T\|)\|x\| : \|x\| \leq 1\} \\ &= \|S\| + \|T\|\end{aligned}$$

and the proof is complete. ■

Proposition 2.8. *Suppose that X is a normed space and Y is a Banach space. Then $\mathcal{B}(X, Y)$ is a Banach space.*

Proof. All that needs to be shown is that $\mathcal{B}(X, Y)$ is complete. To this end, let (A_n) be a Cauchy sequence in $\mathcal{B}(X, Y)$; then

$$\|A_n - A_m\| = \sup\left\{\frac{\|A_n x - A_m x\|}{\|x\|} : x \neq 0\right\} \rightarrow 0,$$

as $n, m \rightarrow \infty$. It follows that for any given $x \in X$, $\|A_n x - A_m x\| \rightarrow 0$, as $n, m \rightarrow \infty$, i.e., $(A_n x)$ is a Cauchy sequence in the Banach space Y . Hence there is some $y \in Y$ such that $A_n x \rightarrow y$ in Y . Set $Ax = y$. We have

$$\begin{aligned}A(\alpha x + x') &= \lim_n A_n(\alpha x + x') \\ &= \lim_n (\alpha A_n x + A_n x') \\ &= \alpha \lim_n A_n x + \lim_n A_n x' \\ &= \alpha Ax + Ax',\end{aligned}$$

for any $x, x' \in X$ and $\alpha \in \mathbb{C}$. It follows that $A : X \rightarrow Y$ is a linear operator. Next we shall check that A is bounded. To see this, we observe first that for sufficiently large $m, n \in \mathbb{N}$, and any $x \in X$,

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \|x\|$$

Taking the limit $n \rightarrow \infty$ gives the inequality $\|Ax - A_m x\| \leq \|x\|$. Hence, for any sufficiently large m ,

$$\begin{aligned}\|Ax\| &\leq \|Ax - A_m x\| + \|A_m x\| \\ &\leq \|x\| + \|A_m\| \|x\|\end{aligned}$$

and we deduce that $\|A\| \leq 1 + \|A_m\|$. Thus $A \in \mathcal{B}(X, Y)$.

We must now show that, indeed, $A_n \rightarrow A$ with respect to the norm in $\mathcal{B}(X, Y)$. Let $\varepsilon > 0$ be given. Then there is $N \in \mathbb{N}$ such that

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\| \leq \varepsilon \|x\|$$

for any $m, n > N$ and for any $x \in X$. Taking the limit $n \rightarrow \infty$, as before, we obtain

$$\|Ax - A_m x\| \leq \varepsilon \|x\|$$

for any $m > N$ and any $x \in X$. Taking the supremum over $x \in X$ with $\|x\| \leq 1$ yields $\|A - A_m\| \leq \varepsilon$ for all $m > N$. In other words, $A_m \rightarrow A$ in $\mathcal{B}(X, Y)$ and the proof is complete. ■

Remark 2.9. Note that

$$|\|A\| - \|A_m\|| \leq \|A - A_m\| \rightarrow 0$$

so that $(\|A_m\|)$ converges to $\|A\|$.

Remark 2.10. If S and T belong to $\mathcal{B}(X)$, then $ST : X \rightarrow X$ is defined by $STx = S(Tx)$, for any $x \in X$. Clearly ST is a linear operator. Also, $\|STx\| \leq \|S\| \|Tx\| \leq \|S\| \|T\| \|x\|$, which implies that ST is bounded and $\|ST\| \leq \|S\| \|T\|$. Thus $\mathcal{B}(X)$ is an example of an algebra with unit (—the unit is the bounded linear operator $\mathbb{1}x = x$, $x \in X$). If X is complete, then so is $\mathcal{B}(X)$. In this case $\mathcal{B}(X)$ is an example of a Banach algebra.

Examples 2.11.

1. Let $A = (a_{ij})$ be any $n \times n$ complex matrix. Then the map $x \mapsto Ax$, $x \in \mathbb{C}^n$, is a linear operator on \mathbb{C}^n . Clearly, this map is continuous (where \mathbb{C}^n is equipped with the usual Euclidean norm), and so therefore it is bounded. By slight abuse of notation, let us also denote by A this map, $x \mapsto Ax$.

To find $\|A\|$, we note that the matrix A^*A is self adjoint and positive, and so there exists a unitary matrix V such that VA^*AV^{-1} is diagonal:

$$VA^*AV^{-1} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & & \lambda_n \end{pmatrix}$$

where each $\lambda_i \geq 0$, and we may assume that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Now, we have

$$\begin{aligned} \|A\|^2 &= \sup\{\|Ax\| : \|x\| = 1\}^2 \\ &= \sup\{\|Ax\|^2 : \|x\| = 1\} \\ &= \sup\{(A^*Ax, x) : \|x\| = 1\} \\ &= \sup\{(VA^*AV^{-1}x, x) : \|x\| = 1\} \\ &= \sup\left\{\sum_{k=1}^n \lambda_k |x_k|^2 : \sum_{k=1}^n |x_k|^2 = 1\right\} \\ &= \lambda_1. \end{aligned}$$

It follows that $\|A\| = \lambda_1$, the largest eigenvalue of A^*A .

2. Let $K : [0, 1] \times [0, 1] \rightarrow \mathbb{C}$ be a given continuous function on the unit square. For $f \in \mathcal{C}([0, 1])$, set

$$(Tf)(s) = \int_0^1 K(s, t)f(t) dt.$$

Evidently, T is a linear operator $T : \mathcal{C}([0, 1]) \rightarrow \mathcal{C}([0, 1])$. Setting $M = \sup\{|K(s, t)| : (s, t) \in [0, 1] \times [0, 1]\}$, we see that

$$\begin{aligned} |Tf(s)| &\leq \int_0^1 |K(s, t)| |f(t)| dt \\ &\leq M \int_0^1 |f(t)| dt. \end{aligned}$$

Thus, $\|Tf\|_1 \leq M\|f\|_1$, so that T is a bounded linear operator on the space $(\mathcal{C}([0, 1]), \|\cdot\|_1)$.

3. With T defined as in example 2, above, it is straightforward to check that

$$\|Tf\|_\infty \leq M\|f\|_1$$

and that

$$\|Tf\|_1 \leq M\|f\|_\infty$$

so we conclude that T is a bounded linear operator from $(\mathcal{C}([0, 1]), \|\cdot\|_1)$ to $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$ and also from $(\mathcal{C}([0, 1]), \|\cdot\|_\infty)$ to $(\mathcal{C}([0, 1]), \|\cdot\|_1)$.

4. Take $X = \ell^1$, and, for any $x = (x_n) \in X$, define Tx to be the sequence $Tx = (x_2, x_3, x_4, \dots)$. Then $Tx \in X$ and satisfies $\|Tx\|_1 \leq \|x\|_1$. Thus T is a bounded linear operator from $\ell^1 \rightarrow \ell^1$, with $\|T\| \leq 1$. In fact, $\|T\| = 1$ (—take $x = (0, 1, 0, 0, \dots)$). T is called the left shift on ℓ^1 .

Similarly, one sees that $T : \ell^\infty \rightarrow \ell^\infty$ is a bounded linear operator, with $\|T\| = 1$.

5. Take $X = \ell^1$, and, for any $x = (x_n) \in X$, define Sx to be the sequence $Sx = (0, x_1, x_2, x_3, \dots)$. Clearly, $\|Sx\|_1 = \|x\|_1$, and so S is a bounded linear operator from $\ell^1 \rightarrow \ell^1$, with $\|S\| = 1$. S is called the right shift on ℓ^1 .

As above, S also defines a bounded linear operator from ℓ^∞ to ℓ^∞ , with norm 1.

Theorem 2.12. *Suppose that X is a normed space and Y is a Banach space, and suppose that $T : X \rightarrow Y$ is a linear operator defined on some dense linear subset $D(T)$ of X . Then if T is bounded (as a linear operator from the normed space $D(T)$ to Y) it has a unique extension to a bounded linear operator from all of X into Y . Moreover, this extension has the same norm as T .*

Proof. By hypothesis, $\|Tx\| \leq \|T\|\|x\|$, for all $x \in D(T)$, where $\|T\|$ is the norm of T as a map $D(T) \rightarrow Y$. Let $x \in X$. Since $D(T)$ is dense in X , there is a sequence (ξ_n) in $D(T)$ such that $\xi_n \rightarrow x$, in X , as $n \rightarrow \infty$. In particular, (ξ_n) is a Cauchy sequence in X . But

$$\|T\xi_n - T\xi_m\| = \|T(\xi_n - \xi_m)\| \leq \|T\| \|\xi_n - \xi_m\|,$$

and so we see that $(T\xi_n)$ is a Cauchy sequence in Y . Since Y is complete, there exists $y \in Y$ such that $T\xi_n \rightarrow y$ in Y . We would like to construct an extension \hat{T} of T by defining $\hat{T}x$ to be this limit, y . However, to be able to do this, we must show that the element y does not depend on the particular sequence (ξ_n) in $D(T)$ converging to x . To see this, suppose that (η_n) is any sequence in $D(T)$ such that $\eta_n \rightarrow x$ in X . Then, as before, we deduce that there is y' , say, in Y , such that $T\eta_n \rightarrow y'$. Now consider the combined sequence $\xi_1, \eta_1, \xi_2, \eta_2, \dots$ in $D(T)$. Clearly, this sequence also converges to x and so once again, as above, we deduce that the sequence $(T\xi_1, T\eta_1, T\xi_2, T\eta_2, \dots)$ converges to some z , say, in Y . But this sequence has the two convergent subsequences $(T\xi_k)$ and $(T\eta_m)$, with limits y and y' , respectively. It follows that $z = y = y'$. Therefore we may unambiguously define the map $\hat{T} : X \rightarrow Y$ by the prescription $\hat{T}x = y$, where y is given as above.

Note that if $x \in D(T)$, then we can take $\xi_n \in D(T)$ above to be $\xi_n = x$ for every $n \in \mathbb{N}$. This shows that $\hat{T}x = Tx$, and hence that \hat{T} is an extension of T . We show that \hat{T} is a bounded linear operator from X to Y .

Let $x, x' \in X$ and let $\alpha \in C$ be given. Then there are sequences (ξ_n) and (ξ'_n) in $D(T)$ such that $\xi_n \rightarrow x$ and $\xi'_n \rightarrow x'$ in X . Hence $\alpha\xi_n + \xi'_n \rightarrow \alpha x + x'$, and by the construction of \widehat{T} , we see that

$$\begin{aligned}\widehat{T}(\alpha x + x') &= \lim_n T(\alpha\xi_n + \xi'_n), \quad \text{using the linearity of } D(T), \\ &= \lim_n \alpha T\xi_n + T\xi'_n \\ &= \alpha\widehat{T}x + \widehat{T}x' .\end{aligned}$$

It follows that \widehat{T} is a linear map. To show that \widehat{T} is bounded and has the same norm as T , we first observe that if $x \in X$ and if (ξ_n) is a sequence in $D(T)$ such that $\xi_n \rightarrow x$, then, by construction, $T\xi_n \rightarrow \widehat{T}x$, and so $\|T\xi_n\| \rightarrow \|\widehat{T}x\|$. Hence, the inequalities $\|T\xi_n\| \leq \|T\| \|\xi_n\|$, for $n \in \mathbb{N}$, imply (—by taking the limit) that $\|\widehat{T}x\| \leq \|T\| \|x\|$. Therefore $\|\widehat{T}\| \leq \|T\|$. But since \widehat{T} is an extension of T we have that

$$\begin{aligned}\|\widehat{T}\| &= \sup\{\|\widehat{T}x\| : x \in X, \quad \|x\| \leq 1\} \\ &\geq \sup\{\|\widehat{T}x\| : x \in D(T), \quad \|x\| \leq 1\} \\ &= \sup\{\|Tx\| : x \in D(T), \quad \|x\| \leq 1\} \\ &= \|T\| .\end{aligned}$$

The equality $\|\widehat{T}\| = \|T\|$ follows.

The uniqueness is immediate; if S is also a bounded linear extension of T to the whole of X , then $S - \widehat{T}$ is a bounded (equivalently, continuous) map on X which vanishes on the dense subset $D(T)$. Thus $S - \widehat{T}$ must vanish on the whole of X , i.e., $S = \widehat{T}$. ■

Remark 2.13. This process of extending a densely-defined bounded linear operator to one on the whole of X is often referred to as “extension by continuity”. If T is densely-defined, as above, but is *not* bounded on $(D(T))$ there is no “obvious” way of extending T to the whole of X . Indeed, such a goal may not even be desirable, as we will see later—for example, the Hellinger-Toeplitz theorem.

3. Baire Category Theorem and all that

We shall begin this section with a result concerning “fullness” of complete metric spaces.

Definition 3.1. A subset of a metric space is said to be nowhere dense if its closure has empty interior.

Example 3.2. Consider the metric space \mathbb{R} with the usual metric, and let S be the set $S = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. Then S has closure $\bar{S} = \{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$ which has empty interior.

We shall denote the open ball of radius r around the point a in a metric space by $B(a; r)$. The statement that a set S is nowhere dense is equivalent to the statement that the closure, \bar{S} of S , contains no open ball $B(a; r)$ of positive radius.

The next theorem, the Baire Category theorem, tells us that countable unions of nowhere dense sets cannot amount to much.

Theorem 3.3. (*Baire Category theorem*) *The complement of any countable union of nowhere dense subsets of a complete metric space X is dense in X .*

Proof. Suppose that A_n , $n \in \mathbb{N}$, is a countable collection of nowhere dense sets in the complete metric space X . Set $A_0 = X \setminus \bigcup_{n \in \mathbb{N}} A_n$. We wish to show that A_0 is dense in X . Now, $X \setminus \bigcup_{n \in \mathbb{N}} \bar{A}_n \subseteq X \setminus \bigcup_{n \in \mathbb{N}} A_n$, and a set is nowhere dense if and only if its closure is. Hence, by taking closures if necessary, we may assume that each A_n , $n = 1, 2, \dots$ is closed. Suppose then, by way of contradiction, that A_0 is not dense in X . Then $X \setminus \bar{A}_0 \neq \emptyset$. Now, $X \setminus \bar{A}_0$ is open, and non-empty, so there is $x_0 \in X \setminus \bar{A}_0$ and $r_0 > 0$ such that $B(x_0; r_0) \subseteq X \setminus \bar{A}_0$, that is, $B(x_0; r_0) \cap \bar{A}_0 = \emptyset$. The idea of the proof is to construct a sequence of points in X with a limit which does not lie in any of the sets A_0, A_1, \dots . This will be a contradiction, since X is the union of the A_n 's.

We start by noticing that since A_1 is nowhere dense, the open ball $B(x_0; r_0)$ is not contained in A_1 . This means that there is some point $x_1 \in B(x_0; r_0) \setminus A_1$. Furthermore, since $B(x_0; r_0) \setminus A_1$ is open, there is $0 < r_1 < 1$ such that $\overline{B(x_1; r_1)} \subseteq B(x_0; r_0)$ and also $B(x_1; r_1) \cap A_1 = \emptyset$.

Now, since A_2 is nowhere dense, the open ball $B(x_1; r_1)$ is not contained in A_2 . Thus, there is some $x_2 \in B(x_1; r_1) \setminus A_2$. Since $B(x_1; r_1) \setminus A_2$ is open, there is $0 < r_2 < \frac{1}{2}$ such that $\overline{B(x_2; r_2)} \subseteq B(x_1; r_1)$ and also $B(x_2; r_2) \cap A_2 = \emptyset$.

Similarly, we argue that there is some point x_3 and $0 < r_3 < \frac{1}{3}$ such that $\overline{B(x_3; r_3)} \subseteq B(x_2; r_2)$ and also $B(x_3; r_3) \cap A_3 = \emptyset$.

Recursively, we obtain a sequence x_0, x_1, x_2, \dots in X and positive real numbers r_0, r_1, r_2, \dots satisfying $0 < r_n < \frac{1}{n}$, for $n \in \mathbb{N}$, such that

$$\overline{B(x_n; r_n)} \subseteq B(x_{n-1}; r_{n-1})$$

and $B(x_n; r_n) \cap A_n = \emptyset$.

For any $m, n > N$, both x_m and x_n belong to $B(x_N; r_N)$, and so

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_N) + d(x_N, x_n) \\ &< \frac{1}{N} + \frac{1}{N}. \end{aligned}$$

Hence (x_n) is a Cauchy sequence in X and therefore there is some $x \in X$ such that $x_n \rightarrow x$. Since $x_n \in B(x_n; r_n) \subseteq B(x_N; r_N)$, for all $n > N$, it follows that $x \in \overline{B(x_N; r_N)}$. But by construction, $\overline{B(x_N; r_N)} \subseteq B(x_{N-1}; r_{N-1})$ and $B(x_{N-1}; r_{N-1}) \cap A_{N-1} = \emptyset$. Hence $x \notin A_{N-1}$ for any N . This is our required contradiction and the result follows. ■

Remark 3.4. The theorem implies, in particular, that a complete metric space cannot be given as a countable union of nowhere dense sets. In other words, if a complete metric space is equal to a countable union of sets, then not all of these can be nowhere dense; that is, at least one of them has a closure with non-empty interior. Another corollary to the theorem is that if a metric space can be expressed as a countable union of nowhere dense sets, then it is not complete.

As a first application of the Baire Category Theorem, we will consider the Banach-Steinhaus theorem, also called the Principle of Uniform Boundedness for obvious reasons.

Theorem 3.5. (Banach-Steinhaus) *Let X be a Banach space and let \mathcal{F} be a family of bounded linear operators from X into a normed space Y such that for each $x \in X$ the set $\{\|Tx\| : T \in \mathcal{F}\}$ is bounded. Then the set of norms $\{\|T\| : T \in \mathcal{F}\}$ is bounded.*

Proof. For each $n \in \mathbb{N}$, let $A_n = \{x : \|Tx\| \leq n \text{ for all } T \in \mathcal{F}\}$. Then each A_n is a closed subset of X . Indeed,

$$A_n = \bigcap_{T \in \mathcal{F}} \{x : \|Tx\| \leq n\} = \bigcap_{T \in \mathcal{F}} T^{-1}(\{y : \|y\| \leq n\})$$

and $T^{-1}(\{y : \|y\| \leq n\})$ is closed because $\{y : \|y\| \leq n\}$ is closed in Y and every T in \mathcal{F} is continuous. Moreover, by hypothesis, each $x \in X$ lies in some A_n . Thus, we may write

$$X = \bigcup_{n=1}^{\infty} A_n.$$

By the Baire Category Theorem, together with the fact that each A_n is closed, it follows that there is some $m \in \mathbb{N}$ such that A_m has non-empty interior. Suppose, then, that x_0 is an interior point of A_m , that is, there is $r > 0$ such that $\{x : \|x - x_0\| < r\} \subseteq A_m$. By the definition of A_m , we may say that if x is such that $\|x - x_0\| < r$ then $\|Tx\| \leq m$, for every $T \in \mathcal{F}$. But then for any $x \in X$ with $\|x\| < r$, we have $\|x + x_0 - x_0\| < r$ and so

$$\begin{aligned} \|Tx\| &= \|T(x + x_0) - Tx_0\| \\ &\leq \|T(x + x_0)\| + \|Tx_0\| \\ &\leq m + m \end{aligned}$$

for every $T \in \mathcal{F}$. Hence, for any $x \in X$, with $x \neq 0$, we see that $rx/2\|x\|$ has norm $r/2 < r$ and so $\|T(rx/2\|x\|)\| \leq 2m$ for any $T \in \mathcal{F}$. It follows that $\|Tx\| \leq 4m\|x\|/r$, for any $x \in X$ and any $T \in \mathcal{F}$. This implies that $\|T\| \leq 4m/r$ for any $T \in \mathcal{F}$, i.e., $\{\|T\| : T \in \mathcal{F}\}$ is bounded. ■

The following is an application of the Uniform Boundedness Principle to the study of the joint continuity of bilinear maps.

Proposition 3.6. Suppose that X and Y are normed spaces and that $B(\cdot, \cdot) : X \times Y \rightarrow \mathbb{C}$ is a separately continuous bilinear mapping. Then $B(\cdot, \cdot)$ is jointly continuous. (That is, if $B(x, \cdot) : Y \rightarrow \mathbb{C}$ is continuous for each fixed $x \in X$, and if $B(\cdot, y) : X \rightarrow \mathbb{C}$ is continuous for each fixed $y \in Y$ then $B(\cdot, \cdot)$ is jointly continuous.)

Proof. For each $x \in X$, define $T_x : Y \rightarrow \mathbb{C}$ by $T_x(y) = B(x, y)$, for $y \in Y$. By hypothesis, each T_x is a bounded linear operator from Y into \mathbb{C} . Since \mathbb{C} is complete, we may extend T_x , by continuity, to a bounded linear map on \overline{Y} , the completion of Y . Thus we have extended $B(\cdot, \cdot)$ to $X \times \overline{Y}$ whilst retaining the separate continuity. In other words, we may assume, without loss of generality, that Y is complete, that is, we may assume that Y is a Banach space.

Now, for each fixed $y \in Y$, $x \mapsto B(x, y)$ is a bounded linear operator from X into \mathbb{C} . Hence

$$|B(x, y)| \leq C_y \|x\| \quad \text{for all } x \in X$$

for some constant $C_y \geq 0$. In terms of T_x , this becomes

$$|T_x(y)| \leq C_y \|x\| \quad \text{for all } x \in X.$$

It follows that the family $\{|T_x(y)| : x \in X, \|x\| = 1\}$ is bounded, for each fixed $y \in Y$. By the Uniform Boundedness Principle, the set $\{\|T_x\| : x \in X, \|x\| = 1\}$ is bounded. That is, there is some $K \geq 0$ such that

$$\|T_x\| \leq K \quad \text{for all } x \in X \text{ with } \|x\| = 1.$$

But T_x is linear in x ($T_{\alpha x} = \alpha T_x$ for $\alpha \in \mathbb{C}$) and so we deduce that

$$\|T_x\| \leq K \|x\| \quad \text{for all } x \in X.$$

Hence, for any given $(x_0, y_0) \in X \times Y$ and $\varepsilon > 0$,

$$\begin{aligned} |B(x, y) - B(x_0, y_0)| &= |B(x, y - y_0) + B(x, y_0) - B(x_0, y_0)| \\ &\leq |B(x, y - y_0)| + |B(x - x_0, y_0)| \\ &= |T_x(y - y_0)| + |B(x - x_0, y_0)| \\ &\leq K \|x\| \|y - y_0\| + C_{y_0} \|x - x_0\| \\ &< \varepsilon \end{aligned}$$

provided $\|x - x_0\|$ and $\|y - y_0\|$ are sufficiently small. ■

The next result we will discuss is the Open Mapping Theorem, but first a few preliminary remarks will help to clarify things.

Let A be a subset of a normed space X . For $x \in X$, we use the notation $x + A$ to denote the set

$$x + A = \{z \in X : z = x + a, \quad a \in A\}$$

and for $\lambda \in \mathbb{C}$, λA denotes the set

$$\lambda A = \{z \in X : z = \lambda a, \quad a \in A\}.$$

Then one readily checks that

$$B(a; r) = a + B(0; r)$$

and that for any $\alpha \in \mathbb{C}$, (with $\alpha \neq 0$)

$$\alpha B(a; r) = \alpha a + \alpha B(0; r) = B(\alpha a; |\alpha|r).$$

Now suppose that X and Y are normed spaces and that $T : X \rightarrow Y$ is a linear operator. Then

$$T(B(a; r)) = Ta + TB(0; r).$$

Furthermore, if now $A \subseteq X$ is any set, then $T(\alpha A) = \alpha T(A)$. Also, if $x \in \overline{T(\alpha A)}$, then there is a sequence (a_n) in A such that $T(\alpha a_n) \rightarrow x$. Hence $\alpha T a_n \rightarrow x$ and so $x \in \alpha \overline{T(A)}$. Conversely, if $x \in \alpha \overline{T(A)}$, there is a sequence (a_n) in A such that $T a_n \rightarrow x/\alpha$. Hence $T(\alpha a_n) \rightarrow x$ and we see that $x \in \overline{T(\alpha A)}$. It follows that, for any $\alpha \in \mathbb{C}$,

$$\overline{T(\alpha A)} = \alpha \overline{T(A)}.$$

We say that a subset $A \subseteq X$ is symmetric if $a \in A$ implies that $-a \in A$. We say that the subset A is convex if $\lambda a + (1 - \lambda)a' \in A$, for any $0 \leq \lambda \leq 1$, whenever a and a' belong to A . If A is symmetric or convex, then the same is true of the sets $T(A)$ and $\overline{T(A)}$.

Proposition 3.7. Suppose that $T : X \rightarrow Y$ is a bounded linear operator, where X is a Banach space and Y is a normed space. Suppose that for some $\rho > 0$ and $R > 0$

$$B(0; \rho) \subseteq \overline{T(B(0; R))}.$$

Then $B(0; \rho) \subseteq T(B(0; R))$.

Proof. Let $\varepsilon > 0$ be given, and let $y \in B(0; \rho)$. Then $y \in \overline{T(B(0; R))}$, by hypothesis, and so there is $y_1 \in T(B(0; R))$ such that $\|y - y_1\| < \varepsilon\rho$. That is, there is $x_1 \in B(0; R)$ such that $y_1 = Tx_1$ satisfies $\|y - Tx_1\| < \varepsilon\rho$. In other words, $y - Tx_1 \in B(0; \varepsilon\rho)$. But $B(0; \rho) \subseteq \overline{T(B(0; R))}$ implies that

$$B(0; \varepsilon\rho) = \varepsilon B(0; \rho) \subseteq \varepsilon \overline{T(B(0; R))} = \overline{T(B(0; \varepsilon R))},$$

i.e., $y - Tx_1 \in B(0; \varepsilon\rho) \subseteq \overline{T(B(0; \varepsilon R))}$.

Hence, as before, there is a point $x_2 \in B(0; \varepsilon R)$ such that

$$\|(y - Tx_1) - Tx_2\| < \varepsilon^2\rho.$$

That is, $y - Tx_1 - Tx_2 \in B(0; \varepsilon^2\rho) \subseteq \overline{T(B(0; \varepsilon^2 R))}$.

Continuing in this way, (i.e., by recursion) we obtain a sequence of points x_1, x_2, x_3, \dots in X such that $x_n \in B(0; \varepsilon^{n-1}R)$ and such that

$$y - Tx_1 - Tx_2 - \dots - Tx_n \in B(0; \varepsilon^n\rho) \subseteq \overline{T(B(0; \varepsilon^n R))}.$$

It follows that $y = \sum_{n=1}^{\infty} Tx_n$. However, $\|x_n\| < \varepsilon^{n-1}R$ which means that $\sum_n \|x_n\| < \infty$. Since X is complete, it follows that $\sum x_n$ converges, i.e., there is some $x \in X$ such that $\sum_{k=1}^n x_k \rightarrow x$, as $n \rightarrow \infty$. But T is continuous and therefore $\sum_{k=1}^n Tx_k = T(\sum_{k=1}^n x_k) \rightarrow Tx$, as $n \rightarrow \infty$. It follows that $y = Tx$. Furthermore,

$$\begin{aligned} \left\| \sum_{k=1}^n x_k \right\| &\leq \sum_{k=1}^n \|x_k\| \\ &< \sum_{k=1}^n \varepsilon^{k-1}R \\ &< R/(1 - \varepsilon) \end{aligned}$$

so that $\|x\| = \lim_n \left\| \sum_{k=1}^n x_k \right\| \leq R/(1 - \varepsilon) < R/(1 - 2\varepsilon)$ (suppose $\varepsilon < 1/2$). Hence $x \in B(0; R/(1 - 2\varepsilon))$, and so $y = Tx \in T(B(0; R/(1 - 2\varepsilon)))$. We have proved, so far, that

$$B(0; \rho) \subseteq T\left(B\left(0; \frac{R}{1 - 2\varepsilon}\right)\right)$$

for any $0 < \varepsilon < 1/2$.

Let $y \in B(0; \rho)$. Then $\|y\| < \rho$. Let $d > 0$ be such that $\|y\| < d < \rho$. Then

$$\begin{aligned} y \in B(0; d) &= \frac{d}{\rho} B(0; \rho) \\ &\subseteq \frac{d}{\rho} T\left(B\left(0; \frac{R}{1-2\varepsilon}\right)\right) \\ &= T\left(B\left(0; R \frac{d}{\rho(1-2\varepsilon)}\right)\right). \end{aligned}$$

Since $d/\rho < 1$, we can choose ε sufficiently small that $Rd/\rho(1-2\varepsilon) < R$. It follows that $y \in T(B(0; R))$. Hence $B(0; \rho) \subseteq T(B(0; R))$. ■

The Open Mapping Theorem is a simple consequence of this last result and the Baire Category theorem.

Theorem 3.8. (Open Mapping Theorem) Suppose that both X and Y are Banach spaces and that $T : X \rightarrow Y$ is a bounded linear operator mapping X onto Y . Then T is an open map, i.e., T maps open sets in X into open sets in Y .

Proof. Let G be an open set in X . We wish to show that $T(G)$ is open in Y . If $G = \emptyset$, then $T(G) = T(\emptyset) = \emptyset$ and there is nothing to prove. So suppose that G is non-empty. Let $y \in T(G)$. Then there is $x \in G$ such that $y = Tx$. Since G is open, there is $r > 0$ such that $B(x; r) \subseteq G$. Hence $T(B(x; r)) \subseteq T(G)$. To show that $T(G)$ is open it is certainly enough to show that $T(B(x; r))$ contains an open ball of the form $B(y; r')$, for some $r' > 0$. Now, $B(y; r') = y + B(0; r')$ and

$$T(B(x; r)) = T(x + B(0; r)) = Tx + T(B(0; r)) = y + T(B(0; r))$$

in Y . Hence, the statement that $B(y; r') \subseteq T(B(x; r))$, for some $r' > 0$, is equivalent to the statement that $y + B(0; r') \subseteq Tx + T(B(0; r))$, for some $r' > 0$, which, in turn, is equivalent to the statement that $B(0; r') \subseteq T(B(0; r))$, for some $r' > 0$. We shall prove this last inclusion.

Any $x \in X$ lies in the open ball $B(0; n)$ whenever $n > \|x\|$, and so the collection $\{B(0; n) : n \in \mathbb{N}\}$ covers X . Since $T : X \rightarrow Y$ maps X onto Y , we deduce that

$$Y = \bigcup_{n=1}^{\infty} T(B(0; n)).$$

By the Baire Category Theorem, not all the sets $T(B(0; n))$ can be nowhere dense, that is, there is some $N \in \mathbb{N}$ such that $\overline{T(B(0; N))}$ has non-empty interior. Thus there is some $y \in Y$ and $\rho > 0$ such that

$$B(y; \rho) \subseteq \overline{T(B(0; N))}.$$

Since $\overline{T(B(0; N))}$ is symmetric, it follows that also

$$B(-y; \rho) \subseteq \overline{T(B(0; N))}.$$

Furthermore, $\overline{T(B(0; N))}$ is convex, so if $\|w\| < \rho$, we have that $w + y \in B(y; \rho)$ and $w - y \in B(-y; \rho)$ and so

$$w = \frac{1}{2}(w + y) + \frac{1}{2}(w - y) \in \overline{T(B(0; N))}.$$

In other words,

$$B(0; \rho) \subseteq \overline{T(B(0; N))}.$$

By the proposition, we deduce that $B(0; \rho) \subseteq T(B(0; N))$. Hence

$$\begin{aligned} B\left(0; \frac{r\rho}{N}\right) &\subseteq \frac{r}{N} T(B(0; N)) \\ &= T(B(0; r)). \end{aligned}$$

Taking $r' = \frac{r\rho}{N}$ completes the proof. ■

As a corollary to the Open Mapping Theorem, we have the following theorem.

Theorem 3.9. (Inverse Mapping Theorem) *Any one-one and onto bounded linear mapping between Banach spaces has a bounded inverse.*

Proof. Suppose that $T : X \rightarrow Y$ is a bounded linear mapping between the Banach spaces X and Y , and suppose that T is both injective and surjective. Then it is straightforward to check that T is invertible and that its inverse, $T^{-1} : Y \rightarrow X$, is a linear mapping. We must show that T^{-1} is bounded.

To see this, note that by the Open Mapping Theorem, the image under T of the open unit ball in X is an open set in Y and contains 0. Hence there is some $r > 0$ such that

$$B(0; r) \subseteq T(B(0; 1)).$$

Thus, for any $y \in Y$ with $\|y\| < r$, there is $x \in X$ with $\|x\| < 1$ such that $y = Tx$. That is, if $\|y\| < r$, then $\|T^{-1}y\| < 1$. It follows that T^{-1} is bounded and that $\|T^{-1}\| \leq \frac{1}{r}$.

(Alternatively, we can simply remark that for any open set G in X , its image, $T(G)$, under T is open, by the Open Mapping Theorem. But $T(G)$ is precisely the pre-image of G under the inverse T^{-1} . It follows that T^{-1} is a continuous mapping.) ■

Suppose that X and Y are normed spaces and that $T : X \rightarrow Y$ is a linear operator defined on a dense linear subspace $D(T)$ of X . The linearity of the domain of definition $D(T)$ of T is of course necessary to even state the linearity of T . The point is that we do not assume, for the moment, that $D(T) = X$, or that T is bounded. We simply think of T as a linear operator from the normed space $D(T)$ into Y .

Definition 3.10. Let X and Y be normed spaces and let $T : X \rightarrow Y$ be a linear operator with dense linear domain $D(T)$. The graph of T , denoted $\Gamma(T)$, is the subset of the direct sum $X \oplus Y$ given by

$$\Gamma(T) = \{x \oplus y \in X \oplus Y : x \in D(T), \quad y = Tx\}.$$

Thus, $\Gamma(T) = \{x \oplus Tx : x \in D(T)\}$.

It is readily seen that $\Gamma(T)$ is a linear subspace of $X \oplus Y$. The space $X \oplus Y$ is equipped with the norm

$$\|x \oplus y\| = \|x\| + \|y\|$$

for $x \oplus y \in X \oplus Y$. $X \oplus Y$ is complete with respect to this norm if and only if both X and Y are complete.

Theorem 3.11. (Closed Graph Theorem) Suppose that X and Y are Banach spaces and that $T : X \rightarrow Y$ is a linear operator with domain $D(T) = X$. Then T is bounded if and only if the graph of T is closed in $X \oplus Y$.

Proof. Suppose first that T is bounded, and suppose that $((x_n \oplus y_n))$ is a sequence in $\Gamma(T)$ such that $(x_n, y_n) \rightarrow (x, y)$ in $X \oplus Y$. It follows that $x_n \rightarrow x$ and $y_n \rightarrow y$ and therefore, in particular, $Tx_n \rightarrow Tx$ in Y . However, $y_n = Tx_n$, and so $Tx_n \rightarrow y$. We conclude that $y = Tx$ and that $(x, y) \in \Gamma(T)$. Thus $\Gamma(T)$ is closed in $X \oplus Y$.

Conversely, suppose that $\Gamma(T)$ is closed in the Banach space $X \oplus Y$. Then $\Gamma(T)$ is itself a Banach space (with respect to the norm inherited from $X \oplus Y$). Define maps $\pi_1 : \Gamma(T) \rightarrow X$, $\pi_2 : \Gamma(T) \rightarrow Y$ by the assignments $\pi_1 : x \oplus Tx \mapsto x$ and $\pi_2 : x \oplus Tx \mapsto Tx$. Evidently, both π_1 and π_2 are norm decreasing and so are bounded linear operators. Moreover, it is clear that π_1 is both injective and surjective. It follows, by the Inverse Mapping Theorem, that $\pi_1^{-1} : X \rightarrow \Gamma(T)$ is bounded. But then $T : X \rightarrow Y$ is given by $T = \pi_2 \circ \pi_1^{-1}$,

$$x \xrightarrow{\pi_1^{-1}} (x, Tx) \xrightarrow{\pi_2} Tx,$$

which is the composition of two bounded linear maps and therefore T is bounded. ■

Remark 3.12. The closed graph theorem can be a great help in establishing the boundedness of linear operators between Banach spaces. Indeed, in order to show that a linear operator $T : X \rightarrow Y$ is bounded, one must establish essentially two things; firstly, that if $x_n \rightarrow x$ in X , then (Tx_n) converges in Y and, secondly, that this limit is Tx . The closed graph theorem says that to prove that T is bounded it is enough to prove that its graph is closed (provided, of course, that X and Y are Banach spaces). This means that we may *assume* that $x_n \rightarrow x$ and $Tx_n \rightarrow y$, for some $y \in Y$, and then need only show that $y = Tx$. In other words, thanks to the closed graph theorem, the convergence of (Tx_n) can be taken as part of the hypothesis rather than forming part of the proof itself.

Example 3.13. Let $X = \mathcal{C}([0, 1])$ equipped with the supremum norm, $\|\cdot\|_\infty$. Define an operator T on X by setting $D(T) = \mathcal{C}^1([0, 1])$, the linear subspace of continuously differentiable functions on $[0, 1]$, and, for $x \in D(T)$, put

$$Tx(t) = \frac{dx}{dt}(t), \quad 0 \leq t \leq 1.$$

Note that $D(T)$ is a dense linear subspace of $\mathcal{C}([0, 1])$ (by the Weierstrass approximation theorem, for example) and T is a linear operator. We shall see that the graph of T is closed, but that T is unbounded. To see that T is unbounded, we observe that if g_n denotes the function $g_n(t) = t^n$, $n \in \mathbb{N}$, $t \in [0, 1]$, then $g_n \in D(T)$ and $Tg_n = ng_{n-1}$ for $n > 1$. But $\|g_n\|_\infty = 1$ and $\|Tg_n\|_\infty = n$, so it is clear that T is unbounded.

To show that $\Gamma(T)$ is closed, suppose that $x_n \rightarrow x$ in X with $x_n \in D(T)$, and suppose that $Tx_n \rightarrow y$ in X . We must show that $y \in D(T)$ and that

$y = Tx$. For any $t \in [0, 1]$, we have

$$\begin{aligned} \int_0^t y(s) ds &= \int_0^t \lim_n \frac{dx_n}{ds} ds \\ &= \lim_n \int_0^t \frac{dx_n}{ds} ds, \quad \text{since convergence is uniform,} \\ &= \lim_n (x_n(t) - x_n(0)) \\ &= x(t) - x(0). \end{aligned}$$

Thus

$$x(t) = x(0) + \int_0^t y(s) ds, \quad \text{for } 0 \leq t \leq 1,$$

with $y \in \mathcal{C}([0, 1])$. Hence $x \in \mathcal{C}^1([0, 1])$ and $\frac{dx}{ds} = y$ on $[0, 1]$. That is, $x \in D(T)$ and $Tx = y$. We conclude that $x \oplus y \in \Gamma(T)$ and therefore $\Gamma(T)$ is closed.

Notice that T is *not* defined on the whole of X and so there is no conflict with the closed graph theorem. In fact, we can deduce from the closed graph theorem, that there is no way in which we can extend the definition of T to include every element of X in its domain of definition without spoiling either linearity or the closedness of its graph.

Definition 3.14. A linear operator $T : X \rightarrow Y$ with dense linear domain $D(T)$ is said to be closed if its graph is a closed subset of $X \oplus Y$.

The concept of closed linear operator plays a very important rôle in the theory of unbounded operators in a Hilbert space.

Theorem 3.15. (Hellinger-Toeplitz) Let $A : \mathcal{H} \rightarrow \mathcal{H}$ be a linear operator on the Hilbert space \mathcal{H} with $D(A) = \mathcal{H}$ and suppose that

$$\langle x, Ay \rangle = \langle Ax, y \rangle$$

for every $x, y \in \mathcal{H}$. Then A is bounded.

Proof. We show that the graph, $\Gamma(A)$, of A is closed. Suppose, then, that $x_n \rightarrow x$, and that $Ax_n \rightarrow y$. For any $z \in \mathcal{H}$, we have

$$\begin{aligned} \langle z, y \rangle &= \lim_n \langle z, Ax_n \rangle \\ &= \lim_n \langle Az, x_n \rangle \\ &= \langle Az, x \rangle \\ &= \langle z, Ax \rangle. \end{aligned}$$

Hence $\langle z, y - Ax \rangle = 0$ for all $z \in \mathcal{H}$. Taking $z = y - Ax$, it follows that $y - Ax = 0$, that is $y = Ax$. Thus $x \oplus y \in \Gamma(A)$, and we conclude that $\Gamma(A)$ is closed. By the closed graph theorem, it follows that A is bounded. ■

Remark 3.16. This theorem says that an everywhere-defined symmetric linear operator on a Hilbert space is necessarily bounded. Thus, unbounded symmetric operators cannot be everywhere defined. The domain of definition is a central issue in the theory of unbounded operators. It should be emphasized that unbounded operators should not be considered as somewhat pathological and of no particular interest. Indeed, the example above, the operator of differentiation, could not really be thought of as especially pathological. It turns out that many examples of operators in applications, for example, in the mathematical theory of quantum mechanics, are unbounded. Indeed, one can show that operators P, Q satisfying the famous Heisenberg commutation relation, $PQ - QP = i$, cannot both be bounded.

4. The Hahn-Banach Theorem

We now turn to a discussion of the Hahn-Banach theorem—this is concerned with the extension of a continuous linear functional on a subspace of a normed space to the whole of the space. First we must discuss Zorn’s lemma.

Definition 4.1. A partially ordered set is a non-empty set P on which is defined a relation \preceq (a partial ordering) satisfying:

- (a) $x \preceq x$, for all $x \in P$;
- (b) if $x \preceq y$ and $y \preceq x$, then $x = y$;
- (c) if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

Note that it can happen that a particular pair of elements of P are not comparable, that is, neither $x \preceq y$ nor $y \preceq x$ need hold.

Examples 4.2. 1. Let P be the set of all subsets of a given set, and let \preceq be given by set inclusion \subseteq .

2. Set $P = \mathbb{R}$, and let \preceq be the usual ordering \leq on \mathbb{R} .

3. Set $P = \mathbb{R}^2$, and define \preceq according to the prescription $(x', y') \preceq (x'', y'')$ provided that both $x' \leq x''$ and $y' \leq y''$ in \mathbb{R} .

4. Any subset of a partially ordered set inherits the partial ordering and so is itself a partially ordered set.

Definition 4.3. An element m in a partially ordered set (P, \preceq) is said to be maximal if $m \preceq x$ implies that $x = m$. Thus, a maximal element cannot be “majorized” by any other element.

Example 4.4. Let P be the half-plane in \mathbb{R}^2 given by $P = \{(x, y) : x + y \leq 0\}$, equipped with the partial ordering as in example 3, above. Then one sees that *each* point *on* the line $x + y = 0$ is a maximal element. Thus P has many maximal elements. Note that P has *no* “largest” element, i.e., there is no element $z \in P$ satisfying $x \preceq z$, for all $x \in P$.

Definition 4.5. An upper bound for a subset A in a partially ordered set (P, \preceq) is an element $x \in P$ such that $a \preceq x$ for all $a \in A$.

A partially ordered set (P, \preceq) is called a chain (or totally ordered, or linearly ordered) if for any pair $x, y \in P$, either $x \preceq y$ or $y \preceq x$ holds. In other words, P is totally ordered if every pair of points in P are comparable.

A subset C of a partially ordered set (P, \preceq) is said to be totally ordered (or a chain in P) if for any pair of points $c', c'' \in C$, either $c' \preceq c''$ or $c'' \preceq c'$; that is, C is totally ordered if any pair of points in C are comparable.

We now have sufficient terminology to state Zorn's lemma which we shall take as an axiom.

Zorn's lemma. Let P be a partially ordered set. If each totally ordered subset of P has an upper bound, then P possesses at least one maximal element.

Remark 4.6. As stated, the intuition behind the statement is perhaps not evident. The idea can be roughly outlined as follows. Suppose that a is any element in P . If a is not itself maximal then there is some $x \in P$ with $a \preceq x$. Again, if x is not a maximal element, then there is some $y \in P$ such that $x \preceq y$. Furthermore, the three elements a, x, y form a totally ordered subset of P . If y is not maximal, add in some greater element, and so on. In this way, one can imagine having obtained a totally ordered subset of P . By hypothesis, this set has an upper bound, α , say. (This means that we rule out situations such as having arrived at, say, the natural numbers $1, 2, 3, \dots$, (with their usual ordering) which one could think of having got by starting with 1, then adding in 2, then 3 and so on.) Now if α is not a maximal element, we add in an element greater than α and proceed as before. Zorn's lemma can be thought of as stating that this process eventually must end with a maximal element.

Zorn's lemma can be shown to be logically equivalent to the axiom of choice (and, indeed, to the Well-Ordering Principle). We recall the axiom of choice.

Axiom of Choice. Let $\{A_\alpha : \alpha \in J\}$ be a family of non-empty sets, indexed by the non-empty set J . Then there is a mapping $\varphi : J \rightarrow \bigcup_{\alpha} A_\alpha$

such that $\varphi(\alpha) \in A_\alpha$ for each $\alpha \in J$.

Thus, the axiom says that we can “choose” a family $\{a_\alpha\}$ with $a_\alpha \in A_\alpha$, for each $\alpha \in J$, namely, the range of φ . As a consequence, this axiom gives substance to the cartesian product $\prod_{\alpha} A_\alpha$.

After these few preliminaries, we return to consideration of the Hahn-Banach theorem whose proof rests on an application of Zorn's lemma.

Definition 4.7. A linear map $\lambda : X \rightarrow \mathbb{C}$, from a linear space X into \mathbb{C} is called a linear functional. A real-linear functional on a real linear space is a real-linear map $\lambda : X \rightarrow \mathbb{R}$.

Example 4.8. Let X be a (complex) linear space, and $\lambda : X \rightarrow \mathbb{C}$ a linear functional on X . Define $\ell : X \rightarrow \mathbb{R}$ by $\ell(x) = \operatorname{Re} \lambda(x)$ for $x \in X$. Then ℓ is a real-linear functional on X if we view X as a real linear space. Substituting ix for x , it is straightforward to check that

$$\lambda(x) = \ell(x) - i\ell(ix)$$

for any $x \in X$. On the other hand, suppose that $u : X \rightarrow \mathbb{R}$ is a real linear functional on the complex linear space X (viewed as a real linear space). Set

$$\mu(x) = u(x) - iu(ix)$$

for $x \in X$. Then one sees that $\mu : X \rightarrow \mathbb{C}$ is (complex) linear. Furthermore, $u = \operatorname{Re} \mu$, and so we obtain a natural correspondence between real and complex linear functionals on a complex linear space X via the above relations.

It is more natural to consider the Hahn-Banach theorem for real normed spaces; the complex case will be treated separately as a corollary.

Theorem 4.9. (Hahn-Banach extension theorem) Suppose that M is a (real-) linear subspace of a real normed space X and that $\lambda : M \rightarrow \mathbb{R}$ is a bounded real-linear functional on M . Then there is a bounded linear functional Λ on X which extends λ and with $\|\Lambda\| = \|\lambda\|$; that is, $\Lambda : X \rightarrow \mathbb{R}$, $\Lambda \upharpoonright M = \lambda$, and

$$\|\Lambda\| = \sup \left\{ \frac{|\Lambda(x)|}{\|x\|} : x \in X, x \neq 0 \right\} = \|\lambda\| = \sup \left\{ \frac{|\lambda(x)|}{\|x\|} : x \in M, x \neq 0 \right\}.$$

Proof. One idea would be to extend λ to the subspace of X obtained by enlarging M by one extra dimension — and then to keep doing this. However, one must then give a convincing argument that eventually one does, indeed, exhaust the whole of X in this way. To circumvent this problem, the idea is to use Zorn's lemma and to show that any maximal extension of λ must, in fact, already be defined on the whole of X .

Let x_0 be a non-zero element of X with $x_0 \notin M$ and let M_1 be the (real) linear space spanned by M and x_0 (—if $M = X$, there is nothing to prove). Then any element of M_1 has the form $x + \alpha x_0$, where $x \in M$ and $\alpha \in \mathbb{R}$. If $x + \alpha x_0 = x' + \alpha' x_0$, then $x - x' + (\alpha - \alpha')x_0 = 0$ and so it follows that $x = x'$ and $\alpha = \alpha'$. In other words, this representation of each element of M_1 is unique. We define $\lambda_1 : M_1 \rightarrow \mathbb{R}$ by

$$\lambda_1(x + \alpha x_0) = \lambda(x) + \alpha a$$

where a is any fixed real number. It is evident that λ_1 is real-linear and that $\lambda_1 \upharpoonright M = \lambda$. Now, extending any bounded linear operator cannot decrease its norm, and so λ_1 has the same norm as λ provided we can find a so that

$$(*) \quad |\lambda(x) + \alpha a| \leq \|\lambda\| \|x + \alpha x_0\|$$

for all $x \in M$ and $\alpha \in \mathbb{R}$. This clearly holds for $\alpha = 0$ and so we may suppose that $\alpha \neq 0$. Then we may replace x by $-\alpha x$, and divide through by $|\alpha|$, to obtain the requirement

$$|\lambda(x) - a| \leq \|\lambda\| \|x - x_0\|$$

for all $x \in M$. We rewrite this as the requirement that

$$\underbrace{\lambda(x) - \|\lambda\| \|x - x_0\|}_{A_x} \leq a \leq \underbrace{\lambda(x) + \|\lambda\| \|x - x_0\|}_{B_x}$$

for all $x \in M$. It is possible to find such a real number a if and only if all the closed intervals $[A_x, B_x]$ have a common point. This is equivalent to $A_x \leq B_y$ for all $x, y \in M$. To see that this holds, we note that (since λ is real-valued)

$$\lambda(x - y) \leq \|\lambda\| \|x - y\|$$

for all $x, y \in M$. Thus

$$\begin{aligned} \lambda(x) - \lambda(y) &\leq \|\lambda\| \|x - y\| \\ &\leq \|\lambda\| \|x - x_0\| + \|\lambda\| \|x_0 - y\|. \end{aligned}$$

and hence we obtain $A_x \leq B_y$ for all $x, y \in M$, as required. Thus there is some $a \in \mathbb{R}$ such that the inequality $(*)$ holds, and we conclude that $\|\lambda_1\| = \|\lambda\|$. We shall now set things up so that Zorn's lemma becomes applicable. Let \mathcal{E} be the collection of extensions $e : M_e \rightarrow \mathbb{R}$ of λ which

satisfy $\|e\| = \|\lambda\|$. \mathcal{E} is partially ordered by declaring $e_1 \preceq e_2$ if e_2 is an extension of e_1 (—that is, if $M_{e_1} \subseteq M_{e_2}$ and $e_2 \upharpoonright M_{e_1} = e_1$). Let \mathcal{C} be a totally ordered subset of \mathcal{E} . Then it follows that $\bigcup_{e \in \mathcal{C}} M_e$ is a linear subspace of X . Define $e' : M' \rightarrow \mathbb{R}$, where $M' = \bigcup_{e \in \mathcal{C}} M_e$, by $e'(x) = e(x)$, for $x \in M'$, where e is such that $x \in M_e \subseteq M'$. It is clear that e' is well-defined and is an extension of λ satisfying $|e'(x)| \leq \|\lambda\| \|x\|$ for all $x \in M'$. Hence e' , defined on M' , is an element of \mathcal{E} and is an upper bound for \mathcal{C} . By Zorn's lemma, it follows that \mathcal{E} possesses a maximal element, Λ , say, defined on some linear subspace M'' , say, of X . Now, if M'' were a proper subspace of X , we could repeat our earlier argument to obtain an extension of Λ , with the same norm, which would therefore be an element of \mathcal{E} and contradict the maximality of Λ . We conclude that Λ is defined on the whole of X , and the proof is complete. ■

Corollary 4.10. (**Hahn-Banach extension theorem for complex vector spaces**) Suppose that M is a linear subspace of a complex normed space X and that $\lambda : M \rightarrow \mathbb{C}$ is a bounded linear functional on M . Then λ can be extended to a bounded linear functional Λ on X with $\|\Lambda\| = \|\lambda\|$.

Proof. Consider X as a real linear space, and let $\ell(x) = \operatorname{Re} \lambda(x)$, for $x \in M$. Then

$$\begin{aligned} |\ell(x)| &= |\operatorname{Re} \lambda(x)| \\ &\leq \|\lambda\| \|x\|, \quad x \in M, \end{aligned}$$

so ℓ is bounded with $\|\ell\| \leq \|\lambda\|$. For $x \in M$, write $\lambda(x) = \rho e^{i\theta}$, where $\rho = |\lambda(x)| \geq 0$. Then $\lambda(e^{-i\theta}x) = \rho$, and so $\ell(e^{-i\theta}x) = \operatorname{Re} \lambda(e^{-i\theta}x) = \rho = |\lambda(x)|$. It follows that $\{|\ell(x)| : x \in M\} \supseteq \{|\lambda(x)| : x \in M\}$ and so $\|\ell\| = \|\lambda\|$. By the theorem, ℓ has a real-linear extension L to X with $\|L\| = \|\ell\|$. For $x \in X$, set

$$\Lambda(x) = L(x) - iL(ix).$$

Then $\Lambda : X \rightarrow \mathbb{C}$ is complex-linear and, since $\lambda(x) = \ell(x) - i\ell(ix)$, $x \in M$, we see that Λ extends λ . We must show that $|\Lambda(x)| \leq \|\lambda\| \|x\|$, for $x \in X$, thus giving $\|\Lambda\| = \|\lambda\|$. To see this, we note that for $x \in X$ there is $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that $\alpha\Lambda(x) = |\Lambda(x)|$. Then

$$\begin{aligned} |\Lambda(x)| &= \alpha\Lambda(x) \\ &= \Lambda(\alpha x) \in \mathbb{R} \text{ since the left hand side is real} \end{aligned}$$

$$\begin{aligned}
&= L(\alpha x) \\
&\leq \|L\| \|\alpha x\| \\
&= \|\ell\| |\alpha| \|x\| \\
&= \|\lambda\| \|x\|
\end{aligned}$$

as required. ■

It is an immediate corollary that any (non-zero) normed space has a non-zero continuous linear functional. In fact, it has many as we shall now show.

Theorem 4.11. *Let X be a normed space, and let $x_0 \in X$, with $x_0 \neq 0$. Then there is a continuous linear functional λ on X with $\|\lambda\| = 1$ such that $\lambda(x_0) = \|x_0\|$.*

Proof. Let M be the subspace $M = \{\alpha x_0 : \alpha \in \mathbb{C}\}$, and define $\lambda_0 : M \rightarrow \mathbb{C}$ by $\lambda_0(\alpha x_0) = \alpha \|x_0\|$, $\alpha \in \mathbb{C}$. Evidently, λ_0 is linear, and

$$|\lambda_0(\alpha x_0)| = |\alpha| \|x_0\| = \|\alpha x_0\|$$

for all $\alpha \in \mathbb{C}$, which implies that $\|\lambda_0\| = 1$ (as a map from M into \mathbb{C}). By the Hahn-Banach theorem, λ_0 has an extension to X with the required properties. ■

Remark 4.12. The result above implies that the set of bounded linear functionals on a normed space X separates the points of X , i.e., if x_1 and x_2 are any two points of X with $x_1 \neq x_2$, then there is a bounded linear functional $\lambda : X \rightarrow \mathbb{C}$ such that $\lambda(x_1) \neq \lambda(x_2)$. One simply applies the above to the non-zero element $x_0 = x_1 - x_2$. Put another way, this result says that if $x \in X$ is such that $\lambda(x) = 0$ for every bounded linear functional λ on X , then $x = 0$. The following theorem says that the set of bounded linear functionals on X also separates points and closed subspaces.

Theorem 4.13. *Let M be a proper closed linear subspace of a normed space X and suppose that $x_0 \notin M$. Then there is a bounded linear functional λ on X such that $\lambda(x) = 0$ for all $x \in M$, but $\lambda(x_0) \neq 0$.*

Proof. Let M' be the subspace of X generated by M and $\{x_0\}$. Then any element x , say, of M' can be written uniquely as $x = z + \alpha x_0$, for $z \in M$ and some $\alpha \in \mathbb{C}$ (—if also $x = z' + \alpha' x_0$, then subtracting, we obtain that $z - z' = (\alpha' - \alpha)x_0$, and this implies that $\alpha' = \alpha$ (since $x_0 \notin M$) and therefore $z = z'$). Define $\lambda : M' \rightarrow \mathbb{C}$ by $\lambda(z + \alpha x_0) = \alpha$, for $z \in M$. Clearly $\lambda : M' \rightarrow \mathbb{C}$ is linear, $\lambda(x_0) = 1$, and λ is zero on M . To see that λ is bounded, we note that since $x_0 \notin M$ and M is closed, there is $r > 0$ such that $B(x_0; r) \cap M = \emptyset$; that is, $\|z - x_0\| > r$ for all $z \in M$. Hence, for $\alpha \neq 0$, and $z \in M$,

$$\begin{aligned}\|z + \alpha x_0\| &= |\alpha| \left\| \frac{z}{\alpha} + x_0 \right\| \\ &> |\alpha| r\end{aligned}$$

since $-z/\alpha \in M$. Thus

$$|\lambda(z + \alpha x_0)| = |\alpha| < \frac{1}{r} \|z + \alpha x_0\|.$$

It follows that $\|\lambda\| \leq 1/r$ and so $\lambda : M' \rightarrow \mathbb{C}$ is bounded. By the Hahn-Banach theorem, we may extend λ to the whole of X , and the result follows. ■

Remark 4.14. Thus, if M is a closed subspace of X and $x_0 \in X$ is such that all bounded linear functionals which vanish on M also vanish on x_0 , then $x_0 \in M$.

Definition 4.15. Suppose that $\ell : X \rightarrow \mathbb{C}$ is a linear functional on the linear space X . The kernel of ℓ is its null space; $\ker \ell = \{x \in X : \ell(x) = 0\}$.

Proposition 4.16. *Let ℓ be a linear functional on the linear space X . Then $\ker \ell$ is a linear subspace of X of codimension one.*

Proof. It is clear that $\ker \ell$ is a linear subspace of X . If ℓ is not zero, there is some $x_0 \in X$ such that $\ell(x_0) \neq 0$. Then, for any $x \in X$,

$$x = \frac{\ell(x)}{\ell(x_0)} x_0 + x - \frac{\ell(x)}{\ell(x_0)} x_0.$$

Clearly, $x - \frac{\ell(x)}{\ell(x_0)} x_0 \in \ker \ell$ and hence $X/\ker \ell$ is one-dimensional. ■

As a corollary, we see that a linear functional is determined by its kernel, up to a constant of proportionality.

Corollary 4.17. *Linear functionals ℓ_1 and ℓ_2 on the linear space X have the same kernel if and only if they are proportional.*

Proof. Suppose that $\ker \ell_1 = \ker \ell_2$ and let $x_0 \notin \ker \ell_1$. (If no such x_0 exists, then both ℓ_1 and ℓ_2 are zero on X .) Without loss of generality, we may suppose that $\ell_1(x_0) = 1$. Then for any $x \in X$, we have $x = \ell_1(x)x_0 + z$, with $z \in \ker \ell_1 = \ker \ell_2$. Hence

$$\ell_2(x) = \ell_1(x)\ell_2(x_0)$$

which shows that ℓ_1 and ℓ_2 are proportional.

The converse is clear. ■

Theorem 4.18. *Suppose that $\ell : X \rightarrow \mathbb{C}$ is a linear functional on the normed space X . Then ℓ is bounded if and only if $\ker \ell$ is closed.*

Proof. It is clear that if ℓ is bounded (equivalently, continuous), then $\ker \ell$ is closed.

Conversely, suppose that $\ker \ell$ is closed. If $\ker \ell = X$, it follows that ℓ is zero and so is certainly bounded. Suppose, then, that $\ker \ell \neq X$. Then there is $x_0 \in X$ such that $x_0 \notin \ker \ell$. By hypothesis, $\ker \ell$ is closed, and so there is $r > 0$ such that $B(x_0; r) \cap \ker \ell = \emptyset$. By replacing x_0 by $x_0/\ell(x_0)$, we may assume that $\ell(x_0) = 1$.

Suppose that $x \in X$, $x \notin \ker \ell$. Then $\ell(x) \neq 0$, and

$$-\frac{x}{\ell(x)} + x_0 \in \ker \ell.$$

It follows that $-\frac{x}{\ell(x)} + x_0 \notin B(x_0; r)$; that is,

$$\left\| \left(-\frac{x}{\ell(x)} + x_0 \right) - x_0 \right\| \geq r,$$

that is,

$$\frac{\|x\|}{|\ell(x)|} \geq r,$$

that is,

$$|\ell(x)| \leq \frac{1}{r} \|x\|$$

for all $x \notin \ker \ell$. But this inequality still holds even if $x \in \ker \ell$, and we conclude that ℓ is bounded. ■

Proposition 4.19. *Suppose that $\phi : X \rightarrow \mathbb{C}$ is a linear functional on the normed space X . Then ϕ is unbounded if and only if $\ker \phi$ is a proper dense subset of X .*

Proof. Suppose that $\ker \phi$ is dense in X . If ϕ is bounded, then it follows that ϕ must vanish on the whole of X . So $\ker \phi \neq X$ demands that ϕ be unbounded.

Conversely, suppose that $\ker \phi$ is not dense in X . Then there is some $x_0 \notin \ker \phi$ and some $r > 0$ such that $B(x_0; r) \cap \ker \phi = \emptyset$. We now argue as before to deduce that ϕ is bounded. It follows that if ϕ is unbounded, then $\ker \phi$ is a dense subset of X . Furthermore, $\ker \phi$ must be a proper subset of X since ϕ cannot vanish on the whole of X — otherwise it would clearly be bounded. ■

5. Hamel Bases

Definition 5.1. A finite set of elements x_1, \dots, x_n in a complex (real) vector space is said to be linearly independent if and only if

$$\alpha_1 x_1 + \dots + \alpha_n x_n = 0$$

with $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ (or \mathbb{R}) implies that $\alpha_1 = \dots = \alpha_n = 0$. A subset A in a vector space is said to be linearly independent if and only if each finite subset of A is.

Definition 5.2. A linearly independent subset A in a vector space X is called a Hamel basis of X if and only if any non-zero element $x \in X$ can be written as

$$x = \alpha_1 u_1 + \dots + \alpha_m u_m$$

for some $m \in \mathbb{N}$, non-zero $\alpha_1, \dots, \alpha_m \in \mathbb{C}$ (or \mathbb{R}) and distinct elements $u_1, \dots, u_m \in A$.

In other words, A is a Hamel basis of X if it is linearly independent and if any element of X can be written as a *finite* linear combination of elements of A .

Note that if A is a linearly independent subset of X and if $x \in X$ can be written as $x = \alpha_1 u_1 + \dots + \alpha_m u_m$, as above, then this representation is unique. To see this, suppose that we also have that $x = \beta_1 v_1 + \dots + \beta_k v_k$, for non-zero $\beta_1, \dots, \beta_k \in \mathbb{C}$ and distinct elements $v_1, \dots, v_k \in A$. Taking the difference, we have that

$$0 = \alpha_1 u_1 + \dots + \alpha_m u_m - \beta_1 v_1 - \dots - \beta_k v_k.$$

Suppose that $m \leq k$. Now v_1 is not equal to any of the other v_j 's and so, by independence, cannot also be different from all the u_i 's. In other words, v_1 is equal to one of the u_i 's. Similarly, we argue that every v_j is equal to some

u_i and therefore we must have $m = k$ and v_1, \dots, v_m is just a permutation of u_1, \dots, u_m . But then, again by independence, β_1, \dots, β_m is the same permutation of $\alpha_1, \dots, \alpha_m$. The uniqueness of the representation of x as a finite linear combination of elements of A follows.

We will use Zorn's lemma to prove the existence of a Hamel basis.

Theorem 5.3. *Every vector space X possesses a Hamel basis.*

Proof. Let \mathcal{S} denote the collection of linearly independent subsets of X , partially ordered by inclusion. Let $\{S_\alpha : \alpha \in J\}$ be a totally ordered subset of \mathcal{S} . Put $S = \bigcup_\alpha S_\alpha$. We claim that S is linearly independent. To see this, suppose that x_1, \dots, x_m are distinct elements of S and suppose that

$$\lambda_1 x_1 + \dots + \lambda_m x_m = 0$$

for non-zero $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ (or \mathbb{R}). Then $x_1 \in S_{\alpha_1}, \dots, x_m \in S_{\alpha_m}$ for some $\alpha_1, \dots, \alpha_m \in J$. Since $\{S_\alpha\}$ is totally ordered, there is some $\alpha' \in J$ such that $S_{\alpha_1} \subseteq S_{\alpha'}, \dots, S_{\alpha_m} \subseteq S_{\alpha'}$. Hence $x_1, \dots, x_m \in S_{\alpha'}$. But $S_{\alpha'}$ is linearly independent and so we must have that $\lambda_1 = \dots = \lambda_m = 0$. We conclude that S is linearly independent, as claimed.

It follows that S is an upper bound for $\{S_\alpha\}$ in \mathcal{S} . Thus every totally ordered subset in \mathcal{S} has an upper bound and so, by Zorn's lemma, \mathcal{S} possesses a maximal element, M , say. We claim that M is a Hamel basis.

To see this, let $x \in X$, $x \neq 0$, and suppose that an equality of the form

$$x = \lambda_1 u_1 + \dots + \lambda_k u_k$$

is impossible for any $k \in \mathbb{N}$, distinct elements $u_1, \dots, u_k \in M$ and non-zero $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ (or \mathbb{R}). Then, for any distinct $u_1, \dots, u_k \in M$, an equality of the form

$$\lambda x + \lambda_1 u_1 + \dots + \lambda_k u_k = 0$$

must entail $\lambda = 0$. But then this means that $\lambda_1 u_1 = \dots = \lambda_k u_k = 0$, by independence. Hence x, u_1, \dots, u_k are linearly independent. It follows that $M \cup \{x\}$ is linearly independent, which contradicts the maximality of M . We conclude that x can be written as

$$x = \lambda_1 u_1 + \dots + \lambda_m u_m$$

for suitable $m \in \mathbb{N}$, $u_1, \dots, u_m \in M$, and non-zero $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ (or \mathbb{R}); that is, M is a Hamel basis of X . ■

The next result is a corollary of the preceding method of proof.

Theorem 5.4. *Let A be a linearly independent subset of a linear space X . Then there is a Hamel basis of X containing A ; that is, any linearly independent subset of a linear space can be extended to a Hamel basis.*

Proof. Let \mathcal{S} denote the collection of linearly independent subsets of X which contain A . Then \mathcal{S} is partially ordered by set-theoretic inclusion. As above, we apply Zorn's lemma to obtain a maximal element of \mathcal{S} , which is a Hamel basis of X and contains A . ■

The existence of a Hamel basis proves useful in the construction of various “pathological” examples, as we shall see. We first consider the existence of unbounded linear functionals. It is easy to give examples on a normed space. For example, let X be the linear space of those complex sequences which are eventually zero — thus $(a_n) \in X$ if and only if $a_n = 0$ for all sufficiently large n (depending on the particular sequence). Equip X with the norm $\|(a_n)\| = \sup |a_n|$, and define $\phi : X \rightarrow \mathbb{C}$ by $(a_n) \mapsto \phi((a_n)) = \sum_n a_n$. Evidently ϕ is an unbounded linear functional on X . Another example is furnished by the functional $f \mapsto f(0)$ on the normed space $\mathcal{C}([0, 1])$ equipped with the norm $\|f\| = \int_0^1 |f(s)| ds$.

It is not quite so easy, however, to find examples of unbounded linear functionals or everywhere-defined unbounded linear operators on Banach spaces. To do this, we shall use a Hamel basis. Indeed, let X be any infinite-dimensional normed space and let M be a Hamel basis. We define a linear operator $T : X \rightarrow X$ via its action on M as follows. Let u_1, u_2, \dots be any sequence of distinct elements of M , and set

$$Tu_k = ku_k, \quad k = 1, 2, \dots$$

and

$$Tv = 0, \text{ for } v \in M, v \neq u_k \text{ for any } k \in \mathbb{N}.$$

Then if $x \in X$, with $x = \lambda w_1 + \dots + \lambda_m w_m$, $w_j \in M$, $\lambda_1, \dots, \lambda_m \in \mathbb{C}$, we put

$$Tx = \lambda_1 Tw_1 + \dots + \lambda_m Tw_m.$$

It is clear that T is a linear operator defined on the whole of X . Moreover, $\|Tu_k\| = k\|u_k\|$, for any $k \in \mathbb{N}$, and it follows that T is unbounded.

We can refine this a little. Define $Tu_k = ku_k$ as above, but now let $Tv = v$ for any $v \in M$ with $v \neq u_k$, any k . Then T is everywhere defined and unbounded. Furthermore, it is easy to see that $T : X \rightarrow X$ is one-one

and onto. If X is a Banach space, the inverse mapping theorem implies that T^{-1} must also be unbounded.

Now let $\mu : M \rightarrow \mathbb{R}$ be the map $w \mapsto \mu(w) = \|w\|$, $w \in M$. By linearity, we can extend μ to a linear map on X . Then the map $\phi : X \rightarrow \mathbb{C}$ given by $x \mapsto \mu(Tx)$ is an everywhere-defined unbounded linear functional on X .

We can use the concept of Hamel basis to give an example of a space which is a Banach space with respect to two inequivalent norms. It is not difficult to give examples of linear spaces with inequivalent norms. For example, $\mathcal{C}[0, 1]$ equipped with the $\|\cdot\|_\infty$ and $\|\cdot\|_1$ norms is such an example. It is a little harder to find examples where the space is complete with respect to each of the two inequivalent norms. To give such an example, we will use the fact that any Hamel basis for an infinite dimensional, separable linear space has cardinality 2^{\aleph_0} . In fact, all we need to know is that if X and Y are separable, infinite dimensional spaces with Hamel bases M_X and M_Y , respectively, then M_X and M_Y are isomorphic as sets.

Example 5.5. Set $X = \ell^1$ and $Y = \ell^2$ and, for $k \in \mathbb{N}$, let e_k be the element $e_k = (\delta_{km})_{m \in \mathbb{N}}$ of ℓ^1 and let f_k denote the corresponding element of ℓ^2 . For $n \in \mathbb{N}$, let $a_n = \sum_{k=1}^n \frac{1}{n} e_k \in \ell^1$ and let $b_n = \sum_{k=1}^n \frac{1}{n} f_k \in \ell^2$. Then $\|a_n\|_1 = 1$, for all $n \in \mathbb{N}$, whereas $\|b_n\|_2 = 1/\sqrt{n}$. Let $A = \{a_n : n \in \mathbb{N}\}$ and $B = \{b_n : n \in \mathbb{N}\}$. Then A and B are linearly independent subsets of ℓ^1 and ℓ^2 , respectively. Hence they may be extended to Hamel bases M_X and M_Y of X and Y . Since both ℓ^1 and ℓ^2 are separable, M_X and M_Y are isomorphic. It follows that the map defined by $\varphi(a_n) = b_n$, for $n \in \mathbb{N}$, extends to an isomorphism mapping M_X onto M_Y . By linearity, this map extends to an isomorphism, which we denote also by φ , from ℓ^1 onto ℓ^2 .

We define a new norm $\|\cdot\|$ on ℓ^1 by setting

$$\|x\| = \|\varphi(x)\|_2$$

for $x \in X = \ell^1$. This is a norm because φ is linear and injective. To see that X is complete with respect to this norm, suppose that (x_n) is a Cauchy sequence with respect to $\|\cdot\|$. Then $(\varphi(x_n))$ is a Cauchy sequence in ℓ^2 . Since ℓ^2 is complete, there is some $y \in \ell^2$ such that $\|\varphi(x_n) - y\|_2 \rightarrow 0$. Now, φ is surjective and so we may write y as $y = \varphi(x)$ for some $x \in \ell^1$. We have

$$\begin{aligned} \|\varphi(x_n) - y\|_2 &= \|\varphi(x_n) - \varphi(x)\|_2 \\ &= \|x_n - x\| \end{aligned}$$

and it follows that $\|x_n - x\| \rightarrow 0$ as $n \rightarrow \infty$. In other words, ℓ^1 is complete with respect to the norm $\|\cdot\|$.

We claim that the norms $\|\cdot\|_1$ and $\|\cdot\|$ are *not* equivalent norms on ℓ^1 . Indeed, we have that $\varphi(a_n) = b_n$ and so $\|a_n\| = \|b_n\|_2 = 1/\sqrt{n} \rightarrow 0$ as $n \rightarrow \infty$. However, $\|a_n\|_1 = 1$ for all n .

6. Projections

Let X be a linear space and suppose that V and W are subspaces of X such that $V \cap W = \{0\}$ and $X = \text{span}\{V, W\}$; then any $x \in X$ can be written uniquely as $x = v + w$ with $v \in V$ and $w \in W$. In other words, $X = V \oplus W$. Define a map $P : X \rightarrow V$ by $Px = v$, where $x = v + w \in X$, with $v \in V$ and $w \in W$, as above. Evidently, P is a well-defined linear operator satisfying $P^2 = P$. P is called the projection onto V along W . We see that $\text{ran } P = V$ (since $Pv = v$ for all $v \in V$), and also $\ker P = W$ (since if $x = v + w$ and $Px = 0$ then we have $0 = Px = v$ and so $x = w \in W$).

Conversely, suppose that $P : X \rightarrow X$ is a linear operator such that $P^2 = P$, that is, P is an idempotent. Set $V = \text{ran } P$ and $W = \ker P$. Evidently, W is a linear subspace of X . Furthermore, for any given $v \in V$, there is $x \in X$ such that $Px = v$. Hence

$$Pv = P^2x = Px = v$$

and we see that $(\mathbb{1} - P)v = 0$. Hence $V = \ker(\mathbb{1} - P)$ and it follows that V is also a linear subspace of X . Now, any $x \in X$ can be written as $x = Px + (\mathbb{1} - P)x$ with $Px \in V = \text{ran } P$ and $(\mathbb{1} - P)x \in W = \ker P$. We have seen above that for any $v \in V$, we have $v = Pv$. If also $v \in W = \ker P$ then $Pv = 0$, so that $v = Pv = 0$. It follows that $V \cap W = \{0\}$ and so $X = V \oplus W$.

Now suppose that X is a normed space and that $P : X \rightarrow X$ is a bounded linear operator such that $P^2 = P$. Then both $V = \text{ran } P = \ker(\mathbb{1} - P)$ and $W = \ker P$ are closed subspaces of X and $X = V \oplus W$.

Conversely, suppose that X is a Banach space and $X = V \oplus W$, where V and W are closed linear subspaces of X . Define $P : X \rightarrow V$ as above so that $P^2 = P$ and $V = \text{ran } P = \ker(\mathbb{1} - P)$ and $W = \ker P$. We wish to show that P is bounded. To see this we will show that P is closed and then appeal to the closed-graph theorem. Suppose, then, that $x_n \rightarrow x$ and $Px_n \rightarrow y$.

Now, $Px_n \in V$ for each n and V is closed, by hypothesis. It follows that $y \in V$ and so $Py = y$. Furthermore, $(\mathbb{1} - P)x_n = x_n - Px_n \rightarrow x - y$ and $(\mathbb{1} - P)x_n \in W$ for each n and W is closed, by hypothesis. Hence $x - y \in W$ and so $P(x - y) = 0$, that is, $Px = Py$. Hence we have $Px = Py = y$ and we conclude that P is closed. Thus P is a closed linear operator from the Banach space X onto the Banach space V . By the closed-graph theorem, it follows that P is bounded. Note that $\|P\| \geq 1$ unless $\text{ran } P = \{0\}$. Indeed,

$$\begin{aligned}\|P\| &= \sup\{\|Px\| : \|x\| = 1\} \\ &\geq \sup\{\|Pv\| : \|v\| = 1, v \in V\} = 1.\end{aligned}$$

We have therefore proved the following theorem.

Theorem 6.1. *Suppose that V is a closed subspace of a Banach space X . Then there is a closed subspace W such that $X = V \oplus W$ if and only if there exists a bounded idempotent P with $\text{ran } P = V$.*

Definition 6.2. We say that a closed subspace V in a normed space is complemented if there is a closed subspace W such that $X = V \oplus W$.

Theorem 6.3. *Suppose that V is a finite-dimensional subspace of a normed space X . Then V is closed and complemented.*

Proof. Let v_1, \dots, v_m be linearly independent elements of X which span V . Using this basis, we may identify V with \mathbb{C}^m . Moreover, the norm on V is equivalent to the usual Euclidean norm on \mathbb{C}^m . In particular, there is some constant K such that if $v \in V$ is given by $v = \sum_{i=1}^m \alpha_i v_i$ then $\sqrt{\sum_{i=1}^m |\alpha_i|^2} \leq K\|v\|$. Define $\ell_i : V \rightarrow \mathbb{C}$ by linear extension of the rule $\ell_i(v_j) = \delta_{ij}$ for $1 \leq i, j \leq m$. Then, with the notation above, for any $v \in V$,

$$|\ell_i(v)| = |\alpha_i| \leq K\|v\|$$

and so we see that each ℓ_i is a bounded linear functional on V . By the Hahn-Banach theorem, we may extend these to bounded linear functionals on X , which we will also denote by ℓ_i . Then if $v \in V$ is given by $v = \alpha_1 v_1 + \dots + \alpha_m v_m$ we have $\ell_i(v) = \alpha_i$ and so

$$v = \ell_1(v)v_1 + \dots + \ell_m(v)v_m.$$

Define $P : X \rightarrow X$ by

$$Px = \ell_1(x)v_1 + \dots + \ell_m(x)v_m.$$

It is clear that P is a bounded linear operator on X with range equal to V . Also we see that $P^2 = P$. Hence $V = \ker(\mathbb{1} - P)$ is closed, since $(\mathbb{1} - P)$ is bounded, and $W = \ker P$ is a closed complementary subspace for V . We note that $W = \bigcap_{i=1}^m \ker \ell_i$. ■

7. The Dual Space

Let X be a normed space. The space of all bounded linear functionals on X , $\mathcal{B}(X, \mathbb{C})$, is denoted by X^* and called the dual space of X . Since \mathbb{C} is complete, X^* is a Banach space.

The Hahn-Banach theorem assures us that X^* is non-trivial; indeed, X^* separates the points of X . Now, X^* is a normed space in its own right, so we may consider *its* dual, X^{**} ; this is called the bidual or double dual of X .

Let $x \in X$, and consider the mapping

$$\ell \in X^* : \ell \mapsto \ell(x).$$

Evidently, this is a linear map $: X^* \rightarrow \mathbb{C}$. Moreover,

$$|\ell(x)| \leq \|\ell\| \|x\|, \text{ for every } \ell \in X^*$$

so we see that this is a bounded linear map from X^* into \mathbb{C} , that is, it defines an element of X^{**} . In fact, this leads to an isometric embedding of X into X^{**} , as we now show.

Theorem 7.1. *Let X be a normed space, and for $x \in X$, let $\varphi_x : X^* \rightarrow \mathbb{C}$ be the evaluation map $\varphi_x(\ell) = \ell(x)$, $\ell \in X^*$. Then $x \mapsto \varphi_x$ is an isometric linear map of X into X^{**} .*

Proof. We have seen that $\varphi_x \in X^{**}$ for each $x \in X$. It is easy to see that $x \mapsto \varphi_x$ is linear;

$$\begin{aligned} \varphi_{\alpha x + y}(\ell) &= \ell(\alpha x + y) = \alpha \ell(x) + \ell(y) \\ &= \alpha \varphi_x(\ell) + \varphi_y(\ell) \end{aligned}$$

for all $x, y \in X$, and $\alpha \in \mathbb{C}$. Also

$$|\varphi_x(\ell)| = |\ell(x)| \leq \|\ell\| \|x\| \text{ for all } \ell \in X^*$$

shows that $\|\varphi_x\| \leq \|x\|$. However, by the Hahn-Banach theorem, for any given $x \in X$, $x \neq 0$, there is $\ell' \in X^*$ such that $\|\ell'\| = 1$ and $\ell'(x) = \|x\|$. For this particular ℓ' , we then have

$$|\varphi_x(\ell')| = |\ell'(x)| = \|x\| = \|x\| \|\ell'\|.$$

We conclude that $\|\varphi_x\| = \|x\|$, and the proof is complete. \blacksquare

Thus, we may consider X as a subspace of X^{**} via the linear isometric embedding $x \mapsto \varphi_x$.

Definition 7.2. A Banach space X is called reflexive if $X = X^{**}$ via the above embedding.

Note that X^{**} is a Banach space, so X must be a Banach space if we are to be able to identify X with X^{**} .

Theorem 7.3. A Banach space X is reflexive if and only if X^* is reflexive.

Proof. If $X = X^{**}$, then $X^* = X^{***}$. This can be seen as follows. To say that $X = X^{**}$ means that each element of X^{**} has the form ϕ_x for some $x \in X$. Now let $\psi_\ell \in X^{***}$ be the corresponding association of X^* into X^{***} :

$$\psi_\ell(z) = z(\ell) \text{ for } \ell \in X^* \text{ and } z \in X^{**}.$$

We have $X^* \subset X^{***}$ via $\ell \mapsto \psi_\ell$. Let $\lambda \in X^{***}$. Any $z \in X^{**}$ has the form ϕ_x , $x \in X$, and so $\lambda(z) = \lambda(\phi_x)$.

Define $\ell: X \rightarrow \mathbb{C}$ by $\ell(x) = \lambda(\phi_x)$. Then

$$\begin{aligned} |\ell(x)| &= |\lambda(\phi_x)| \\ &\leq \|\lambda\| \|\phi_x\| \\ &= \|\lambda\| \|x\|. \end{aligned}$$

It follows that $\ell \in X^*$. Moreover,

$$\begin{aligned} \psi_\ell(\phi_x) &= \phi_x(\ell) \\ &= \ell(x) = \lambda(\phi_x) \end{aligned}$$

and so $\psi_\ell = \lambda$; i.e., $X^* = X^{***}$ via ψ .

Now suppose that $X^* = X^{***}$ and suppose that $X \neq X^{**}$. Then there is $\lambda \in X^{***}$ such that $\lambda \neq 0$ but λ vanishes on X in X^{**} ; i.e., $\lambda(\phi_x) = 0$ for all $x \in X$.

But then λ can be written as $\lambda = \psi_\ell$ for some $\ell \in X^*$, since $X^* = X^{***}$ and so

$$\begin{aligned}\lambda(\phi_x) &= \psi_\ell(\phi_x) \\ &= \phi_x(\ell) = \ell(x)\end{aligned}$$

which gives $0 = \lambda(\phi_x) = \ell(x)$ for all $x \in X$, i.e., $\ell = 0$ in X^* .

It follows that $\psi_\ell = 0$ in X^{***} and so $\lambda = 0$. This is a contradiction. Hence $X = X^{**}$. ■

Corollary 7.4. *Suppose that the Banach space X is not reflexive. Then the natural inclusions $X \subseteq X^{**} \subseteq X^{****} \subseteq \dots$ and $X^* \subseteq X^{***} \subseteq \dots$ are all strict.*

Proof. If, for sake of argument $X^{***} = X^{*****}$, then this says that X^{***} is reflexive. But this implies that X^{**} is reflexive, which in turn means that X^* is. Finally we conclude that X is reflexive. ■

We shall now compute the duals of some of the classical Banach spaces. For any $p \in \mathbb{R}$ with $1 \leq p < \infty$, the space ℓ^p is the space of complex sequences $x = (x_n)$ such that

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{\frac{1}{p}} < \infty.$$

We shall consider the cases $1 < p < \infty$ and show that for such p , $\|\cdot\|_p$ is a norm and that ℓ^p is a Banach space with respect to this norm. We will also show that the dual of ℓ^p is ℓ^q , where q is given by the formula $\frac{1}{p} + \frac{1}{q} = 1$. It therefore follows that these spaces are reflexive. At this stage, it is not even clear that ℓ^p is a linear space, never mind whether or not $\|\cdot\|_p$ is a norm. We need some classical inequalities.

Proposition 7.5. *Let $a, b \geq 0$ and $\alpha, \beta > 0$ with $\alpha + \beta = 1$. Then*

$$a^\alpha b^\beta \leq \alpha a + \beta b$$

with equality if and only if $a = b$.

Proof. We note that the function $t \mapsto e^t$ is strictly convex; for any $x, y \in \mathbb{R}$, $e^{(\alpha x + \beta y)} < \alpha e^x + \beta e^y$. Putting $a = e^x$, $b = e^y$ gives the required result. ■

The next result we shall need is Hölder's inequality.

Theorem 7.6. *Let $p > 1$ and let q be such that $\frac{1}{p} + \frac{1}{q} = 1$, (q is called the exponent conjugate to p .) Then, for any $x = (x_n) \in \ell^p$ and $y = (y_n) \in \ell^q$,*

$$\sum_{n=1}^{\infty} |x_n y_n| \leq \|x\|_p \|y\|_q.$$

If $p = 1$, the above is valid if we set $q = \infty$.

Proof. The case for $p = 1$ and $q = \infty$ is easy to see. So suppose that $p > 1$. Without loss of generality, we may suppose that $\|x\|_p = \|y\|_q = 1$. Then let $\alpha = \frac{1}{p}$, $\beta = \frac{1}{q}$, $a = |x_n|^p$, $b = |y_n|^q$ and use the previous proposition. ■

Proposition 7.7. *For any $x = (x_n) \in \ell^p$, with $p > 1$,*

$$\|x\|_p = \sup \left\{ \left| \sum_{n=1}^{\infty} x_n y_n \right| : \|y\|_q = 1 \right\}.$$

The equality holds for $p = 1$ and $q = \infty$, and also for the pair $p = \infty$, $q = 1$.

Proof. Hölder's inequality implies that the right hand side is not greater than the left hand side. For the converse, consider $y = (y_n)$ with $y_n = \overline{\operatorname{sgn} x_n} |x_n|^{p/q} / \|x\|_p^{p/q}$ if $1 < p < \infty$, with $y_n = \overline{\operatorname{sgn} x_n}$ if $p = 1$, and with $y_n = \delta_{nm}$, $m \in \mathbb{N}$, if $p = \infty$. ■

As an immediate corollary, we obtain Minkowski's inequality.

Corollary 7.8. *For any $x, y \in \ell^p$, $p \geq 1$, we have $x + y \in \ell^p$ and*

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

Proof. This follows directly from the triangle inequality and the preceding proposition. ■

Theorem 7.9. For any $1 \leq p \leq \infty$, ℓ^p is a Banach space. Moreover, if $1 \leq p < \infty$, the dual of ℓ^p is ℓ^q , where q is the exponent conjugate to p . Furthermore, for each $1 < p < \infty$, the space ℓ^p is reflexive.

Proof. We have already discussed these spaces for $p = 1$ and $p = \infty$. For the rest, it follows from the preceding results that ℓ^p is a linear space and that $\|\cdot\|_p$ is a norm on ℓ^p . The completeness of ℓ^p , for $1 < p < \infty$, follows in much the same way as that of the proof for $p = 1$.

To show that $\ell^{p*} = \ell^q$, we use the pairing as in Hölder's inequality. Indeed, for any $y = (y_n) \in \ell^q$, define ψ_y on ℓ^p by $\psi_y : x = (x_n) \mapsto \sum_n x_n y_n$. Then Hölder's inequality implies that ψ_y is a bounded linear functional on ℓ^p and the subsequent proposition (with the rôles of p and q interchanged) shows that $\|\psi_y\| = \|y\|_q$.

To show that every bounded linear functional on ℓ^p has the above form, for some $y \in \ell^q$, let $\lambda \in \ell^{p*}$, where $1 \leq p < \infty$. Let $y_n = \lambda(e_n)$, where $e_n = (\delta_{nm})_{m \in \mathbb{N}} \in \ell^p$. Then for any $x = (x_n) \in \ell^p$,

$$\lambda(x) = \lambda\left(\sum_n x_n e_n\right) = \left(\sum_n x_n \lambda(e_n)\right) = \sum_n x_n y_n.$$

Hence, replacing x_n by $\overline{\operatorname{sgn}(x_n y_n)} x_n$, we see that

$$\sum_n |x_n y_n| \leq \|\lambda\| \|x\|_p.$$

For any $N \in \mathbb{N}$, denote by y' the truncated sequence $(y_1, y_2, \dots, y_N, 0, 0, \dots)$. Then

$$\sum_n |x_n y'_n| \leq \|\lambda\| \|x\|_p$$

and, taking the supremum over x with $\|x\|_p = 1$, we obtain the estimate

$$\|y'\|_q \leq \|\lambda\|.$$

It follows that $y \in \ell^q$ (— and that $\|y\|_q \leq \|\lambda\|$). But then, by definition, $\psi_y = \lambda$, and we deduce that $y \mapsto \psi_y$ is an isometric mapping onto ℓ^{p*} . Thus, the association $y \mapsto \psi_y$ is an isometric linear isomorphism between ℓ^q and ℓ^{p*} .

Finally, we note that the above discussion shows that ℓ^p is reflexive, for all $1 < p < \infty$. ■

We shall now consider c_0 , the linear space of all complex sequences which converge to 0, equipped with the supremum norm

$$\|x\|_\infty = \sup\{|x_n| : n \in \mathbb{N}\}, \text{ for } x = (x_n) \in c_0.$$

One checks that c_0 is a Banach space (— a closed subspace of ℓ^∞). We shall show that the dual of c_0 is ℓ^1 , that is, there is an isometric isomorphism between c_0^* and ℓ^1 . To see this, suppose first that $z = (z_n) \in \ell^1$. Define $\psi_z : c_0 \rightarrow \mathbb{C}$ to be $\psi : x \mapsto \sum_n z_n x_n$, $x = (x_n) \in c_0$. It is clear that ψ_z is well-defined for any $z \in \ell^1$ and that

$$|\psi_z(x)| \leq \sum_n |z_n| |x_n| \leq \|z\|_1 \|x\|_\infty.$$

Thus we see that ψ_z is a bounded linear functional with norm $\|\psi_z\| \leq \|z\|_1$. By taking x to be the element of c_0 whose first m terms are equal to 1, and whose remaining terms are 0, we see that $\|\psi_z\| \geq \sum_{n=1}^m |z_n|$. It follows that $\|\psi_z\| = \|z\|_1$, and therefore $z \mapsto \psi_z$ is an isometric mapping of ℓ^1 into c_0^* . We shall show that every element of c_0^* is of this form and hence $z \mapsto \psi_z$ is onto.

To see this, let $\lambda \in c_0^*$, and, for $n \in \mathbb{N}$, let $z_n = \lambda(e_n)$, where $e_n \in c_0$ is the sequence all of whose terms are zero except for the n^{th} term which is equal to 1, i.e., e_n is the sequence $(\delta_{nm})_{m \in \mathbb{N}}$. For any given $N \in \mathbb{N}$, let

$$v = \sum_{k=1}^N \overline{\text{sgn } z_k} e_k.$$

Then $v \in c_0$ and $\|v\|_\infty = 1$ and

$$|\lambda(v)| = \sum_{k=1}^N |z_k| \leq \|\lambda\| \|v\|_\infty = \|\lambda\|.$$

It follows that $z = (z_n) \in \ell^1$ and that $\|z\|_1 \leq \|\lambda\|$. Furthermore, for any element $x = (x_n) \in c_0$,

$$\begin{aligned} \lambda(x) &= \lambda\left(\sum_{n=1}^{\infty} x_n e_n\right) \\ &= \left(\sum_{n=1}^{\infty} x_n \lambda(e_n)\right) \\ &= \psi_z(x). \end{aligned}$$

Hence $\lambda = \psi_z$, and the proof is complete.

Remark 7.10. Exactly as above, we see that the map $z \mapsto \psi_z$ is a linear isometric mapping of ℓ^1 into the dual of ℓ^∞ . Furthermore, if λ is any element of the dual of ℓ^∞ , then, in particular, it defines a bounded linear functional on c_0 . Thus, *the restriction* of λ to c_0 is of the form ψ_z for some $z \in \ell^1$. It does *not* follow, however, that λ has this form on the whole of ℓ^∞ . Indeed, c_0 is a closed linear subspace of ℓ^∞ , and, for example, the element $y = (y_n)$, where $y_n = 1$ for all $n \in \mathbb{N}$, is an element of ℓ^∞ which is not an element of c_0 . Then we know, from the Hahn-Banach theorem, that there is a bounded linear functional λ , say, such that $\lambda(x) = 0$ for all $x \in c_0$ and such that $\lambda(y) = 1$. Thus, λ is an element of the dual of ℓ^∞ which is clearly *not* determined by an element of ℓ^1 . The dual of ℓ^∞ is strictly larger than ℓ^1 .

Theorem 7.11. *Suppose that X is a Banach space and that X^* is separable. Then X is separable.*

Proof. Let $\{\lambda_n : n = 1, 2, \dots\}$ be a countable dense subset of X^* . For each $n \in \mathbb{N}$, let $x_n \in X$ be such that $\|x_n\| = 1$ and $|\lambda_n(x_n)| \geq \frac{1}{2}\|\lambda_n\|$. Let S be the set of finite linear combinations of the x_n 's with rational complex coefficients. Then S is countable. We claim that S is dense in X . To see this, suppose the contrary; that is, suppose that \overline{S} is a proper closed linear subspace of X . Then there exists a non-zero bounded linear functional $\Lambda \in X^*$ such that Λ vanishes on \overline{S} . Since $\Lambda \in X^*$ and $\{\lambda_n : n \in \mathbb{N}\}$ is dense in X^* , there is some subsequence (λ_{n_k}) such that $\lambda_{n_k} \rightarrow \Lambda$ as $k \rightarrow \infty$, that is,

$$\|\Lambda - \lambda_{n_k}\| \rightarrow 0$$

as $n \rightarrow \infty$. However,

$$\begin{aligned} \|\Lambda - \lambda_{n_k}\| &\geq |(\Lambda - \lambda_{n_k})(x_{n_k})|, \text{ since } \|x_{n_k}\| = 1, \\ &= |\lambda_{n_k}(x_{n_k})|, \text{ since } \Lambda \text{ vanishes on } \overline{S}, \\ &\geq \frac{1}{2}\|\lambda_{n_k}\| \end{aligned}$$

and so it follows that $\|\lambda_{n_k}\| \rightarrow 0$, as $k \rightarrow \infty$. But $\lambda_{n_k} \rightarrow \Lambda$ implies that $\|\lambda_{n_k}\| \rightarrow \|\Lambda\|$ and therefore $\|\Lambda\| = 0$. This forces $\Lambda = 0$, which is a contradiction. We conclude that \overline{S} is dense in X and that, consequently, X is separable. ■

Theorem 7.12. For $1 \leq p < \infty$ the space ℓ^p is separable, but the space ℓ^∞ is non-separable.

Proof. Let S denote the set of sequences of complex numbers (z_n) such that (z_n) is eventually zero (i.e., $z_n = 0$ for all sufficiently large n , depending on the sequence) and such that z_n has rational real and imaginary parts, for all n . Then S is a countable set and it is straightforward to verify that S is dense in each ℓ^p , for $1 \leq p < \infty$.

Note that S is also a subset of ℓ^∞ , but it is not a dense subset. Indeed, if x denotes that element of ℓ^∞ all of whose terms are equal to 1, then $\|x - \zeta\|_\infty \geq 1$ for any $\zeta \in S$.

To show that ℓ^∞ is not separable, consider the subset A of elements whose components consist of the numbers $0, 1, \dots, 9$. Then A is uncountable and the distance between any two distinct elements of A is at least 1. It follows that the balls $\{x : \|x - a\|_\infty < 1/2\} : a \in A\}$ are pairwise disjoint. Now, if B is any dense subset of ℓ^∞ , each ball will contain an element of B , and these will all be distinct. It follows that B must be uncountable. ■

Remark 7.13. The example of ℓ^1 shows that a separable Banach space need not have a separable dual—we have seen that the dual of ℓ^1 is ℓ^∞ , which is not separable. This also shows that ℓ^1 is not reflexive. Indeed, this would require that ℓ^1 be isometrically isomorphic to the dual of ℓ^∞ . Since ℓ^1 is separable, an application of the earlier theorem would lead to the false conclusion that ℓ^∞ is separable.

8. Topological Spaces

Definition 8.1. Let X be a nonempty set and suppose that \mathcal{T} is a collection of subsets of X . \mathcal{T} is called a topology on X if the following hold;

- (i) $\emptyset \in \mathcal{T}$, and $X \in \mathcal{T}$;
- (ii) if $\{U_\alpha : \alpha \in J\}$ is an arbitrary collection of elements of \mathcal{T} , labelled by J , then $\bigcup_{\alpha \in J} U_\alpha \in \mathcal{T}$;
- (iii) if, for any $k \in \mathbb{N}$, $U_1, U_2, \dots, U_k \in \mathcal{T}$, then $\bigcap_{i=1}^k U_i \in \mathcal{T}$.

The elements of \mathcal{T} are called *open sets*, or \mathcal{T} -open sets. The pair (X, \mathcal{T}) is called a topological space.

Examples 8.2.

1. $\mathcal{T} = \{\emptyset, X\}$ — called the indiscrete topology.
2. \mathcal{T} is the set of all subsets of X — discrete topology.
3. $X = \{0, 1, 2\}$ and $\mathcal{T} = \{\emptyset, X, \{0\}, \{1, 2\}\}$.
4. X any metric space, \mathcal{T} the set of open sets in the usual metric space sense.

Thus a topological space is a generalization of a metric space.

Definition 8.3. For any non-empty subset A of a topological space (X, \mathcal{T}) , the induced topology, \mathcal{T}_A , on A is defined to be that given by the collection $A \cap \mathcal{T} = \{A \cap U : U \in \mathcal{T}\}$ of subsets of A . (It is readily verified that \mathcal{T}_A is a topology on A .)

Definition 8.4. A topological space (X, \mathcal{T}) is said to be metrizable if there is a metric on X such that \mathcal{T} is as in example 4 above.

Remark 8.5. Not every topology is metrizable. For example, example 1, above (—provided X consists of more than one point).

Many of the usual concepts in metric space theory appear in the theory of topological spaces — but suitably rephrased in terms of open sets.

Definition 8.6. A subset F of a topological space X is said to be closed if and only if its complement $X \setminus F$ is open, i.e., belongs to \mathcal{T} .

A point $a \in X$ is an interior point of a set $A \subseteq X$ if there is $U \in \mathcal{T}$ such that $a \in U$ and $U \subseteq A$. (Thus, a set G is open if and only if each of its points is an interior point of G .)

The set of interior points of the set A is denoted by $\overset{\circ}{A}$.

The point x is a limit point (accumulation point) of the set A if and only if for every open set U , with $x \in U$, it is true that $U \cap A$ contains some point distinct from x , i.e., the set $A \cap \{U \setminus \{x\}\} \neq \emptyset$.

The point $a \in A$ is said to be an isolated point of A if there is an open set U such that $a \in U$ but $U \cap \{A \setminus \{a\}\} = \emptyset$.

The closure of the set A , written \overline{A} , is the union of A and its set of limit points,

$$\overline{A} = A \cup \{\text{limit points of } A\}.$$

Proposition 8.7. The closure \overline{A} of A is the smallest closed set containing A , that is,

$$\overline{A} = \bigcap \{F : F \text{ is closed and } F \supseteq A\}.$$

Proof. We shall first show that \overline{A} is closed. Let $y \in X \setminus \overline{A}$. Then $y \notin A$, and y is not a limit point of A . Hence there is an open set U such that $y \in U$ and $U \cap A = \emptyset$. But then no point of U can be a limit point of A so we deduce that $U \cap \overline{A} = \emptyset$. It follows that $X \setminus \overline{A}$ is open and so, by definition, \overline{A} is closed.

Now suppose that F is closed and that $A \subseteq F$. Then $X \setminus F$ is open. Let $z \in X \setminus F$. Then there is some open set U such that $z \in U$ and $U \subseteq X \setminus F$. In particular, $U \cap A = \emptyset$, and so $z \notin A$ and z is not a limit point of A . Hence $X \setminus F \subseteq X \setminus \overline{A}$ and we see that $\overline{A} \subseteq F$. The result follows. ■

Definition 8.8. A family of open sets $\{U_\alpha : \alpha \in J\}$ is said to be an open cover of a set $B \subseteq X$ if $B \subseteq \bigcup_\alpha U_\alpha$.

Definition 8.9. A subset K in a topological space is said to be compact if every open cover of K contains a finite subcover.

By taking complements, open sets become closed sets, unions are replaced by intersections and the notion of compactness can be rephrased as follows.

Theorem 8.10. *Let K be a subset of a topological space (X, \mathcal{T}) . The following statements are equivalent.*

- (i) K is compact.
- (ii) If $\{F_\alpha\}_{\alpha \in J}$ is any family of closed sets in X such that $K \cap \bigcap_{\alpha \in J} F_\alpha = \emptyset$, then $K \cap \bigcap_{\alpha \in I} F_\alpha = \emptyset$ for some finite subset $I \subseteq J$.
- (iii) If $\{F_\alpha\}_{\alpha \in J}$ is any family of closed sets in X such that $K \cap \bigcap_{\alpha \in I} F_\alpha \neq \emptyset$, for every finite subset $I \subseteq J$, then $K \cap \bigcap_{\alpha \in J} F_\alpha \neq \emptyset$.

Remark 8.11. The statements (ii) and (iii) are contrapositives. The property in statement (iii) is called the finite intersection property (of the family $\{F_\alpha\}_{\alpha \in J}$ of closed sets).

Definition 8.12. A set N is a neighbourhood of a point x in a topological space (X, \mathcal{T}) if and only if there is $U \in \mathcal{T}$ such that $x \in U$ and $U \subseteq N$.

Thus, a set U belongs to \mathcal{T} if and only if U is a neighbourhood of each of its points. Note that N need not itself be open. For example, in any metric space (X, d) , the closed sets $\{z \in X : d(a, z) \leq r\}$, for $r > 0$, are neighbourhoods of the point a .

Definition 8.13. A topological space (X, \mathcal{T}) is said to be a Hausdorff topological space if and only if for any pair of distinct points $x, y \in X$, ($x \neq y$), there exist sets $U, V \in \mathcal{T}$ such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

We can paraphrase the Hausdorff property by saying that any pair of distinct points can be separated by disjoint open sets. Example 3 above is an example of a non-Hausdorff topological space.

Proposition 8.14. *A non-empty subset A of the topological space (X, \mathcal{T}) is compact if and only if A is compact with respect to the induced topology, that is, if and only if (A, \mathcal{T}_A) is compact.*

If (X, \mathcal{T}) is Hausdorff then so is (A, \mathcal{T}_A) .

Proof. Suppose first that A is compact in (X, \mathcal{T}) , and let $\{G_\alpha\}$ be an open cover of A in (A, \mathcal{T}_A) . Then each G_α has the form $G_\alpha = A \cap U_\alpha$ for some $U_\alpha \in \mathcal{T}$. It follows that $\{U_\alpha\}$ is an open cover of A in (X, \mathcal{T}) . By hypothesis, there is a finite subcover, U_1, \dots, U_n , say. But then G_1, \dots, G_n is an open cover of A in (A, \mathcal{T}_A) ; that is, (A, \mathcal{T}_A) is compact.

Conversely, suppose that (A, \mathcal{T}_A) is compact. Let $\{U_\alpha\}$ be an open cover of A in (X, \mathcal{T}) . Set $G_\alpha = A \cap U_\alpha$. Then $\{G_\alpha\}$ is an open cover of (A, \mathcal{T}_A) . By hypothesis, there is a finite subcover, say, G_1, \dots, G_m . Clearly, U_1, \dots, U_m is an open cover for A in (X, \mathcal{T}) . That is, A is compact in (X, \mathcal{T}) .

Suppose that (X, \mathcal{T}) is Hausdorff, and let a_1, a_2 be any two distinct points of A . Then there is a pair of disjoint open sets U, V in X such that $a_1 \in U$ and $a_2 \in V$. Evidently, $G_1 = A \cap U$ and $G_2 = A \cap V$ are open in (A, \mathcal{T}_A) , are disjoint, and $a_1 \in G_1$ and $a_2 \in G_2$. Hence (A, \mathcal{T}_A) is Hausdorff, as claimed. ■

Remark 8.15. Note that it is quite possible for (A, \mathcal{T}_A) to be Hausdorff whilst (X, \mathcal{T}) is not. A simple example is provided by example 3 above with A given by $A = \{0, 1\}$. In this case, the induced topology on A coincides with the discrete topology on A .

Proposition 8.16. *Let (X, \mathcal{T}) be a Hausdorff topological space and let $K \subseteq X$ be compact. Then K is closed.*

Proof. Let $z \in X \setminus K$. Then for each $x \in K$, there are open sets U_x, V_x such that $x \in U_x$, $z \in V_x$ and $U_x \cap V_x = \emptyset$. Evidently, $\{U_x : x \in K\}$ is an open cover of K and therefore there is a finite number of points $x_1, x_2, \dots, x_n \in K$ such that $K \subseteq U_{x_1} \cup \dots \cup U_{x_n}$.

Put $V = V_{x_1} \cap \dots \cap V_{x_n}$. Then V is open, and $z \in V$. Furthermore, $V \subseteq V_{x_i}$ for each i implies that $V \cap U_{x_i} = \emptyset$ for $1 \leq i \leq n$. Hence $V \cap K = \emptyset$, and therefore $z \in V$ and $V \subseteq X \setminus K$. Thus, $X \setminus K$ is open, and K is closed. ■

Example 8.17. Let $X = \{0, 1, 2\}$ and $\mathcal{T} = \{\emptyset, X, \{0\}, \{1, 2\}\}$. Then the set $K = \{2\}$ is compact, but $X \setminus K = \{0, 1\}$ is not an element of \mathcal{T} . Thus, K is not closed. As we have already noted, (X, \mathcal{T}) is not Hausdorff.

Proposition 8.18. *Let (X, \mathcal{T}) be a topological space and let K be compact. Suppose that F is closed and $F \subseteq K$. Then F is compact. (In other words, closed subsets of compact sets are compact.)*

Proof. Let $\{U_\alpha : \alpha \in J\}$ be any given open cover of F . We augment this collection by the open set $X \setminus F$. This gives an open cover of K ;

$$K \subseteq (X \setminus F) \cup \left(\bigcup_{\alpha \in J} U_\alpha \right).$$

Since K is compact, there are elements $\alpha_1, \alpha_2, \dots, \alpha_m$ in J such that

$$K \subseteq (X \setminus F) \cup U_{\alpha_1} \cup U_{\alpha_2} \cup \dots \cup U_{\alpha_m}.$$

But then we see that

$$F \subseteq U_{\alpha_1} \cup U_{\alpha_2} \cup \dots \cup U_{\alpha_m}.$$

We conclude that F is compact. ■

We now consider continuity of mappings between topological spaces. The definition is the obvious rewriting of the standard result from metric space theory.

Definition 8.19. Let (X, \mathcal{T}) and (Y, \mathcal{S}) be topological spaces and suppose that $f : X \rightarrow Y$ is a given mapping. We say that f is continuous if and only if $f^{-1}(V) \in \mathcal{T}$ for any $V \in \mathcal{S}$.

Many of the standard results concerning continuity in metric spaces have analogues in this more general setting.

Theorem 8.20. Let (X, \mathcal{T}) and (Y, \mathcal{S}) be topological spaces with (X, \mathcal{T}) compact. Suppose that $f : X \rightarrow Y$ is a continuous surjection. Then (Y, \mathcal{S}) is compact. (In other words, the image of a compact space under a continuous mapping is compact.)

Proof. Let $\{V_\alpha\}$ be any given open cover of Y . Then $\{f^{-1}(V_\alpha)\}$ is an open cover of X . Hence there are indices $\alpha_1, \alpha_2, \dots, \alpha_m$ such that

$$X = f^{-1}(V_{\alpha_1}) \cup \dots \cup f^{-1}(V_{\alpha_m}).$$

Since f is onto, it follows that

$$Y = V_{\alpha_1} \cup \dots \cup V_{\alpha_m}$$

and so Y is compact. ■

Theorem 8.21. Let X be a compact topological space, Y a Hausdorff topological space and $f : X \rightarrow Y$ a continuous injective surjection. Then $f^{-1} : Y \rightarrow X$ exists and is continuous (—and so X is also Hausdorff).

Proof. Clearly f^{-1} exists as a mapping from Y onto X . Let F be a closed subset of X . To show that f^{-1} is continuous, it is enough to show that $f(F)$ is closed in Y . Now, F is compact in X and, as above, it follows that $f(F)$ is compact in Y . But Y is Hausdorff and so $f(F)$ is closed in Y . ■

Definition 8.22. A continuous bijection $f : X \rightarrow Y$, between topological spaces (X, \mathcal{T}) and (Y, \mathcal{S}) , with a continuous inverse is called a homeomorphism.

A homeomorphism $f : X \rightarrow Y$ sets up a one-one correspondence between the open sets in X and those in Y , via $U \longleftrightarrow f(U)$. The previous theorem says that a continuous bijection from a compact space onto a Hausdorff space is a homeomorphism. It follows that both spaces are compact and Hausdorff.

Definition 8.23. Let $\mathcal{T}_1, \mathcal{T}_2$ be topologies on a set X . We say that \mathcal{T}_1 is weaker (or coarser or smaller) than \mathcal{T}_2 if $\mathcal{T}_1 \subseteq \mathcal{T}_2$ (—alternatively, we say that \mathcal{T}_2 is stronger (or finer or larger) than \mathcal{T}_1).

The *stronger* (or finer) a topology the *more* open sets there are. It is immediately clear that if $f : (X, \mathcal{T}) \rightarrow (Y, \mathcal{S})$ is continuous, then f is also continuous with respect to any topology \mathcal{T}' on X which is *stronger* than \mathcal{T} , or any topology \mathcal{S}' on Y which is *weaker* than \mathcal{S} . In particular, if X has the discrete topology or Y has the indiscrete topology, then *every* map $f : X \rightarrow Y$ is continuous.

Let X be a given (non-empty) set, let (Y, \mathcal{S}) a topological space and let $f : X \rightarrow Y$ be a given map. We wish to investigate topologies on X which make f continuous. Now, if f is to be continuous, then $f^{-1}(V)$ should be open in X for all V open in Y . Let $\mathcal{T} = \bigcap \mathcal{T}'$, where the intersection is over all topologies \mathcal{T}' on X which contain all the sets $f^{-1}(V)$, for $V \in \mathcal{S}$. (The discrete topology on X is one such.) Then \mathcal{T} is a topology on X (—any intersection of topologies is also a topology). Moreover, \mathcal{T} is evidently the weakest topology on X with respect to which f is continuous. We can generalise this to an arbitrary collection of functions. Suppose that $\{(Y_\alpha, \mathcal{S}_\alpha) : \alpha \in I\}$ is a collection of topological spaces, indexed by I , and that $\mathcal{F} = \{f_\alpha : X \rightarrow Y_\alpha\}$ is a family of maps from X into the topological spaces $(Y_\alpha, \mathcal{S}_\alpha)$. Let \mathcal{T} be the intersection of all those topologies on X which contain all sets of the form $f_\alpha^{-1}(V_\alpha)$, for $f_\alpha \in \mathcal{F}$ and $V_\alpha \in \mathcal{S}_\alpha$. Then \mathcal{T} is a topology on X and it is the weakest topology on X with respect to which every $f_\alpha \in \mathcal{F}$ is continuous.

\mathcal{T} is called the $\sigma(X, \mathcal{F})$ -topology on X .

Theorem 8.24. Suppose that each $(Y_{\alpha_0}, \mathcal{S}_{\alpha_0})$ is Hausdorff and that \mathcal{F} separates points of X , i.e., if $a, b \in X$ with $a \neq b$, then there is some $f_\alpha \in \mathcal{F}$ such that $f_\alpha(a) \neq f_\alpha(b)$. Then the $\sigma(X, \mathcal{F})$ -topology is Hausdorff.

Proof. Suppose that $a, b \in X$, with $a \neq b$. Then, by hypothesis, there is some $\alpha \in I$ such that $f_\alpha(a) \neq f_\alpha(b)$. Since $(Y_\alpha, \mathcal{S}_\alpha)$ is Hausdorff, there exist elements $U, V \in \mathcal{S}_\alpha$ such that $f_\alpha(a) \in U$, $f_\alpha(b) \in V$ and $U \cap V = \emptyset$. But then $f_\alpha^{-1}(U)$ and $f_\alpha^{-1}(V)$ are open with respect to the $\sigma(X, \mathcal{F})$ -topology and $a \in f_\alpha^{-1}(U)$, $b \in f_\alpha^{-1}(V)$ and $f_\alpha^{-1}(U) \cap f_\alpha^{-1}(V) = \emptyset$. ■

To describe the $\sigma(X, \mathcal{F})$ -topology somewhat more explicitly, it is convenient to introduce some terminology.

Definition 8.25. A collection \mathcal{B} of open sets is said to be a base for the topology \mathcal{T} on a space X if and only if each element of \mathcal{T} can be written as a union of elements of \mathcal{B} .

Examples 8.26.

1. The open sets $\{x : d(x, a) < r\}$, $a \in X$, $r \in \mathbb{Q}$, $r > 0$, are a base for the usual topology in a metric space (X, d) .
2. The rectangles $\{(x, y) \in \mathbb{R}^2 : |x - a| < \frac{1}{n}, |y - b| < \frac{1}{m}\}$, with $(a, b) \in \mathbb{R}^2$ and $n, m \in \mathbb{N}$, form a base for the usual Euclidean topology on \mathbb{R}^2 .
3. The singleton sets $\{x\}$, $x \in X$, form a base for the discrete topology on any non-empty set X .

Proposition 8.27. *The collection of open sets \mathcal{B} is a base for the topology \mathcal{T} on a space X if and only if for each non-empty set $G \in \mathcal{T}$ and $x \in G$ there is some $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq G$.*

Proof. Suppose that \mathcal{B} is a base for the topology \mathcal{T} and suppose that $G \in \mathcal{T}$ is non-empty. Then G can be written as a union of elements of \mathcal{B} . In particular, for any $x \in G$, there is some $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq G$.

Conversely, suppose that for any non-empty set $G \in \mathcal{T}$ and for any $x \in G$, there is some $B_x \in \mathcal{B}$ such that $x \in B_x$ and $B_x \subseteq G$. Then $G \subseteq \bigcup_{x \in G} B_x \subseteq G$, which shows that \mathcal{B} is a base for \mathcal{T} . ■

Definition 8.28. A collection \mathcal{S} of subsets of a topology \mathcal{T} on X is said to be a sub-base for \mathcal{T} if and only if the collection of intersections of finite families of members of \mathcal{S} is a base for \mathcal{T} .

Example 8.29. The collection of subsets of \mathbb{R} consisting of those intervals of the form (a, ∞) or $(-\infty, b)$, $a, b \in \mathbb{R}$, is a sub-base for the usual topology on \mathbb{R} .

Proposition 8.30. *Let X be any non-empty set and let \mathcal{S} be any collection of subsets of X which covers X , i.e., for any $x \in X$, there is some $A \in \mathcal{S}$ such that $x \in A$. Let \mathcal{B} be the collection of intersections of finite families of elements of \mathcal{S} . Then the collection \mathcal{T} of subsets of X consisting of \emptyset together with arbitrary unions of elements of members of \mathcal{B} is a topology on X , and is the weakest topology on X containing the collection of sets \mathcal{S} . Moreover, \mathcal{S} is a sub-base for \mathcal{T} , and \mathcal{B} is a base for \mathcal{T} .*

Proof. Clearly, $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$, and any union of elements of \mathcal{T} is also a member of \mathcal{T} . It remains to show that any finite intersection of elements of \mathcal{T} is also an element of \mathcal{T} . It is enough to show that if $A, B \in \mathcal{T}$, then $A \cap B \in \mathcal{T}$. If A or B is the empty set, there is nothing more to prove, so suppose that

$A \neq \emptyset$ and $B \neq \emptyset$. Then we have that $A = \bigcup_{\alpha} A_{\alpha}$ and $B = \bigcup_{\beta} B_{\beta}$ for families of elements $\{A_{\alpha}\}$ and $\{B_{\beta}\}$ belonging to \mathcal{B} . Thus

$$A \cap B = \bigcup_{\alpha} A_{\alpha} \cap \bigcup_{\beta} B_{\beta} = \bigcup_{\alpha, \beta} (A_{\alpha} \cap B_{\beta}).$$

Now, each A_{α} is an intersection of a finite number of elements of \mathcal{S} , and the same is true of B_{β} . It follows that the same is true of every $A_{\alpha} \cap B_{\beta}$, and so we see that $A \cap B \in \mathcal{T}$, which completes the proof that \mathcal{T} is a topology on X .

If \mathcal{T}' is any topology on X which contains the collection \mathcal{S} , then certainly \mathcal{T}' must also contain \mathcal{B} . But then \mathcal{T}' must contain arbitrary unions of families of subsets of \mathcal{B} , that is, \mathcal{T}' must contain \mathcal{T} . It follows that \mathcal{T} is the weakest topology on X containing \mathcal{S} . From the definitions, it is clear that \mathcal{S} is a sub-base for \mathcal{T} and that \mathcal{B} is a base for \mathcal{T} . ■

Remark 8.31. The topology \mathcal{T} can also be described as follows: a non-empty set G belongs to \mathcal{T} if and only if for any $x \in G$ there is some $S \in \mathcal{B}$ such that $x \in S$ and $S \subseteq G$. Indeed, if G has this property then it is clearly a union of members of \mathcal{S} . Conversely, if G is a union of elements $\{S_{\alpha}\}$, say, of \mathcal{S} and $x \in G$, then certainly there is some α_0 such that $x \in S_{\alpha_0}$. Evidently, we also have that $S_{\alpha_0} \subseteq G$.

Remark 8.32. We can therefore describe the $\sigma(X, \mathcal{F})$ -topology on X determined by the family of maps $\{f_{\alpha} : \alpha \in I\}$, discussed earlier, as the topology with sub-base given by the collection $\{f_{\alpha}^{-1}(V) : \alpha \in I, V \in \mathcal{S}_{\alpha}\}$.

9. Weak and weak*-topologies

Consider a normed space X with dual space X^* . In particular, these are both topological spaces with respect to the topologies induced by the norms. We wish to consider topologies different from these norm topologies. First we will see how we can use X^* to define a topology on X . The dual, X^* , of X is a collection of (all bounded linear) maps from X into the same topological space, \mathbb{C} . Thus, we can consider the $\sigma(X, X^*)$ -topology on X —called the weak topology on the normed space X . It is Hausdorff since X^* separates the points of X .

The weak topology on X is therefore the weakest topology on X making every member of the dual, X^* , continuous. A non-empty set G in X is open with respect to the weak topology if and only if for each point $a \in G$ there is B such that $a \in B$ and $B \subseteq G$, and where B has the form

$$B = \ell_1^{-1}(U_1) \cap \cdots \cap \ell_n^{-1}(U_n)$$

for some $n \in \mathbb{N}$, $\ell_1, \dots, \ell_n \in X^*$ and U_1, \dots, U_n open sets in \mathbb{C} . (This is just the statement that the sets of the form $\ell(U)$, with $\ell \in X^*$ and U open in \mathbb{C} form a sub-base.) Now, $a \in \ell_j^{-1}(U_j)$ is equivalent to $\ell_j(a) \in U_j$, and if U_j is open in \mathbb{C} , there is some $\varepsilon_j > 0$ such that the open ball $B(\ell_j(a); \varepsilon_j)$ is contained in U_j . It follows that the set G is open with respect to the weak topology if and only if there is B as above but of the form

$$B = \ell_1^{-1}(B(\ell_1(a); \varepsilon_1)) \cap \cdots \cap \ell_n^{-1}(B(\ell_n(a); \varepsilon_n))$$

By taking $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_n\}$, this is equivalent to the existence of some B as above of the form

$$B = \ell_1^{-1}(B(\ell_1(a); \varepsilon)) \cap \cdots \cap \ell_n^{-1}(B(\ell_n(a); \varepsilon)).$$

However,

$$\ell_j^{-1}(B(\ell_j(a); \varepsilon)) = \{x \in X : |\ell_j(x) - \ell_j(a)| < \varepsilon\}$$

so that

$$\begin{aligned} \ell_1^{-1}(B(\ell_1(a); \varepsilon)) \cap \cdots \cap \ell_n^{-1}(B(\ell_n(a); \varepsilon)) \\ = \{x \in X : |\ell_j(x) - \ell_j(a)| < \varepsilon, 1 \leq j \leq n\}. \end{aligned}$$

Finally we arrive at the following characterisation of the weak topology on the normed space X . A non-empty subset G of X is open with respect to the weak topology if and only if for any $a \in G$ there is $n \in \mathbb{N}$, $\ell_1, \dots, \ell_n \in X^*$ and $\varepsilon > 0$ such that

$$\{x \in X : |\ell_j(x) - \ell_j(a)| < \varepsilon, 1 \leq j \leq n\} \subseteq G.$$

Equivalently, we can say that a set is open with respect to the weak topology if and only if it is a union of sets of the above form. We introduce the following notation; for given $\ell_1, \dots, \ell_n \in X^*$ and $\varepsilon > 0$, we write

$$N(a; \ell_1, \dots, \ell_n, \varepsilon) = \{x \in X : |\ell_j(x) - \ell_j(a)| < \varepsilon, 1 \leq j \leq n\}.$$

Then each $N(a; \ell_1, \dots, \ell_n, \varepsilon)$ contains a and is $\sigma(X, X^*)$ open. A non-empty subset G in X is weakly open if and only if for any $a \in G$ we have $N(a; \ell_1, \dots, \ell_n, \varepsilon) \subseteq G$ for some $n \in \mathbb{N}$, $\ell_1, \dots, \ell_n \in X^*$ and $\varepsilon > 0$. The sets $N(a; \ell_1, \dots, \ell_n, \varepsilon)$ play a rôle analogous to that of the open balls in a metric space.

Definition 9.1. Let x be a point in a topological space (X, \mathcal{T}) . A collection \mathcal{N}_x of neighbourhoods of x is said to be a neighbourhood base at x if and only if for each open set G with $x \in G$ there is some $N \in \mathcal{N}_x$ such that $x \in N$ and $N \subseteq G$. (Note that the members of \mathcal{N}_x need not themselves be open sets.)

Thus, we see that the family

$$\mathcal{N}_a = \{N(a; \ell_1, \dots, \ell_n, \varepsilon) : n \in \mathbb{N}, \varepsilon > 0, \ell_1, \dots, \ell_n \in X^*\}$$

is an *open* neighbourhood base at a for the weak topology.

In order to discuss the relationship between the norm and the weak topologies on X , we shall need the following result.

Proposition 9.2. *Let ℓ_1, \dots, ℓ_n be bounded linear functionals on an infinite dimensional normed space X . Then there exists $x \in X$ such that $x \neq 0$ and $\ell_1(x) = \dots = \ell_n(x) = 0$.*

Proof. First we note that X^* is infinite dimensional. In fact, if X^* were finite dimensional, then its dual, X^{**} , would also be finite dimensional. But we have seen that X is isometrically isomorphic to a subspace of X^{**} . Since X is infinite dimensional the same must be true of X^{**} and hence also of X^* .

Let $\lambda_1, \dots, \lambda_m$ be linearly independent elements spanning the finite dimensional subspace generated by ℓ_1, \dots, ℓ_n . Then $m \leq n$. Since X^* is infinite dimensional, there is $\ell \in X^*$ independent of $\lambda_1, \dots, \lambda_m$. Define $T : X \rightarrow \mathbb{C}^{m+1}$ by

$$Tx = (\ell(x), \lambda_1(x), \dots, \lambda_m(x)).$$

It is clear that T is a linear operator and so its range $\text{ran } T$ is a linear subspace of \mathbb{C}^{m+1} . If $\text{ran } T \neq \mathbb{C}^{m+1}$, there would exist a vector $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_m)$ in \mathbb{C}^{m+1} orthogonal to the range of T , that is,

$$\alpha_0 \ell(x) + \alpha_1 \lambda_1(x) + \dots + \alpha_m \lambda_m(x) = 0$$

for all $x \in X$. But this is precisely the statement that

$$\alpha_0 \ell + \alpha_1 \lambda_1 + \dots + \alpha_m \lambda_m = 0$$

in X^* . This contradicts the assumed linear independence of $\ell, \lambda_1, \dots, \lambda_m$. It follows that $\text{ran } T = \mathbb{C}^{m+1}$. In particular, there is $x \in X$ such that $Tx = (1, 0, 0, \dots, 0)$. That is, there is $x \in X$ such that $\ell(x) = 1$ and $\lambda_j(x) = 0$ for all $1 \leq j \leq m$. Hence $x \neq 0$ and also $\ell_i(x) = 0$ for $1 \leq i \leq n$ and the proof is complete. ■

Theorem 9.3. *The weak topology on a normed space is weaker than the norm topology. If X is infinite dimensional, then the weak topology on X is strictly weaker than the norm topology on X .*

Proof. Every element of X^* is continuous when X is equipped with the norm topology. The weak topology is the weakest topology on X with this property, so it is immediately clear that the weak topology is weaker than the norm topology.

Now suppose that X is infinite-dimensional. We shall exhibit a set which is open with respect to the norm topology but not with respect to the weak topology. We consider the “open” unit ball

$$G = \{x \in X : \|x\| < 1\}.$$

Clearly G is open with respect to the norm topology on X . We claim that G is not weakly open. If, on the contrary, G were weakly open, then, since $0 \in G$, there would be $n \in \mathbb{N}$, $\ell_1, \dots, \ell_n \in X^*$ and $\varepsilon > 0$ such that

$$N(0; \ell_1, \dots, \ell_n, \varepsilon) \subseteq G.$$

By the previous proposition, there is $x \in X$ with $x \neq 0$ such that $\ell_i(x) = 0$ for all $1 \leq i \leq n$. Put $y = 2x/\|x\|$. Then $\|y\| = 2$ and $\ell_i(y) = 0$ for all $1 \leq i \leq n$. It follows that $y \in N(0; \ell_1, \dots, \ell_n, \varepsilon)$ but $y \notin G$. Hence $N(0; \ell_1, \dots, \ell_n, \varepsilon)$ cannot be contained in G , which is a contradiction. We conclude that G is not weakly open and therefore the weak topology on X is strictly weaker than the norm topology. ■

Remark 9.4. If we put $y = kx/\|x\|$ in the argument above, we see that $y \in N(0; \ell_1, \dots, \ell_n, \varepsilon)$ and $\|y\| = k$. It follows that no weak neighbourhood $N(0; \ell_1, \dots, \ell_n, \varepsilon)$ can be norm bounded. Moreover, for any $a \in X$, we have $N(a; \ell_1, \dots, \ell_n, \varepsilon) = a + N(0; \ell_1, \dots, \ell_n, \varepsilon)$ and so *none* of the non-empty weakly open sets can be norm bounded.

We now turn to a discussion of a topology on X^* , the dual of the normed space X . The idea is to consider X as a family of maps $\ell : X^* \rightarrow \mathbb{C}$ given by $x : \ell \mapsto \ell(x)$, for $x \in X$ and $\ell \in X^*$ — this is the map ψ_x defined earlier.

Definition 9.5. The weak*-topology on X^* , the dual of the normed space X , is the $\sigma(X^*, X)$ -topology, where X is considered as a collection of maps from $X^* \rightarrow \mathbb{C}$ as above.

Remark 9.6. The weak*-topology is also called the w^* -topology on X^* . Since X separates points of X^* , the w^* -topology is Hausdorff. In view of the identification of X as a subset of X^* , we see that the w^* -topology is weaker than the $\sigma(X^*, X^{**})$ -topology on X^* . Of course, we have equality if X is reflexive. The converse is also true.

Theorem 9.7. *The normed space X is reflexive if and only if the weak and weak* topologies on X^* coincide.*

Proof. We will not do this here. ■

By repeating our earlier analysis, we see that a non-empty subset $G \subseteq X^*$ is w^* -open if and only if for each $\ell \in G$ there is $m \in \mathbb{N}$, elements x_1, \dots, x_m in X and $\varepsilon > 0$ such that

$$N(\ell; x_1, \dots, x_m, \varepsilon) \cong \{\lambda \in X^* : |\lambda(x_i) - \ell(x_i)| < \varepsilon, \quad 1 \leq i \leq m\} \subseteq G.$$

An open neighbourhood base at 0 for the w^* -topology on X^* is given by

$$\mathcal{N}_0 = \{N(0; x_1, \dots, x_m, \varepsilon) : m \in \mathbb{N}, x_1, \dots, x_m \in X, \varepsilon > 0\}.$$

An open neighbourhood base at $\ell \in X^*$ is given by

$$\mathcal{N}_\ell = \{N + \ell : N \in \mathcal{N}_0\}.$$

One can show the following.

Theorem 9.8. *Let X be a normed space and Y a family of bounded linear functionals on X which separates points of X . Then the bounded linear functional ℓ on X is $\sigma(X, Y)$ -continuous if and only if $\ell \in Y$. In particular, the only w^* -continuous linear functionals on X^* are the elements of X .*

A subset K of a metric space is compact if and only if any sequence in K has a subsequence which converges to an element of K , but this need no longer be true in a topological space. We will see an example of this later. We have seen that in a general topological space nets can be used rather than sequences, so the natural question is whether there is a sensible notion of that of “subnet” of a net, generalising that of subsequence of a sequence. Now, a subsequence of a sequence is obtained simply by leaving out various terms—the sequence is labelled by the natural numbers and the subsequence is labelled by a subset of these. The notion of a subnet is somewhat more subtle than this.

Definition 9.9. A map $F : J \rightarrow I$ between directed sets I and J is said to be cofinal if for any $\alpha \in I$ there is some $\beta' \in J$ such that $J(\beta) \succeq \alpha$ whenever $\beta \succeq \beta'$. In other words, F is eventually greater than any given $\alpha \in I$. Suppose that $(x_\alpha)_{\alpha \in I}$ is a net indexed by I and that $F : J \rightarrow I$ is a cofinal map from the directed set J into I . The net $(y_\beta)_{\beta \in J} = (x_{F(\beta)})_{\beta \in J}$ is said to be a subnet of the net $(x_\alpha)_{\alpha \in I}$.

It is important to notice that there is *no* requirement that the index set for the subnet be the same as that of the original net.

Example 9.10. If we set $I = J = \mathbb{N}$, equipped with the usual ordering, and let $F : J \rightarrow I$ be any increasing map, then the subnet $(y_n) = (x_{F(n)})$ is a subsequence of the sequence (x_n) .

Example 9.11. Let $I = \mathbb{N}$ with the usual order, and let $J = \mathbb{N}$ equipped with the usual ordering on the even and odd elements separately but where any even number is declared to be greater than any odd number. Thus I and J are directed sets. Define $F : J \rightarrow I$ by $F(\beta) = 3\beta$. Let $\alpha \in I$ be given. Set $\beta' = 2\alpha$ so that if $\beta \succeq \beta'$ in J , we must have that β is even and greater than β' in the usual sense. Hence $F(\beta) = 3\beta \geq \beta \geq \beta' = 2\alpha \geq \alpha$ in I and so F is cofinal. Let $(x_n)_{n \in I}$ be any sequence of real numbers, say. Then $(x_{F(m)})_{m \in J} = (x_{3m})_{m \in J}$ is a subnet of $(x_n)_{n \in I}$. It is not a subsequence because the ordering of the index set is not the usual one. Suppose that $x_{2k} = 0$ and $x_{2k-1} = 2k - 1$ for $k \in I = \mathbb{N}$. Then (x_n) is the sequence $(1, 0, 3, 0, 5, 0, 7, 0, \dots)$. The subsequence $(x_{3m})_{m \in \mathbb{N}}$ is $(3, 0, 9, 0, 15, 0, \dots)$ which clearly does not converge in \mathbb{R} . However, the subnet $(x_{3m})_{m \in J}$ does converge, to 0. Indeed, for $m \succeq 2$ in J , we have $x_{3m} = 0$.

Proposition 9.12. Let $(x_\alpha)_I$ be a net in the space X and let \mathcal{A} be a family of subsets of X such that

- (i) $(x_\alpha)_I$ is frequently in each member of \mathcal{A} ;
- (ii) for any $A, B \in \mathcal{A}$ there is $C \in \mathcal{A}$ such that $C \subseteq A \cap B$.

Then there is a subnet $(x_{F(\beta)})_J$ of the net $(x_\alpha)_I$ such that $(x_{F(\beta)})_J$ is eventually in each member of \mathcal{A} .

Proof. Equip \mathcal{A} with the ordering given by reverse inclusion, that is, we define $A \preceq B$ to mean $B \subseteq A$ for $A, B \in \mathcal{A}$. For any $A, B \in \mathcal{A}$, there is $C \in \mathcal{A}$ with $C \subseteq A \cap B$, by (ii). This means that $C \succeq A$ and $C \succeq B$ and we see that \mathcal{A} is directed with respect to this partial ordering.

Let \mathcal{E} denote the collection of pairs $(\alpha, A) \in I \times \mathcal{A}$ such that $x_\alpha \in A$;

$$\mathcal{E} = \{(\alpha, A) : \alpha \in I, A \in \mathcal{A}, x_\alpha \in A\}.$$

Define $(\alpha', A') \preceq (\alpha'', A'')$ to mean that $\alpha' \preceq \alpha''$ in I and $A' \preceq A''$ in \mathcal{A} . Then \preceq is a partial order on \mathcal{E} . Furthermore, for given $(\alpha', A'), (\alpha'', A'')$ in \mathcal{E} , there is $\alpha \in I$ with $\alpha \succeq \alpha'$ and $\alpha \succeq \alpha''$, and there is $A \in \mathcal{A}$ such that $A \succeq A'$ and $A \succeq A''$. But (x_α) is frequently in A , by (i), and therefore there is $\beta \succeq \alpha \in I$ such that $x_\beta \in A$. Thus $(\beta, A) \in \mathcal{E}$ and $(\beta, A) \succeq (\alpha, A')$, $(\beta, A) \succeq (\alpha, A'')$ and it follows that \mathcal{E} is directed. \mathcal{E} will be the index set for the subnet.

Next, we must construct a cofinal map from \mathcal{E} to I . Define $F : \mathcal{E} \rightarrow I$ by $F((\alpha, A)) = \alpha$. To show that F is cofinal, let $\alpha_0 \in I$ be given. For any $A \in \mathcal{A}$ there is $\alpha \succeq \alpha_0$ such that $x_\alpha \in A$ (since (x_α) is frequently in each $A \in \mathcal{A}$). Hence $(\alpha, A) \in \mathcal{E}$ and $F((\alpha, A)) = \alpha \succeq \alpha_0$. So if $(\alpha', A') \succeq (\alpha, A)$ in \mathcal{E} , then we have

$$F((\alpha', A')) = \alpha' \succeq \alpha \succeq \alpha_0.$$

This shows that F is cofinal and therefore $(x_{F((\alpha, A))})_{\mathcal{E}}$ is a subnet of $(x_\alpha)_I$.

It remains to show that this subnet is eventually in every member of \mathcal{A} . Let $A \in \mathcal{A}$ be given. Then there is $\alpha \in I$ such that $x_\alpha \in A$ and so $(\alpha, A) \in \mathcal{E}$. For any $(\alpha', A') \in \mathcal{E}$ with $(\alpha', A') \succeq (\alpha, A)$, we have

$$x_{F((\alpha', A'))} = x_{\alpha'} \in A' \subseteq A.$$

Thus $(x_{F((\alpha, A))})_{\mathcal{E}}$ is eventually in A . ■

Theorem 9.13. *A point x in a topological space X is a cluster point of the net $(x_\alpha)_I$ if and only if some subnet converges to x .*

Proof. Suppose that x is a cluster point of the net $(x_\alpha)_I$ and let \mathcal{N} be the family of neighbourhoods of x . Then if $A, B \in \mathcal{N}$, we have $A \cap B \in \mathcal{N}$, and also (x_α) is frequently in each member of \mathcal{N} . By the preceding proposition, there is a subnet $(y_\beta)_J$ eventually in each member of \mathcal{N} , that is, the subnet (y_β) converges to x .

Conversely, suppose that $(y_\beta)_{\beta \in J} = (x_{F(\beta)})_{\beta \in J}$ is a subnet of $(x_\alpha)_I$ converging to x . We must show that x is a cluster point of $(x_\alpha)_I$. Let N be any neighbourhood of x . Then there is $\beta_0 \in J$ such that $x_{F(\beta)} \in N$ whenever $\beta \succeq \beta_0$. Since F is cofinal, for any given $\alpha' \in I$ there is $\beta' \in J$ such that $F(\beta) \succeq \alpha'$ whenever $\beta \succeq \beta'$. Let $\beta \succeq \beta_0$ and $\beta \succeq \beta'$. Then $F(\beta) \succeq \alpha'$ and $y_\beta = x_{F(\beta)} \in N$. Hence $(x_\alpha)_I$ is frequently in N and we conclude that x is a cluster point of the net $(x_\alpha)_I$, as claimed. ■

In a metric space, compactness is equivalent to sequential compactness (— the Bolzano-Weierstrass property). In a general topological space, this need no longer be the case. However, there is an analogue in terms of nets.

Theorem 9.14. *A topological space (X, \mathcal{T}) is compact if and only if every net in X has a convergent subnet.*

Proof. Suppose that every net has a convergent subnet. Let $\{G_\alpha\}_I$ be an open cover of X with no finite subcover. Let \mathcal{F} be the collection of finite subfamilies of the open cover, ordered by set-theoretic inclusion. For each $F = \{G_{\alpha_1}, \dots, G_{\alpha_m}\} \in \mathcal{F}$, let x_F be any point in X such that $x_F \notin \bigcup_{j=1}^m G_{\alpha_j}$. Note that such x_F exists since $\{G_\alpha\}$ has no finite subcover. By hypothesis, the net $(x_F)_{F \in \mathcal{F}}$ has a convergent subnet or, equivalently, by the previous theorem, a cluster point x , say. Now, since $\{G_\alpha\}$ is a cover of X , there is some α' such that $x \in G_{\alpha'}$. But then, by definition of cluster point, $(x_F)_{\mathcal{F}}$ is frequently in $G_{\alpha'}$. Thus, for any $F' \in \mathcal{F}$, there is $F \succeq F' \in \mathcal{F}$ such that $x_F \in G_{\alpha'}$. In particular, if we take $F' = \{G_{\alpha'}\}$, we deduce that there is $F = \{G_{\alpha_1}, \dots, G_{\alpha_k}\}$ such that $F \succeq \{G_{\alpha'}\}$, that is, $\{G_{\alpha'}\} \subseteq F$, and such that $x_F \in G_{\alpha'}$. Hence $G_{\alpha'} = G_{\alpha_i}$ for some $1 \leq i \leq k$, and

$$x_F \in G_{\alpha_i} \subseteq \bigcup_{j=1}^k G_{\alpha_j}.$$

But $x_F \notin \bigcup_{j=1}^k G_{\alpha_j}$, by construction. This contradiction implies that every open cover has a finite subcover, and so (X, \mathcal{T}) is compact.

For the converse, suppose that (X, \mathcal{T}) is compact and let $(x_\alpha)_I$ be a net in X . Suppose that $(x_\alpha)_I$ has no cluster points. Then, for any $x \in X$, there is an open neighbourhood U_x of x and $\alpha_x \in I$ such that $x_\alpha \notin U_x$ whenever $\alpha \succeq \alpha_x$. The family $\{U_x : x \in X\}$ is an open cover of X and so there exists $x_1, \dots, x_n \in X$ such that $\bigcup_{i=1}^n U_{x_i} = X$. Since I is directed there is $\alpha \succeq \alpha_i$ for each $i = 1, \dots, n$. But then $x_\alpha \notin U_{x_i}$ for all $i = 1, \dots, n$, which is impossible since the U_{x_i} 's cover X . We conclude that $(x_\alpha)_I$ has a cluster point, or, equivalently, a convergent subnet. ■

Definition 9.15. A universal net in a topological space (X, \mathcal{T}) is a net with the property that, for any subset A of X , it is either eventually in A or eventually in $X \setminus A$, the complement of A .

The concept of a universal net leads to substantial simplification of the proofs of various results, as we will see.

Proposition 9.16. *If a universal net has a cluster point, then it converges (to the cluster point). In particular, a universal net in a Hausdorff space can have at most one cluster point.*

Proof. Suppose that x is a cluster point of the universal net $(x_\alpha)_I$. Then for each neighbourhood N of x , (x_α) is frequently in N . However, (x_α) is either eventually in N or eventually in $X \setminus N$. Evidently, the former must be the case and we conclude that (x_α) converges to x . The last part follows because in a Hausdorff space a net can converge to at most one point. ■

At this point, it is not at all clear that universal nets exist!

Examples 9.17.

1. It is clear that any eventually constant net is a universal net. In particular, any net with finite index set is a universal net. Indeed, if $(x_\alpha)_I$ is a net in X with finite index set I , then I has a maximum element, α' , say. The net is therefore eventually equal to $x_{\alpha'}$. For any subset $A \subseteq X$, we have that $(x_\alpha)_I$ is eventually in A or eventually in $X \setminus A$ depending on whether $x_{\alpha'}$ belongs to A or not.
2. No sequence can be a universal net, unless it is eventually constant. To see this, suppose that $(x_n)_{n \in \mathbb{N}}$ is a sequence which is not eventually constant. Then the set $S = \{x_n : n \in \mathbb{N}\}$ is an infinite set. Let A be any infinite subset of S such that $S \setminus A$ also infinite. Then (x_n) cannot be eventually in either of A or its complement. That is, the sequence $(x_n)_{\mathbb{N}}$ cannot be universal.

We shall show that every net has a universal subnet. First we need the following lemma.

Lemma 9.18. *Let $(x_\alpha)_I$ be a net in a topological space X . Then there is a family \mathcal{C} of subsets of X such that*

- (i) (x_α) is frequently in each member of \mathcal{C} ;
- (ii) if $A, B \in \mathcal{C}$ then $A \cap B \in \mathcal{C}$;
- (iii) for any $A \subseteq X$, either $A \in \mathcal{C}$ or $X \setminus A \in \mathcal{C}$.

Proof. Let Φ denote the collection of families of subsets of X satisfying the conditions (i) and (ii):

$$\Phi = \{\mathcal{F} : \mathcal{F} \text{ satisfies (i) and (ii)}\}.$$

Evidently $\{X\} \in \Phi$ so $\Phi \neq \emptyset$. The collection Φ is partially ordered by set inclusion:

$$\mathcal{F}_1 \preceq \mathcal{F}_2 \text{ if and only if } \mathcal{F}_1 \subseteq \mathcal{F}_2, \text{ for } \mathcal{F}_1, \mathcal{F}_2 \in \Phi.$$

Let $\{\mathcal{F}_\gamma\}$ be a totally ordered family in Φ , and put $\widehat{\mathcal{F}} = \bigcup_\gamma \mathcal{F}_\gamma$. We shall show that $\widehat{\mathcal{F}} \in \Phi$. Indeed, if $A \in \widehat{\mathcal{F}}$, then there is some γ such that $A \in \mathcal{F}_\gamma$, and so (x_α) is frequently in A and condition (i) holds.

Now, for any $A, B \in \widehat{\mathcal{F}}$, there is γ_1 and γ_2 such that $A \in \mathcal{F}_{\gamma_1}$, and $B \in \mathcal{F}_{\gamma_2}$. Suppose, without loss of generality, that $\mathcal{F}_{\gamma_1} \preceq \mathcal{F}_{\gamma_2}$. Then $A, B \in \mathcal{F}_{\gamma_2}$ and therefore $A \cap B \in \mathcal{F}_{\gamma_2} \subseteq \widehat{\mathcal{F}}$, and we see that condition (ii) is satisfied. Thus $\widehat{\mathcal{F}} \in \Phi$ as claimed.

By Zorn's lemma, we conclude that Φ has a maximal element, \mathcal{C} , say. We shall show that \mathcal{C} also satisfies condition (iii).

To see this, let $A \subseteq X$ be given. Suppose, first, that it is true that (x_α) is frequently in $A \cap B$ for all $B \in \mathcal{C}$. Define \mathcal{F}' by

$$\mathcal{F}' = \{C \subseteq X : A \cap B \subseteq C, \text{ for some } B \in \mathcal{C}\}.$$

Then $C \in \mathcal{F}'$ implies that $A \cap B \subseteq C$ for some B in \mathcal{C} and so (x_α) is frequently in C . Also, if $C_1, C_2 \in \mathcal{F}'$, then there is B_1 and B_2 in \mathcal{C} such that $A \cap B_1 \subseteq C_1$ and $A \cap B_2 \subseteq C_2$. It follows that $A \cap (B_1 \cap B_2) \subseteq C_1 \cap C_2$. Since $B_1 \cap B_2 \in \mathcal{C}$, we deduce that $C_1 \cap C_2 \in \mathcal{F}'$. Thus $\mathcal{F}' \in \Phi$.

However, it is clear that $A \in \mathcal{F}'$ and also that if $B \in \mathcal{C}$ then $B \in \mathcal{F}'$. But \mathcal{C} is maximal in Φ , and so $\mathcal{F}' = \mathcal{C}$ and we conclude that $A \in \mathcal{C}$, and (iii) holds.

Now suppose that it is false that (x_α) is frequently in every $A \cap B$, for $B \in \mathcal{C}$. Then there is some $B_0 \in \mathcal{C}$ such that (x_α) is not frequently in $A \cap B_0$. Thus there is α_0 such that $x_\alpha \in X \setminus (A \cap B_0)$ for all $\alpha \succeq \alpha_0$. That is, (x_α) is eventually in $X \setminus (A \cap B_0) \equiv \widetilde{A}$, say. It follows that (x_α) is frequently in $\widetilde{A} \cap B$ for every $B \in \mathcal{C}$. Thus, as above, we deduce that $\widetilde{A} \in \mathcal{C}$. Furthermore, for any $B \in \mathcal{C}$, $B \cap B_0 \in \mathcal{C}$ and so $\widetilde{A} \cap B \cap B_0 \in \mathcal{C}$. But

$$\begin{aligned} \widetilde{A} \cap B \cap B_0 &= (X \setminus (A \cap B_0)) \cap (B \cap B_0) \\ &= ((X \setminus A) \cup (X \setminus B_0)) \cap B \cap B_0 \\ &= \{(X \setminus A) \cap B \cap B_0\} \cup \underbrace{\{(X \setminus B_0) \cap B \cap B_0\}}_{=\emptyset} \\ &= (X \setminus A) \cap B \cap B_0 \end{aligned}$$

and so we see that (x_α) is frequently in $(X \setminus A) \cap B \cap B_0$ and hence is frequently in $(X \setminus A) \cap B$ for any $B \in \mathcal{C}$. Again, by the above argument, we deduce that $X \setminus A \in \mathcal{C}$. This proves the claim and completes the proof of the lemma. \blacksquare

Theorem 9.19. *Every net has a universal subnet.*

Proof. To prove the theorem, let $(x_\alpha)_I$ be any net in X , and let \mathcal{C} be a family of subsets as given by the lemma. Then, in particular, the conditions of Proposition 9.12 hold, and we deduce that $(x_\alpha)_I$ has a subnet $(y_\beta)_J$ such that $(y_\beta)_J$ is eventually in each member of \mathcal{C} . But, for any $A \subseteq X$, either $A \in \mathcal{C}$ or $X \setminus A \in \mathcal{C}$, hence the subnet $(y_\beta)_J$ is either eventually in A or eventually in $X \setminus A$; that is, $(y_\beta)_J$ is universal. ■

Theorem 9.20. *A topological space is compact if and only if every universal net converges.*

Proof. Suppose that (X, \mathcal{T}) is a compact topological space and that (x_α) is a universal net in X . Since X is compact, (x_α) has a convergent subnet, with limit $x \in X$, say. But then x is a cluster point of the universal net (x_α) and therefore the net (x_α) itself converges to x .

Conversely, suppose that every universal net in X converges. Let (x_α) be any net in X . Then (x_α) has a subnet which is universal and must therefore converge. In other words, we have argued that (x_α) has a convergent subnet and therefore X is compact. ■

Corollary 9.21. *A non-empty subset K of a topological space is compact if and only if every universal net in K converges in K .*

Proof. The subset K of the topological space (X, \mathcal{T}) is compact if and only if it is compact with respect to the induced topology \mathcal{T}_K on K . The result now follows by applying the theorem to (K, \mathcal{T}_K) . ■

10. Product Spaces

Suppose that (X_1, \mathcal{T}_1) and (X_2, \mathcal{T}_2) are topological spaces and consider the cartesian product

$$Y = X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\}.$$

We would like to give Y a topology using those on X_1 and X_2 . Define \mathcal{T} to be the collection of subsets of Y given as follows: $\emptyset \in \mathcal{T}$ and the non-empty set G belongs to \mathcal{T} if and only if for any point $(a_1, a_2) \in G$ there are open sets $U \in \mathcal{T}_1$ and $V \in \mathcal{T}_2$ such that $(a_1, a_2) \in U \times V$. In other words, G belongs to \mathcal{T} if and only if each of its points is contained in an “open rectangle” $U \times V$ which itself lies in G . One checks that \mathcal{T} is indeed a topology — called the product topology on the cartesian product $X_1 \times X_2$ given by \mathcal{T}_1 and \mathcal{T}_2 . A set is open (with respect to \mathcal{T}) if and only if it is a union of sets of the form $U \times V$; that is, the sets of the form $U \times V$ form a base for the product topology. Since $U \times V = U \times X_2 \cap X_1 \times V$, we see that the sets $\{U \times X_2, X_1 \times V : U \in \mathcal{T}_1, V \in \mathcal{T}_2\}$ constitute a sub-base for the topology \mathcal{T} .

Example 10.1. Let $(X_1, \mathcal{T}_1) = (X_2, \mathcal{T}_2) = (\mathbb{R}, \mathcal{T}_{\text{Euclidean}})$. Then the cartesian product $X_1 \times X_2$ is just \mathbb{R}^2 , and we see that a non-empty set $G \subseteq \mathbb{R}^2$ is open with respect to the product topology if and only if each of its points is contained in an open rectangle also lying in G . Thus, the product topology is precisely the usual Euclidean topology on \mathbb{R}^2 . Rectangles are as good as discs.

The projection maps, p_1 and p_2 , on the cartesian product $X_1 \times X_2$, are defined by

$$\begin{aligned} p_1 : X_1 \times X_2 &\rightarrow X_1, & (x_1, x_2) &\mapsto x_1 \\ p_2 : X_1 \times X_2 &\rightarrow X_2, & (x_1, x_2) &\mapsto x_2. \end{aligned}$$

For any open set $U \subseteq X_1$ (that is, $U \in \mathcal{T}_1$), we have $p_1^{-1}(U) = U \times X_2 \in \mathcal{T}$ and so it follows that $p_1 : X_1 \times X_2 \rightarrow X_1$ is continuous. Similarly, $p_2 : X_1 \times X_2 \rightarrow X_2$ is continuous. This property characterises the product topology as we now show.

Proposition 10.2. *The product topology \mathcal{T} is the weakest topology on the cartesian product $X_1 \times X_2$ such that both p_1 and p_2 are continuous.*

Proof. Suppose that \mathcal{S} is a topology on $Y = X_1 \times X_2$ with respect to which both p_1 and p_2 are continuous. Then for any $U \in \mathcal{T}_1$, $p_1^{-1}(U) \in \mathcal{S}$. But $p_1^{-1}(U) = U \times X_2$ and so $U \times X_2 \in \mathcal{S}$ for all $U \in \mathcal{T}_1$. Similarly, $X_1 \times V \in \mathcal{S}$ for all $V \in \mathcal{T}_2$. Since these sets form a sub-base for \mathcal{T} we deduce that $\mathcal{T} \subseteq \mathcal{S}$, as required. ■

We would like to generalise this to an arbitrary cartesian product of topological spaces. Let $\{(X_\alpha, \mathcal{T}_\alpha) : \alpha \in I\}$ be a collection of topological spaces indexed by the set I . We recall that $X = \prod_\alpha X_\alpha$, the cartesian product of the X_α 's, is defined to be the collection of maps γ from I into the union $\bigcup_\alpha X_\alpha$ satisfying $\gamma(\alpha) \in X_\alpha$ for each $\alpha \in I$. We can think of the value $\gamma(\alpha)$ as the α^{th} coordinate of the point γ in X . The idea is to construct a topology on $X = \prod_\alpha X_\alpha$ built from the individual topologies \mathcal{T}_α . Two possibilities suggest themselves. The first is to construct a topology on X such that it is the weakest topology with respect to which all the projection maps $p_\alpha \rightarrow X_\alpha$ are continuous. The second is to construct the topology on X whose open sets are unions of “super rectangles”, that is, sets of the form $\prod_\alpha U_\alpha$, where $U_\alpha \in \mathcal{T}_\alpha$ for every $\alpha \in I$.

In general, these two topologies are *not* the same, as we will see. Consider the first construction. We wish to define a topology \mathcal{T} on X making every projection map p_α continuous. This means that \mathcal{T} must contain all sets of the form $p_\alpha^{-1}(U_\alpha)$, for $U_\alpha \in \mathcal{T}_\alpha$, and also finite intersections of such sets, and also arbitrary unions of such finite intersections. So we define \mathcal{T} to be the topology on X with sub-base given by the sets $p_\alpha^{-1}(U_\alpha)$, for $U_\alpha \in \mathcal{T}_\alpha$. For reasons we will discuss later, *this* topology turns out to be the more appropriate and is taken as the definition of the product topology.

Definition 10.3. The product topology on the cartesian product of the topological spaces $\{(X_\alpha, \mathcal{T}_\alpha) : \alpha \in I\}$ is that with sub-base given by the sets $p_\alpha^{-1}(U_\alpha)$, for $U_\alpha \in \mathcal{T}_\alpha$.

Clearly, this agrees with the case discussed earlier for the product of just two spaces. Moreover, this definition is precisely the statement that \mathcal{T} is the $\sigma(\prod_\alpha X_\alpha, \mathcal{F})$ -topology, where \mathcal{F} is the family of projection maps

$\{p_\alpha : \alpha \in I\}$. In other words, \mathcal{T} is the weakest topology on $\prod_\alpha X_\alpha$ with respect to which every projection map p_α is continuous.

Remark 10.4. Let G be a non-empty open set in X , equipped with the product topology, and let $\gamma \in G$. Then, by definition of the topology, there exist $\alpha_1, \dots, \alpha_n \in I$ and open sets U_{α_i} in X_{α_i} , $1 \leq i \leq n$, such that

$$\gamma \in p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \dots \cap p_{\alpha_n}^{-1}(U_{\alpha_n}) \subseteq G.$$

Hence there are open sets S_α , $\alpha \in I$, such that $\gamma \in \prod_\alpha S_\alpha \subseteq G$ and where all but a finite number of the S_α are equal to the whole space X_α . This means that G can differ from X in at most a finite number of components.

Now let us consider the second candidate for a topology on X . Let \mathcal{S} be the topology on X with base given by the sets of the form $\prod_\alpha V_\alpha$, where $V_\alpha \in \mathcal{T}_\alpha$ for $\alpha \in I$. Thus, a non-empty set G in X belongs to \mathcal{S} if and only if for any point x in G there exist $V_\alpha \in \mathcal{T}_\alpha$ such that

$$x \in \prod_\alpha V_\alpha \subseteq G.$$

Here there is *no* requirement that all but a finite number of the V_α are equal to X_α .

Definition 10.5. The topology on the cartesian product $\prod_\alpha X_\alpha$ constructed in this way is called the box-topology on X .

Evidently, in general, \mathcal{S} is strictly finer than the product topology \mathcal{T} .

We shall use the notation $\mathcal{T}_{\text{product}}$ and \mathcal{T}_{box} for the product and box topologies, respectively.

Proposition 10.6. A net (x_λ) converges in $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{product}})$ if and only if $(p_\alpha(x_\lambda))$ converges in $(X_\alpha, \mathcal{T}_\alpha)$ for each $\alpha \in I$.

Proof. Suppose that $x_\lambda \rightarrow x$ in $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{product}})$. Then $p_\alpha(x_\lambda) \rightarrow p_\alpha(x)$ for each α , since p_α is continuous.

Conversely, suppose that $p_\alpha(x_\lambda) \rightarrow z_\alpha$ in $(X_\alpha, \mathcal{T}_\alpha)$ for each $\alpha \in I$. We shall show that $x_\lambda \rightarrow z$ in $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{product}})$ where z is given by $z(\alpha) = z_\alpha$. Indeed, let G be any neighbourhood of z in $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{product}})$. Then there are $\alpha_1, \dots, \alpha_n$ and open sets $U_{\alpha_1}, \dots, U_{\alpha_n}$ such that

$$z \in p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \dots \cap p_{\alpha_n}^{-1}(U_{\alpha_n}) \subseteq G.$$

Now, $p_\alpha(x_\lambda) \rightarrow z_\alpha$, for each α , so there is λ_j such that $p_{\alpha_j}(x_\lambda) \in U_{\alpha_j}$ whenever $\lambda \succeq \lambda_j$, $1 \leq j \leq n$. Let $\lambda' \succeq \lambda_j$ for all $1 \leq j \leq n$. Then if $\lambda \succeq \lambda'$,

we have $p_{\alpha_j}(x_\lambda) \in U_{\alpha_j}$ and so $x_\lambda \in p_{\alpha_j}^{-1}(U_{\alpha_j})$ for all $1 \leq j \leq n$. In other words, for $\lambda \succeq \lambda'$,

$$x_\lambda \in p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \cdots \cap p_{\alpha_n}^{-1}(U_{\alpha_n}) \subseteq G.$$

Hence $x_\lambda \rightarrow z$ in $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{product}})$ as required. \blacksquare

Example 10.7. Let $I = \mathbb{N}$, let X_α be the open interval $(-2, 2)$ for each $\alpha \in \mathbb{N}$, and let \mathcal{T}_α be the usual (Euclidean) topology on X_α . Let $x_n \in \prod_k X_k$ be the element $x_n = (\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \dots)$; that is, $p_k(x_n) = \frac{1}{n}$ for all $k \in I = \mathbb{N}$. Clearly $p_k(x_n) \rightarrow 0$ as $n \rightarrow \infty$, for each k and so the sequence (x_n) converges to z in $(\prod_k X_k, \mathcal{T}_{\text{product}})$ where z is given by $p_k(z) = 0$ for all k .

However, (x_n) does *not* converge to z with respect to the box-topology. Indeed, to see this, let $G = \prod_k A_k$ where A_k is the open set $A_k = (-\frac{1}{k}, \frac{1}{k}) \in \mathcal{T}_k$. Then G is open with respect to the box-topology and is a neighbourhood of z but $x_n \notin G$ for any $n \in \mathbb{N}$. It follows that, in fact, (x_n) does not converge at all with respect to the box-topology (—if it did, then the limit would have to be the same as that for the product topology, namely z).

Remark 10.8. This is a first indication that the box-topology may not be very useful (apart from being a possible source of counter-examples).

Suppose that each $(X_\alpha, \mathcal{T}_\alpha)$, $\alpha \in I$, is compact. What can be said about the product space $\prod_\alpha X_\alpha$ with respect to the product and the box topologies?

Example 10.9. Let $I = \mathbb{N}$ and let $X_k = [0, 1]$ for each $k \in I = \mathbb{N}$ and equip each X_k with the Euclidean topology. Then each (X_k, \mathcal{T}_k) is compact. However, the product space $\prod_k X_k$ is *not* compact with respect to the box-topology. To see this, let $I_k(t)$ be the open disk in X_k with centre t and radius $\frac{1}{k}$;

$$I_k(t) = [0, 1] \cap \left(t - \frac{1}{k}, t + \frac{1}{k}\right) \subseteq [0, 1].$$

Evidently, the diameter of $I_k(t)$ is at most $\frac{2}{k}$. For each $x \in \prod_k X_k$ let G_x be the set

$$G_x = \prod_k I_k(x(k))$$

—so G_x is the product of the open sets $I_k(x(k))$, each centred on the k^{th} component of x and with diameter at most $\frac{2}{k}$. The set G_x is open with respect to the box-topology and can be pictured as an ever narrowing “tube” centred on $x = (x(k))$.

Clearly, $\{G_x : x \in \prod_k X_k\}$ is an open cover of $\prod_k X_k$ (for the box-topology). We shall argue that this cover has no finite subcover — this because the tails of the G_x 's become too narrow. Indeed, for any points x_1, \dots, x_n in $\prod_k X_k$, and any $m \in \mathbb{N}$, we have

$$p_m(G_{x_1} \cup \dots \cup G_{x_n}) = I_m(x_1(m)) \cup \dots \cup I_m(x_n(m)).$$

Each of the n intervals $I_m(x_j(m))$ has diameter not greater than $\frac{2}{m}$, so any interval covered by their union cannot have length greater than $\frac{2n}{m}$. If we choose $m > 3n$, then this union cannot cover any interval of length greater than $\frac{2}{3}$, and in particular, it cannot cover X_m . It follows that G_{x_1}, \dots, G_{x_n} is not a cover for $\prod_k X_k$ and, consequently, $\prod_k X_k$ is not compact with respect to the box-topology.

That this behaviour cannot occur with the product topology — this being the content of Tychonov's theorem which shall now discuss. It is convenient to first prove a result on the existence of a certain family of sets satisfying the finite intersection property (fip).

Proposition 10.10. *Suppose that \mathcal{F} is any collection of subsets of a given set X satisfying the fip. Then there is a maximal collection \mathcal{D} containing \mathcal{F} and satisfying the fip, i.e., if $\mathcal{F} \subseteq \mathcal{F}'$ and if \mathcal{F}' satisfies the fip, then $\mathcal{F}' \subseteq \mathcal{D}$. Furthermore;*

- (i) *if $A_1, \dots, A_n \in \mathcal{D}$, then $A_1 \cap \dots \cap A_n \in \mathcal{D}$;*
- (ii) *if A is any subset of X such that $A \cap D \neq \emptyset$ for all $D \in \mathcal{D}$, then $A \in \mathcal{D}$.*

Proof. As might be expected, we shall use Zorn's lemma. Let \mathcal{C} denote the collection of those families of subsets of X which contain \mathcal{F} and satisfy the fip. Then $\mathcal{F} \in \mathcal{C}$, so \mathcal{C} is not empty. Evidently, \mathcal{C} is ordered by set-theoretic inclusion. Suppose that Φ is a totally ordered set of families in \mathcal{C} . Let $A = \bigcup_{\mathcal{S} \in \Phi} \mathcal{S}$. Then $\mathcal{F} \subseteq A$, since $\mathcal{F} \subseteq \mathcal{S}$, for all $\mathcal{S} \in \Phi$. We shall show

that A satisfies the fip. To see this, let $S_1, \dots, S_n \in A$. Then each S_i is an element of some family \mathcal{S}_i that belongs to Φ . But Φ is totally ordered and so there is \mathcal{S}_0 such that $\mathcal{S}_i \subseteq \mathcal{S}_0$ for all $1 \leq i \leq n$. Hence $S_1, \dots, S_n \in \mathcal{S}_0$ and so $S_1 \cap \dots \cap S_n \neq \emptyset$ since \mathcal{S}_0 satisfies the fip. It follows that A is an upper bound for Φ in \mathcal{C} . Hence, by Zorn's lemma, \mathcal{C} contains a maximal element, \mathcal{D} , say.

- (i) Now suppose that $A_1, \dots, A_n \in \mathcal{D}$ and let $B = A_1 \cap \dots \cap A_n$. Let $\mathcal{D}' = \mathcal{D} \cup \{B\}$. Then any finite intersection of members of \mathcal{D}' is equal to a finite intersection of members of \mathcal{D} . Thus \mathcal{D}' satisfies the fip. Clearly, $\mathcal{F} \subseteq \mathcal{D}'$, and so, by maximality, we deduce that $\mathcal{D}' = \mathcal{D}$. Thus $B \in \mathcal{D}$.

(ii) Suppose that $A \subseteq X$ and that $A \cap D \neq \emptyset$ for every $D \in \mathcal{D}$. Let $\mathcal{D}' = \mathcal{D} \cup \{A\}$, and let $D_1, \dots, D_m \in \mathcal{D}'$. If $D_i \in \mathcal{D}$, for all $1 \leq i \leq m$, then $D_1 \cap \dots \cap D_m \neq \emptyset$ since \mathcal{D} satisfies the fip. If some $D_i = A$ and some $D_j \neq A$, then $D_1 \cap \dots \cap D_m$ has the form $D_1 \cap \dots \cap D_k \cap A$ with $D_1, \dots, D_k \in \mathcal{D}$. By (i), $D_1 \cap \dots \cap D_k \in \mathcal{D}$ and so, by hypothesis, $A \cap (D_1 \cap \dots \cap D_k) \neq \emptyset$. Hence \mathcal{D}' satisfies the fip and, again by maximality, we have $\mathcal{D}' = \mathcal{D}$ and thus $A \in \mathcal{D}$. ■

We are now ready to prove Tychonov's theorem. In fact we shall present three proofs. The first is based on the previous proposition, the second uses the idea of partial cluster points together with Zorn's lemma, and the third uses universal nets.

Theorem 10.11. (Tychonov's Theorem) *Let $\{(X_\alpha, \mathcal{T}_\alpha) : \alpha \in I\}$ be any given collection of compact topological spaces. Then the Cartesian product $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{product}})$ is compact.*

Proof. (1st proof) Let \mathcal{F} be any family of closed subsets of $\prod_\alpha X_\alpha$ satisfying the fip. We must show that $\bigcap_{F \in \mathcal{F}} F \neq \emptyset$. By the previous proposition, there is a maximal family \mathcal{D} of subsets of $\prod_\alpha X_\alpha$ satisfying the fip and with $\mathcal{F} \subseteq \mathcal{D}$. (Note that the members of \mathcal{D} need not all be closed sets.)

For each $\alpha \in I$, consider the family $\{p_\alpha(D) : D \in \mathcal{D}\}$. Then this family satisfies the fip because \mathcal{D} does. Hence $\{\overline{p_\alpha(D)} : D \in \mathcal{D}\}$ satisfies the fip. But this is a collection of closed sets in the compact space $(X_\alpha, \mathcal{T}_\alpha)$, and so

$$\bigcap_{D \in \mathcal{D}} \overline{p_\alpha(D)} \neq \emptyset.$$

That is, there is some $x_\alpha \in X_\alpha$ such that $x_\alpha \in \overline{p_\alpha(D)}$ for every $D \in \mathcal{D}$.

Let $x \in \prod_\alpha X_\alpha$ be given by $p_\alpha(x) = x_\alpha$, i.e., the α^{th} coordinate of x is x_α . Now, for any $\alpha \in I$, and for any $D \in \mathcal{D}$, $x_\alpha \in \overline{p_\alpha(D)}$ implies that for any neighbourhood U_α of x_α we have $U_\alpha \cap p_\alpha(D) \neq \emptyset$. Hence $p_\alpha^{-1}(U_\alpha) \cap D \neq \emptyset$ for every $D \in \mathcal{D}$. By the previous proposition, it follows that $p_\alpha^{-1}(U_\alpha) \in \mathcal{D}$. Hence, again by the previous proposition, for any $\alpha_1, \dots, \alpha_n \in I$ and neighbourhoods $U_{\alpha_1}, \dots, U_{\alpha_n}$ of $x_{\alpha_1}, \dots, x_{\alpha_n}$, respectively,

$$p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \dots \cap p_{\alpha_n}^{-1}(U_{\alpha_n}) \in \mathcal{D}.$$

Furthermore, since \mathcal{D} has the fip, we have that

$$p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \dots \cap p_{\alpha_n}^{-1}(U_{\alpha_n}) \cap D \neq \emptyset$$

for every $D \in \mathcal{D}$, every finite family $\alpha_1, \dots, \alpha_n \in I$ and neighbourhoods $U_{\alpha_1}, \dots, U_{\alpha_n}$ of $x_{\alpha_1}, \dots, x_{\alpha_n}$, respectively.

We shall show that $x \in \overline{D}$ for every $D \in \mathcal{D}$. To see this, let G be any neighbourhood of x . Then, by definition of the product topology, there is a finite family $\alpha_1, \dots, \alpha_m \in I$ and open sets $U_{\alpha_1}, \dots, U_{\alpha_m}$ such that

$$x \in p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \dots \cap p_{\alpha_m}^{-1}(U_{\alpha_m}) \subseteq G.$$

But we have shown that for any $D \in \mathcal{D}$,

$$D \cap p_{\alpha_1}^{-1}(U_{\alpha_1}) \cap \dots \cap p_{\alpha_m}^{-1}(U_{\alpha_m}) \neq \emptyset$$

and therefore $D \cap G \neq \emptyset$. We deduce that $x \in \overline{D}$, the closure of D , for any $D \in \mathcal{D}$. In particular, $x \in \overline{F} = F$ for every $F \in \mathcal{F}$. Thus

$$\bigcap_{F \in \mathcal{F}} F \neq \emptyset$$

— it contains x . The result follows. ■

Proof. (2nd proof) Let $(\gamma_\alpha)_{\alpha \in A}$ be any given net in $X = \prod_{i \in I} X_i$. We shall show that (γ_α) has a cluster point. For each $i \in I$, $(\gamma_\alpha(i))$ is a net in the compact space X_i and therefore has a cluster point z_i , say, in X_i . However, the element $\gamma \in X$ given by $\gamma(i) = z_i$ need *not* be a cluster point of (γ_α) . (For example, let $I = \{1, 2\}$, $X_1 = X_2 = [-1, 1]$ with the usual topology and let (γ_n) be the sequence $((x_n, y_n)) = (((-1)^n, (-1)^{n+1}))$ in $X_1 \times X_2$. Then 1 is a cluster point of both (x_n) and (y_n) but $(1, 1)$ is not a cluster point of the sequence $((x_n, y_n))$.) The idea of the proof is to consider the set of partial cluster points, that is, cluster points of the net (γ_α) with respect to some subset of components. These are naturally partially ordered, and an appeal to Zorn's lemma assures the existence of a maximal such element. One shows that this is truly a cluster point of (γ_α) in the usual sense.

For given $\gamma \in X$ and $J \subseteq I$, $J \neq \emptyset$, let $\gamma \upharpoonright J$ denote the element of the partial cartesian product $\prod_{j \in J} X_j$ whose j^{th} component is given by $\gamma \upharpoonright J(j) = \gamma(j)$, for $j \in J$. In other words, $\gamma \upharpoonright J$ is obtained from γ by simply ignoring the components in each X_j for $j \notin J$. Let $g \in \prod_{j \in J} X_j$. We shall say that g is a partial cluster point of (γ_α) if g is a cluster point of the net $(\gamma_\alpha \upharpoonright J)_{\alpha \in A}$ in the topological space $\prod_{j \in J} X_j$. Let \mathcal{P} denote the collection of partial cluster points of (γ_α) . Now, for any $j \in I$, X_j is compact, by hypothesis. Hence, $(\gamma_\alpha(j))_{\alpha \in A}$ has a cluster point, x_j , say, in X_j . Set

$J = \{j\}$ and define $g \in \prod_{i \in \{j\}} X_i = X_j$ by $g(j) = x_j$. Then g is a partial cluster point of (γ_α) , and therefore \mathcal{P} is not empty.

The collection \mathcal{P} is partially ordered by extension: that is, if g_1 and g_2 are elements of \mathcal{P} , where $g_1 \in \prod_{j \in J_1} X_j$ and $g_2 \in \prod_{j \in J_2} X_j$, we say that $g_1 \preceq g_2$ if $J_1 \subseteq J_2$ and $g_1(j) = g_2(j)$ for all $j \in J_1$. Let $\{g_\lambda \in \prod_{j \in J_\lambda} X_j : \lambda \in \Lambda\}$ be any totally ordered family in \mathcal{P} . Set $J = \bigcup_{\lambda \in \Lambda} J_\lambda$ and define $g \in \prod_{j \in J} X_j$ by setting $g(j) = g_\lambda(j)$, $j \in J$, where λ is such that $j \in J_\lambda$. Then g is well-defined because $\{g_\lambda : \lambda \in \Lambda\}$ is totally ordered. It is clear that $g \succeq g_\lambda$ for each $\lambda \in \Lambda$. We claim that g is a partial cluster point of (γ_α) . Indeed, let G be any neighbourhood of g in $X_J = \prod_{j \in J} X_j$. Then there is a finite set F in J and open sets $U_j \in X_j$, for $j \in F$, such that $g \in \bigcap_{j \in F} p_j^{-1}(U_j) \subseteq G$. By definition of the partial order on \mathcal{P} , it follows that there is some $\lambda \in \Lambda$ such that $F \subseteq J_\lambda$, and therefore $g(j) = g_\lambda(j)$, for $j \in F$. Now, g_λ belongs to \mathcal{P} and so is a cluster point of the net $(\gamma_\alpha \upharpoonright J_\lambda)_{\alpha \in A}$. It follows that for any $\alpha \in A$ there is $\alpha' \succeq \alpha$ such that $p_j(\gamma_{\alpha'}) \in U_j$ for every $j \in F$. Thus $\gamma_{\alpha'} \in G$, and we deduce that g is a cluster point of $(\gamma_\alpha \upharpoonright J)_\alpha$. Hence g is a partial cluster point of (γ_α) and so belongs to \mathcal{P} .

We have shown that any totally ordered family in \mathcal{P} has an upper bound and hence, by Zorn's lemma, \mathcal{P} possesses a maximal element, γ , say. We shall show that the maximality of γ implies that it is, in fact, not just a partial cluster point but a cluster point of the net (γ_α) . To see this, suppose that $\gamma \in \prod_{j \in J} X_j$, with $J \subseteq I$, so that γ is a cluster point of $(\gamma_\alpha \upharpoonright J)_{\alpha \in A}$. We shall show that $J = I$. By way of contradiction, suppose that $J \neq I$ and let $k \in I \setminus J$. Since γ is a cluster point of $(\gamma_\alpha \upharpoonright J)_{\alpha \in A}$ in $\prod_{j \in J} X_j$, it is the limit of some subnet $(\gamma_{\phi(\beta)} \upharpoonright J)_{\beta \in B}$, say. Now, $(\gamma_{\phi(\beta)}(k))_{\beta \in B}$ is a net in the compact space X_k and therefore has a cluster point, $\xi \in X_k$, say. Let $J' = J \cup \{k\}$ and define $\gamma' \in \prod_{j \in J'} X_j$ by

$$\gamma'(j) = \begin{cases} \gamma(j) & j \in J \\ \xi & j = k \end{cases}.$$

We shall show that γ' is a cluster point of $(\gamma_\alpha \upharpoonright J')_{\alpha \in A}$. Let F be any finite subset in J and, for $j \in F$, let U_j be any open neighbourhood of $\gamma'(j)$ in X_j , and let V be any open neighbourhood of $\gamma'(k) = \xi$ in X_k . Since $(\gamma_{\phi(\beta)})_{\beta \in B}$ converges to γ in $\prod_{j \in J} X_j$, there is $\beta_1 \in B$ such that $\gamma_{\phi(\beta)}(j)_{\beta \in B} \in U_j$ for each $j \in F$ for all $\beta \succeq \beta_1$. Furthermore, $(\gamma_{\phi(\beta)}(k))_{\beta \in B}$ is frequently in V . Let $\alpha_0 \in A$ be given. There is $\beta_0 \in B$ such that if $\beta \succeq \beta_0$ then $\phi(\beta) \succeq \alpha_0$. Let $\beta_2 \in B$ be such that $\beta_2 \succeq \beta_0$ and $\beta_2 \succeq \beta_1$. Since $(\gamma_{\phi(\beta)}(k))_{\beta \in B}$ is frequently in V , there is $\beta \succeq \beta_2$ such that $\gamma_{\phi(\beta)}(k) \in V$. Set $\alpha = \phi(\beta) \in A$. Then $\alpha \succeq \alpha_0$, $\gamma_\alpha(k) \in V$ and, for $j \in F$, $\gamma_\alpha(j) = \gamma_{\phi(\beta)}(j) \in U_j$. It follows that

γ' is a cluster point of the net $(\gamma_\alpha \upharpoonright J')_{\alpha \in A}$, as required. This means that $\gamma' \in \mathcal{P}$. However, it is clear that $\gamma \preceq \gamma'$ and that $\gamma \neq \gamma'$. This contradicts the maximality of γ in \mathcal{P} and we conclude that, in fact, $J = I$ and therefore γ is a cluster point of $(\gamma_\alpha)_{\alpha \in A}$.

We have seen that any net in X has a cluster point and therefore it follows that X is compact. ■

Finally, we will consider a proof using universal nets.

Proof. (3rd proof) Let $(\gamma_\alpha)_{\alpha \in A}$ be any universal net in $X = \prod_{i \in I} X_i$. For any $i \in I$, let S_i be any given subset of X_i and let S be the subset of X given by

$$S = \{\gamma \in X : \gamma(i) \in S_i\}.$$

Then (γ_α) is either eventually in S or eventually in $X \setminus S$. Hence we have that either $(\gamma_\alpha(i))$ is either eventually in S_i or eventually in $X_i \setminus S_i$. In other words, $(\gamma_\alpha(i))_{\alpha \in A}$ is a universal net in X_i . Since X_i is compact, by hypothesis, $(\gamma_\alpha(i))$ converges; $\gamma_\alpha(i) \rightarrow x_i$, say, for $i \in I$. Let $\gamma \in X$ be given by $\gamma(i) = x_i$, $i \in I$. Then we have that $p_i(\gamma_\alpha) = \gamma_\alpha(i) \rightarrow x_i = \gamma(i)$ for each $i \in I$ and therefore $\gamma_\alpha \rightarrow \gamma$ in X . Thus every universal net in X converges, and we conclude that X is compact. ■

We will use Tychonov's theorem to show that the unit ball in the dual space of a normed space is compact in the w^* -topology. To do this, it is necessary to consider the unit ball of the dual space as a suitable cartesian product. By way of a preamble, let us discuss the dual space X^* of the normed space X as a cartesian product. Each element ℓ in X^* is a (linear) function on X . The collection of values $\ell(x)$, as x runs over X , can be thought of as an element of a cartesian product with components given by the $\ell(x)$. Specifically, for each $x \in X$, let Y_x be a copy of \mathbb{C} , equipped with its usual topology. Let $Y = \prod_{x \in X} Y_x = \prod_{x \in X} \mathbb{C}$, equipped with the product topology. To each element $\ell \in X^*$, we associate the element $\gamma_\ell \in Y$ given by $\gamma_\ell(x) = \ell(x)$, i.e., the x^{th} -coordinate of γ_ℓ is $\ell(x) \in \mathbb{C} = Y_x$.

If $\ell_1, \ell_2 \in X^*$, and if $\gamma_{\ell_1} = \gamma_{\ell_2}$, then γ_{ℓ_1} and γ_{ℓ_2} have the same coordinates so that $\ell_1(x) = \gamma_{\ell_1}(x) = \gamma_{\ell_2}(x) = \ell_2(x)$ for all $x \in X$. In other words, $\ell_1 = \ell_2$, and so the correspondence $\ell \mapsto \gamma_\ell$ of $X^* \rightarrow Y$ is one-one. Thus X^* can be thought of as a subset of $Y = \prod_{x \in X} \mathbb{C}$.

Suppose now that $\{\ell_\alpha\}$ is a net in X^* such that $\ell_\alpha \rightarrow \ell$ in X^* with respect to the w^* -topology. This is equivalent to the statement that $\ell_\alpha(x) \rightarrow \ell(x)$ for each $x \in X$. But then $\ell_\alpha(x) = p_x(\gamma_{\ell_\alpha}) \rightarrow \ell(x) = p_x(\gamma_\ell)$ for all $x \in X$,

which, in turn, is equivalent to the statement that $\gamma_{\ell_\alpha} \rightarrow \gamma_\ell$ with respect to the product topology on $\prod_{x \in X} \mathbb{C}$.

We see, then, that the correspondence $\ell \leftrightarrow \gamma_\ell$ respects the convergence of nets when X^* is equipped with the w^* -topology and Y with the product topology. It will not come as a surprise that this also respects compactness.

Consider now X_1^* , the unit ball in the dual of the normed space X . For any $x \in X$ and $\ell \in X_1^*$, we have that $|\ell(x)| \leq \|x\|$. Let B_x denote the ball in \mathbb{C} given by

$$B_x = \{\zeta \in \mathbb{C} : |\zeta| \leq \|x\|\}.$$

Then the above remark is just the observation that $\ell(x) \in B_x$ for every $\ell \in X_1^*$. We equip B_x with its usual metric topology, so that it is compact. Let $Y = \prod_{x \in X} B_x$ equipped with the product topology. Then, by Tychonov's theorem, Y is compact.

Let $\ell \in X_1^*$. Then, as above, ℓ determines an element γ_ℓ of Y by setting

$$\gamma_\ell(x) = p_x(\gamma_\ell) = \ell(x) \in B_x.$$

The mapping $\ell \mapsto \gamma_\ell$ is one-one. Let \widehat{Y} denote the image of X_1^* under this map ;

$$\widehat{Y} = \{\gamma \in Y : \gamma = \gamma_\ell \text{ some } \ell \in X_1^*\}.$$

Proposition 10.12. \widehat{Y} is closed in Y .

Proof. Let (γ_λ) be a net in \widehat{Y} such that $\gamma_\lambda \rightarrow \gamma$ in Y . Then $p_x(\gamma_\lambda) \rightarrow p_x(\gamma)$ in B_x , for each $x \in X$. Each γ_λ is of the form γ_{ℓ_λ} for some $\ell_\lambda \in X_1^*$. Hence

$$p_x(\gamma_{\ell_\lambda}) = \ell_\lambda(x) \rightarrow \gamma(x)$$

for each $x \in X$. It follows that for any $a \in \mathbb{C}$ and elements $x_1, x_2 \in X$

$$\begin{aligned} \gamma(ax_1 + x_2) &= \lim \ell_\lambda(ax_1 + x_2) \\ &= \lim a\ell_\lambda(x_1) + \ell_\lambda(x_2) \\ &= a\gamma(x_1) + \gamma(x_2). \end{aligned}$$

That is, the map $x \mapsto \gamma(x)$ is linear on X . Furthermore, $\gamma(x) = p_x(\gamma) \in B_x$, i.e., $|\gamma(x)| \leq \|x\|$, for $x \in X$. We conclude that the mapping $x \mapsto \gamma(x)$ defines an element of X_1^* . In other words, if we set $\ell(x) = \gamma(x)$, $x \in X$, then $\ell \in X_1^*$ and $\gamma_\ell = \gamma$. That is, $\gamma \in \widehat{Y}$ and so \widehat{Y} is closed, as required. ■

Remark 10.13. We know that Y is compact and since \widehat{Y} is closed in Y , we conclude that \widehat{Y} is also compact.

Theorem 10.14. (Banach-Alaoglu) *Let X be a normed space and let X_1^* denote the unit ball in X^* , the dual of X ;*

$$X_1^* = \{\ell \in X^* : \|\ell\| \leq 1\}.$$

Then X_1^ is a compact subset of X^* with respect to the w^* -topology.*

Proof. First let us note that X_1^* is closed in X^* with respect to the w^* -topology. To see this, let (ℓ_α) be a net in X_1^* such that $\ell_\alpha \rightarrow \ell$ in X^* in the w^* -topology. This means that $\ell_\alpha(x) \rightarrow \ell(x)$ for each $x \in X$. But $|\ell_\alpha(x)| \leq \|x\|$ for each α and so we also have that $|\ell(x)| \leq \|x\|$. Thus $\ell \in X_1^*$ and therefore X_1^* is w^* -closed, as claimed.

To show compactness, we will exploit the above identification with \widehat{Y} . Suppose $\{F_\beta\}$ is a family of closed sets in X_1^* satisfying the finite intersection property. The proof is complete if we can show that the whole family has a non-empty intersection.

Let A_β denote the image of F_β in Y — so $A_\beta \in \widehat{Y}$. We claim that A_β is closed in Y . To see this, suppose that (γ_λ) is a net in A_β such that $\gamma_\lambda \rightarrow \gamma$ in Y . Each γ_λ has the form $\gamma_\lambda = \gamma_{\ell_\lambda}$ for some $\ell_\lambda \in F_\beta$. Now, $A_\beta \subseteq \widehat{Y}$ and \widehat{Y} is closed and so we have that $\gamma \in \widehat{Y}$; that is, $\gamma = \gamma_\ell$ for some $\ell \in X_1^*$. But $\gamma_{\ell_\lambda} \rightarrow \gamma_\ell$ in Y implies that $\ell_\lambda \rightarrow \ell$ in X_1^* with respect to the w^* -topology. Since $\ell_\lambda \in F_\beta$ and F_β is w^* -closed we deduce that $\ell \in F_\beta$. Hence $\gamma_\ell \in A_\beta$ and therefore A_β is closed.

Now, $\{A_\beta\}$ also has the finite intersection property and since A_β is closed, for each β , and \widehat{Y} is compact, we deduce that $\bigcap_\beta A_\beta \neq \emptyset$ and therefore $\bigcap_\beta F_\beta \neq \emptyset$. It follows that X_1^* is w^* -compact. ■

Example 10.15. Let $X = \ell^\infty$ and for each $n \in \mathbb{N}$ let $\ell_m : X \rightarrow \mathbb{C}$ be the map $\ell_m(x) = x_m$, where $x = (x_n) \in \ell^\infty$. Thus ℓ_m is simply the evaluation of the m^{th} coordinate on ℓ^∞ .

We have $|\ell_m(x)| = |x_m| \leq \|x\|_\infty$ and so we see that $\ell_m \in X_1^*$ for each $m \in \mathbb{N}$. We claim that the sequence $(\ell_m)_{m \in \mathbb{N}}$ in X_1^* has *no* w^* -convergent subsequence, despite the fact that X_1^* is w^* -compact. Indeed, let $(\ell_{m_k})_{k \in \mathbb{N}}$ be *any* subsequence. Then $\ell_{m_k} \rightarrow \ell$ in the w^* -topology if and only if $\ell_{m_k}(x) \rightarrow \ell(x)$ in \mathbb{C} for every $x \in X = \ell^\infty$. Let z be the particular element of $X = \ell^\infty$

given by $z = (z_n)$ where

$$z_n = \begin{cases} 1, & n = m_{2j}, \quad j \in \mathbb{N}, \\ -1, & \text{otherwise.} \end{cases}$$

Then $\ell_{m_k}(z) = 1$ if k is even and is equal to -1 if k is odd. So the subsequence $(\ell_{m_k}(z))$ cannot converge in \mathbb{C} .