1. Topological Spaces

The notion of open set plays an important rôle in the theory of metric spaces. Our starting point is to extract and abstract the basic properties enjoyed by such sets. This leads to a whole new area of study—topological spaces.

Definition 1.1 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 provided that the following hold;

(i) $\varnothing \in \mathcal{T}$, and $X \in \mathcal{T}$;

(ii) if $\{U_\alpha : \alpha \in J\}$ is any collection of elements of $\mathcal{T}$, labelled by some index set 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*. A topological space is a pair $(X, \mathcal{T})$ where X is a non-empty set and $\mathcal{T}$ is a topology on X.

Examples 1.2

1. For any X, let $\mathcal{T}$ be the set of all subsets of X. This topology is called the discrete topology on X.

2. For any X, let $\mathcal{T} = \{\varnothing, X\}$. $\mathcal{T}$ is called the indiscrete topology on X.

3. Let $X = \{0, 1, 2\}$ and let $\mathcal{T} = \{\,\varnothing, X, \{0\}, \{1, 2\}\,\}$.

4. Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces, and let Z be the cartesian product $Z = X \times Y$. Let $\mathcal{T}'$ consist of $\varnothing$ together with all those non-empty subsets G of Z with the property that for any $z = (x, y) \in G$ there are sets $U \in \mathcal{T}$ and $V \in \mathcal{S}$ such that $z \in U \times V \subseteq G$. One checks that $\mathcal{T}'$ is a topology on $X \times Y$. This topology is called the product topology.

5. Let X be any metric space and let $\mathcal{T}$ be the set of open sets in the usual metric space sense. Then it is a theorem in metric space theory that $\mathcal{T}$ is, indeed, a topology on X. Thus, the notion of topological space generalizes that of metric space.

6. Suppose that $\{\mathcal{T}_\lambda\}_{\lambda \in \Lambda}$ is a family of topologies on a set X. Then one readily checks that $\mathcal{T} = \bigcap_\lambda \mathcal{T}_\lambda$ is a topology on X.

Definition 1.3 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 5 above.

Remark 1.4 Not every topology is metrizable. The space in example 1 above is not metrizable whenever X consists of more than one point. (If it were metrizable and contained distinct points a and b, say, then the set $\{x : d(x,b) < d(a,b)\}$ would be a non-empty proper open subset.) Of course, even if a topological space is metrizable, the metric will be far from unique—for example, proportional metrics generate the same collection of open sets.

Definition 1.5 For any non-empty subset A of a topological space $(X, \mathcal{T})$, the induced (or relative) 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.)

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

Definition 1.6 Suppose that $(X, \mathcal{T})$ is a topological space. A subset F of X is said to be closed if and only if its complement $X \setminus F$ is open, that is, belongs to $\mathcal{T}$. (It follows immediately that if $\{F_\alpha\}$ is any collection of closed sets then $\bigcap_\alpha F_\alpha$ is closed. Indeed, $X \setminus \bigcap_\alpha F_\alpha = \bigcup_\alpha (X \setminus F_\alpha)$, which belongs to $\mathcal{T}$.)

A point $a \in X$ is an interior point of the subset A of 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. To see this, suppose that each point of G is an interior point of G. Then for each $x \in G$ there is $U_x \in \mathcal{T}$ such that $x \in U_x \subseteq G$. Hence $G = \bigcup_{x \in G} U_x \in \mathcal{T}$. The converse is clear—take $U = G$.)

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

The point x is a limit point (or accumulation point) of the set A if and only if for every open set U containing x, it is true that $U \cap A$ contains some point distinct from x, i.e., $A \cap \{U \setminus \{x\}\} \neq \varnothing$. Note that x need not belong to A.

The point $a \in A$ is said to be an isolated point of A if there is an open set U such that $U \cap A = \{a\}$. (In other words, there is some open set containing a but no other points of A.)

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

$$\overline{A} = A \cup \{x \in X : x \text{ is a limit point of } A\}.$$

It follows from the definition that $x \in \overline{A}$ if and only if $A \cap U \neq \varnothing$ for any open set U containing x. Indeed, suppose that $x \in \overline{A}$ and that U is some open set containing x. Then either $x \in A$ or x is a limit point of A (or both), in which case $A \cap U \neq \varnothing$. On the other hand, suppose that $A \cap U \neq \varnothing$ for any open set U containing x. Then if x is not an element of A it is certainly a limit point. Thus $x \in \overline{A}$.

Proposition 1.7 *The closure 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 $y \in X \setminus \overline{A}$. By the above remark, there is an open set U such that $y \in U$ and $U \cap A = \varnothing$. But then, by the same remark, no point of U can belong to $\overline{A}$. In other words, $U \subseteq X \setminus \overline{A}$ and we see that y is an interior point of $X \setminus \overline{A}$. Thus all points of $X \setminus \overline{A}$ are interior points and therefore $X \setminus \overline{A}$ is open. It follows, by definition, that $\overline{A}$ is closed. Since $A \subseteq \overline{A}$ we see that $\bigcap \{F : F \text{ is closed and } F \supseteq A\} \subseteq \overline{A}$.

Now suppose that F is closed and that $A \subseteq F$. We claim that $\overline{A} \subseteq F$. Indeed, $X \setminus F$ is open and $(X \setminus F) \cap A = \varnothing$ so that no point of $X \setminus F$ can be a member of $\overline{A}$. Hence $\overline{A} \subseteq F$, as claimed. Thus $\overline{A} \subseteq \bigcap \{F : F \text{ is closed and } F \supseteq A\}$ and the result follows. ■

Corollary 1.8 *A subset A of a topological space is closed if and only if $A = \overline{A}$. Moreover, for any subset A, $\overline{\overline{A}} = \overline{A}$.*

Proof If A is closed, then A is surely the smallest closed set containing A. Thus $A = \overline{A}$. On the other hand, if $A = \overline{A}$ then A is closed because $\overline{A}$ is. Now let A be arbitrary. Then $\overline{A}$ is closed and so is equal to its closure, as above. That is, $\overline{\overline{A}} = \overline{A}$. ■

Definition 1.9 A family of open sets $\{U_\alpha : \alpha \in J\}$ is said to be an open cover of a set B in a topological space if $B \subseteq \bigcup_\alpha U_\alpha$.

Definition 1.10 A subset K of 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.

Proposition 1.11 *For a subset K 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 $K \cap \bigcap_{\alpha \in J} F_\alpha = \varnothing$, then $K \cap \bigcap_{\alpha \in I} F_\alpha = \varnothing$ for some finite subset $I \subseteq J$.*

(iii) *If $\{F_\alpha\}_{\alpha \in J}$ is any family of closed sets in X such that $K \cap \bigcap_{\alpha \in I} F_\alpha \neq \varnothing$ for every finite subset $I \subseteq J$, then $K \cap \bigcap_{\alpha \in J} F_\alpha \neq \varnothing$.*

Proof The statements (ii) and (iii) are contrapositives. We shall show that (i) and (ii) are equivalent. The proof rests on the observation that if $\{U_\alpha\}$ is any collection of sets, then $K \subseteq \bigcup_\alpha U_\alpha$ if and only if $K \cap \bigcap_\alpha (X \setminus U_\alpha) = \varnothing$. We first show that (i) implies (ii). Suppose that K is compact and let $\{F_\alpha\}$ be a given family of closed sets such that $K \cap \bigcap_\alpha F_\alpha = \varnothing$. Put $U_\alpha = X \setminus F_\alpha$. Then each U_α is open, and, by the above observation, $K \subseteq \bigcup_\alpha U_\alpha$. But then there is a finite set I such that $K \subseteq \bigcup_{\alpha \in I} U_\alpha$, and so $K \cap \bigcap_{\alpha \in I} F_\alpha = \varnothing$, which proves (ii).

Now suppose that (ii) holds, and let $\{U_\alpha\}$ be an open cover of K. Then each $X \setminus U_\alpha$ is closed and $K \cap \bigcap_\alpha (X \setminus U_\alpha) = \varnothing$. By (ii), there is a finite set I such that $K \cap \bigcap_{\alpha \in I} (X \setminus U_\alpha) = \varnothing$. This is equivalent to the statement that $K \subseteq \bigcup_{\alpha \in I} U_\alpha$. Hence K is compact. ∎

Remark 1.12 We say that a family $\{A_\alpha\}_{\alpha \in J}$ has the finite intersection property if $\bigcap_{\alpha \in I} A_\alpha \neq \varnothing$ for each finite subset I in J. Thus, we can say that a topological space $(X, \mathcal{T})$ is compact if and only if any family of closed sets $\{F_\alpha\}_{\alpha \in J}$ in X having the finite intersection property is such that $\bigcap_{\alpha \in J} F_\alpha \neq \varnothing$.

Definition 1.13 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 $U \subseteq N$.

Note that N need not itself be open. For example, in any metric space (X, d), the closed sets $\{x \in X : d(a, x) \le r\}$, for $r > 0$, are neighbourhoods of the point a.

We note also that a set U belongs to $\mathcal{T}$ if and only if U is a neighbourhood of each of its points. (Indeed, to say that U is a neighbourhood of x is to say that x is an interior point of U. We have already observed that a set is open if and only if each of its points is an interior point and so this is the same as saying that it is a neighbourhood of each of its points.)

Definition 1.14 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 $U \cap V = \varnothing$.

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—take $\{x, y\}$ to be $\{1, 2\}$.

Proposition 1.15 *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 required. ∎

Remark 1.16 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 1.17 *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 $U_x \cap V_x = \varnothing$. 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 $V \cap U_{x_i} = \varnothing$ for $1 \le i \le n$. Hence $V \cap K = \varnothing$. We have $z \in V$ and $V \subseteq X \setminus K$ which implies that z is an interior point of $X \setminus K$. Thus, $X \setminus K$ is open, and K is closed. ∎

Example 1.18 Let $X = \{0, 1, 2\}$ and $\mathcal{T} = \{\varnothing, X, \{0\}, \{1, 2\}\}$. 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. (In fact, if a topology on a set X only contains a finite number of elements, i.e., there are only finitely-many open sets, then *all* subsets of X are compact. For such spaces the concept of compactness is not very interesting!)

Example 1.19 It is perhaps surprising to discover that the closure of a compact set need not be compact. For example, let $X = \mathbb{N}$ and let $\mathcal{T}$ be the (non-Hausdorff) topology on $\mathbb{N}$ whose non-empty sets are precisely those subsets of $\mathbb{N}$ which contain 1. Set $K = \{1\}$. Then K is compact, trivially, and $\overline{K} = \mathbb{N}$—every open neighbourhood of any given $n \in \mathbb{N}$ contains 1 and so meets K and so $n \in \overline{K}$. However, $\mathbb{N}$ is clearly not compact—for example, the open cover $\{G_n : n \in \mathbb{N}\}$, where $G_n = \{1, n\}$, has no finite subcover.

Proposition 1.20 *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 with the open set $X \setminus F$. This gives an open cover of K;

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

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

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

It follows that

$$F \subseteq U_{\alpha_1} \cup U_{\alpha_2} \cup \cdots \cup U_{\alpha_m}$$

and 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 1.21 Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces and suppose that $f : X \to Y$ is a given mapping. We say that f is continuous if and only if $f^{-1}(V) \in \mathcal{T}$ for any $V \in \mathcal{S}$. (By taking complements, this is equivalent to $f^{-1}(F)$ being closed for every F closed in Y.)

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

Theorem 1.22 *Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces with $(X, \mathcal{T})$ compact, and suppose that $f : X \to Y$ is continuous. Then $f(X)$ is compact in $(Y, \mathcal{S})$. (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 $f(X)$. Then $\{f^{-1}(V_\alpha)\}$ is an open cover of X and so, by compactness, there are indices $\alpha_1, \alpha_2, \ldots, \alpha_m$ such that

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

It follows that

$$f(X) \subseteq V_{\alpha_1} \cup \cdots \cup V_{\alpha_m}$$

and hence $f(X)$ is compact. ■

Theorem 1.23 *Suppose that $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ are topological spaces and that $(Y, \mathcal{S})$ is a Hausdorff space. Suppose that $f : X \to Y$ is a continuous injection. Then $(X, \mathcal{T})$ is Hausdorff.*

Proof Suppose that $a, b \in X$ with $a \neq b$. The injectivity of f implies that $f(a) \neq f(b)$. Since $(Y, \mathcal{S})$ is Hausdorff, there are disjoint elements V_1, V_2 in $\mathcal{S}$ such that $f(a) \in V_1$ and $f(b) \in V_2$. Put $U_1 = f^{-1}(V_1)$ and $U_2 = f^{-1}(V_2)$. Then $a \in U_1$, $b \in U_2$ and $U_1 \cap U_2 = \varnothing$. Furthermore, the continuity of f implies that U_1 and U_2 both belong to $\mathcal{T}$. We conclude that $(X, \mathcal{T})$ is Hausdorff. ∎

As a corollary to this result, we can say that there can be *no* continuous injection from a non-Hausdorff space into a Hausdorff space.

Theorem 1.24 *Let X be a compact topological space, Y a Hausdorff topological space and $f : X \to Y$ a continuous injective surjection. Then $f^{-1} : Y \to X$ exists and is continuous.*

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

A homeomorphism $f : X \to Y$ sets up a one-one correspondence between the open sets in X and those in Y, via $U \leftrightsquigarrow f(U)$. 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 1.26 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 $\mathcal{T}_1 \subseteq \mathcal{T}_2$ (or, 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 $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ 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 $f : X \to Y$ is continuous.

Theorem 1.27 *Suppose that $\mathcal{T}_1, \mathcal{T}_2$ are topologies on a set X such that $\mathcal{T}_1 \subseteq \mathcal{T}_2$, $\mathcal{T}_1$ is a Hausdorff topology and such that $(X, \mathcal{T}_2)$ is compact. Then $\mathcal{T}_1 = \mathcal{T}_2$.*

Proof Let $U \in \mathcal{T}_2$ and set $F = X \setminus U$. Then F is $\mathcal{T}_2$-closed and hence $\mathcal{T}_2$-compact. But any $\mathcal{T}_1$-open cover is also a $\mathcal{T}_2$-open cover, so it follows that F is $\mathcal{T}_1$-compact. Since $(X, \mathcal{T}_1)$ is Hausdorff, it follows that F is $\mathcal{T}_1$-closed and therefore U is $\mathcal{T}_1$-open. Thus $\mathcal{T}_2 \subseteq \mathcal{T}_1$ and the result follows.

■

Remark 1.28 Thus we see that if $(X, \mathcal{T})$ is a compact Hausdorff topological space, then $\mathcal{T}$ cannot be enlarged without spoiling compactness or reduced without spoiling the Hausdorff property. This expresses a rigidity of compact Hausdorff spaces.

Let X be a given (non-empty) set, let $(Y, \mathcal{S})$ be a topological space and let $f : X \to Y$ 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, since 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 mappings. Suppose that $\{(Y_\alpha, \mathcal{S}_\alpha) : \alpha \in I\}$ is a collection of topological spaces, indexed by I, and that $\mathcal{F} = \{f_\alpha : X \to Y_\alpha\}$ 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.

Definition 1.29 The topology $\mathcal{T}$, described above, is called the $\sigma(X, \mathcal{F})$-topology on X.

Theorem 1.30 *Suppose that each $(Y_\alpha, \mathcal{S}_\alpha)$ is Hausdorff and that $\mathcal{F}$ separates points of X, i.e., for any $a, b \in X$ with $a \neq b$, 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 $a \neq b$. 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 $U \cap V = \varnothing$. 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 $f_\alpha^{-1}(U) \cap f_\alpha^{-1}(V) = \varnothing$. ■

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

Definition 1.31 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 non-empty element of $\mathcal{T}$ is a union of elements of $\mathcal{B}$.

Examples 1.32

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.

4. Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces. The sets $U \times V$, with $U \in \mathcal{T}$ and $V \in \mathcal{S}$, form a base for the product topology on $X \times Y$.

Proposition 1.33 *A 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 1.34 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 1.35 The collection of subsets of $\mathbb{R}$ consisting of those intervals of the form (a, ∞) or $(-\infty, b)$, with $a, b \in \mathbb{R}$, is a sub-base for the usual topology on $\mathbb{R}$.

Proposition 1.36 *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 $\varnothing$ together with arbitrary unions of elements of members of $\mathcal{B}$ is a topology on X, and it 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, $\varnothing \in \mathcal{T}$ 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 $A \neq \varnothing$ and $B \neq \varnothing$. 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.

It is clear that $\mathcal{T}$ contains $\mathcal{S}$. Suppose that $\mathcal{T}'$ is any topology on X which also 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 $\mathcal{B}$ is a base and that $\mathcal{S}$ is a sub-base for $\mathcal{T}$. ■

Remark 1.37 We can describe the $\sigma(X, \mathcal{F})$-topology on X determined by the family of maps $\{f_\alpha \colon \alpha \in I\}$, discussed earlier, as the topology with sub-base given by the collection $\{f_\alpha^{-1}(V) : \alpha \in I,\ V \in \mathcal{S}_\alpha\}$.

Example 1.38 Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces and let $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ be the projection maps defined by $p_X\big((x,y)\big) = x$ and $p_Y\big((x,y)\big) = y$ for $(x,y) \in X \times Y$. For any $U \subseteq X$ and $V \subseteq Y$, $p_X^{-1}(U) = U \times Y$ and $p_Y^{-1}(V) = X \times V$, so that $p_X^{-1}(U) \cap p_Y^{-1}(V) = U \times V$. Now, the collection $\{U \times V : U \in \mathcal{T},\, V \in \mathcal{S}\}$ is a base for the product topology on $X \times Y$ and it is clear it has as a sub-base the collection $\{p_X^{-1}(U),\, p_Y^{-1}(V) : U \in \mathcal{T},\, V \in \mathcal{S}\}$. It follows that the product topology is the same as the $\sigma(X \times Y, \{p_X, p_Y\})$-topology and therefore the product topology is the weakest topology on $X \times Y$ such that the projection maps are continuous.

We will adopt this later as our definition for the product topology on an arbitrary cartesian product of topological spaces.

The local version of a base is often useful, not least for purposes of economy.

Definition 1.39 Let $(X, \mathcal{T})$ be a topological space, and, for $x \in X$, let $\mathcal{V}_x$ denote the collection of all neighbourhoods of x. A subset $\mathcal{N}_x \subseteq \mathcal{V}_x$ is said to be a (local) neighbourhood base at x (or a fundamental system of neighbourhoods of x) if any element of $\mathcal{V}_x$ contains some element of the subcollection $\mathcal{N}_x$.
A topological space is said to be first countable if each of its points possesses a countable neighbourhood base. It is said to be second countable if it possesses a countable base. (Evidently, a second countable space is first countable—those members of the base containing any given point constitute a countable neighbourhood base for that particular point.)

Example 1.40 In any metric space (X, d), the countable collection of disks $\big\{\{x' : d(x, x') < \frac{1}{n}\}\big\}_{n \in \mathbb{N}}$ is a neighbourhood base at the point $x \in X$. Consequently, every metrizable topological space is first countable.

The space $\mathbb{R}^n$, equipped with the usual (metric) topology, is an example of a second countable space. Indeed, a countable base is given by the collection of open balls with rational radii and centres with rational coordinates.

Proposition 1.41 *Suppose that $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ are topological spaces. A map $f : X \to Y$ is continuous if and only if for each $x \in X$ and local neighbourhood bases $\mathcal{N}_x$ and $\mathcal{N}_{f(x)}$ at x and $f(x)$, respectively, it is true that for any $V \in \mathcal{N}_{f(x)}$ there is $U \in \mathcal{N}_x$ such that $f(U) \subseteq V$.*

Proof Suppose that f is continuous. Let $x \in X$ and let $\mathcal{N}_x$ and $\mathcal{N}_{f(x)}$ be neighbourhood bases at x and $f(x)$, respectively. Let $V \in \mathcal{N}_{f(x)}$. Then there is an open set G such that $f(x) \in G \subseteq V$. By continuity, $f^{-1}(G)$ is an open neighbourhood of x and so there is some $U \in \mathcal{N}_x$ such that $x \in U \subseteq f^{-1}(G)$. Thus $f(U) \subseteq V$.

Conversely, suppose that for any $x \in X$ and local neighbourhood bases $\mathcal{N}_x$ and $\mathcal{N}_{f(x)}$ and for any $V \in \mathcal{N}_{f(x)}$, there is some $U \in \mathcal{N}_x$ such that $f(U) \subseteq V$. We wish to show that f is continuous. Let G be an open set in Y. If $f^{-1}(G) = \varnothing$, there is no more to prove, so suppose that $x \in f^{-1}(G)$. Then $f(x) \in G$ and there is some V in $\mathcal{N}_{f(x)}$ such that $V \subseteq G$. But then, by hypothesis, there is some $U \in \mathcal{N}_x$ such $f(U) \subseteq V$. Let W be any open neighbourhood of x with $W \subseteq U$. Then $f(W) \subseteq V$ and so certainly it is true that $W \subseteq f^{-1}(G)$. Hence every point of $f^{-1}(G)$ is an interior point and we conclude that $f^{-1}(G)$ is open and so f is continuous. ∎

The following is a generalisation of the standard ε-δ definition of the continuity at a given point of a real function of a real variable to the present context.

Definition 1.42 Let $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ be topological spaces and suppose that $f : X \to Y$ is a given map. We say that f is continuous at the point $x \in X$ if for any neighbourhood V of $f(x)$ in Y there is a neighbourhood U of x in X such that $U \subseteq f^{-1}(V)$, i.e., $f(U) \subseteq V$.

The previous proposition then states that a mapping from one topological space into another is continuous if and only if it is continuous at every point.

2. Nets

In a metric space, a point belongs to the closure of a given set if and only if it is the limit of some sequence of points belonging to that set. The convergence of the sequence $(a_n)_{n\in\mathbb{N}}$ to the point x is defined by the requirement that for any $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that the distance between a_n and x is less than ε whenever $n \geq N$. This is equivalent to the requirement that for any neighbourhood U of x there is some $N \in \mathbb{N}$ such that a_n belongs to U whenever $n \geq N$. We take this latter formulation, word for word, as the definition of convergence of the sequence $(a_n)_{n\in\mathbb{N}}$ in an arbitrary topological space. The resulting theory is not as straightforward as one might suspect, as is illustrated by the following example.

Example 2.1 Let X be the real interval $[0,1]$ and let $\mathcal{T}$ be the co-countable topology on X; that is, $\mathcal{T}$ consists of X and $\varnothing$ together with all those subsets G of X whose complement $X \setminus G$ is a countable set. Let $A = [0,1)$, and consider $\overline{A}$. Now, $\{1\} \notin \mathcal{T}$ because $X \setminus \{1\} = [0,1)$ is not countable. It follows that the complement of $\{1\}$ is not closed. That is to say, A is not closed. However, the closure of A is closed and contains A. This means that we must have $\overline{A} = [0,1]$, since $[0,1]$ is the only closed set containing A. Since 1 is not an element of A, it must be a limit point of A.

Is there some sequence in A which converges to 1? Suppose that $(a_n)_{n\in\mathbb{N}}$ is any sequence in A whatsoever. Let $B = \{a_1, a_2, \dots\}$ and let $G = X \setminus B$. Then $1 \in G$. Since B is countable, it follows from the definition of the topology on X that G is open. Thus G is an open neighbourhood of 1 which contains *no* member of the sequence $(a_n)_{n\in\mathbb{N}}$. Clearly, $(a_n)_{n\in\mathbb{N}}$ cannot converge to 1. *No* sequence in A can converge to the limit point 1. We have exhibited a topological space with a subset possessing a limit point which is not the limit of *any* sequence from the subset.

This example indicates that sequences may not be as useful in topological spaces as they are in metric spaces. One can, at this point, choose to abandon the use of sequences, except perhaps in favourable situations, or to seek some form of replacement. It turns out that one can keep the intuition of sequences but at the cost of an increase in formalism. As perhaps suggested by the previous example, the problem seems to be that a point in a topological space may simply have too many neighbourhoods to be penetrated by a sequence, or, put differently, a sequence may not, in general, have enough elements to populate all the neighbourhoods of the putative limit point. It is necessary to generalize the notion of a sequence allowing an index set inherently more general than the natural numbers, but retaining some concept of 'eventually greater than'. The concept of directed set is made for just this purpose. First we need the notion of a partial order.

Definition 2.2 A partially ordered set is a non-empty set P on which is defined a relation $\preceq$ (called 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 $x \preceq z$.
We sometimes write $x \succeq y$ to mean $y \preceq x$.

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. In fact, a partial order on P is more properly defined as a subset S, say, of the cartesian product $P \times P$, such that:
(i) $(x, x) \in S$, for all $x \in P$;
(ii) if (x, y) and (y, x) belong to S, then $x = y$;
(iii) if $(x, y) \in S$ and $(y, z) \in S$, then $(x, z) \in S$.
By writing $x \preceq y$ if and only if $(x, y) \in S$, we recover our original formulation above.

Examples 2.3

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

2. Let $P = \mathbb{R}$, and take $\preceq$ to be the usual ordering $\leq$ on $\mathbb{R}$.

3. Let $P = \mathbb{R}^2$, 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 2.4 A directed set is a partially ordered set I, with partial order $\preceq$, say, such that for any pair of elements $\alpha, \beta \in I$ there is some $\gamma \in I$ such that $\alpha \preceq \gamma$ and $\beta \preceq \gamma$.

It follows, by induction, that if $x_1, \ldots, x_n$ is any finite number of elements of a directed set I, there is some $z \in I$ such that $x_i \preceq z$ for all $i = 1, \ldots, n$.

Examples 2.5

1. Let $I = \mathbb{N}$ (or $\mathbb{Z}$ or $\mathbb{R}$) furnished with the usual order.

2. $\mathbb{N}$ is a directed set when equipped with the ordering given by the usual ordering on the even numbers, the usual ordering on the odd numbers, but where any even number is declared to be 'greater than' every odd number. More precisely, $\preceq$ is defined by the assignment that $2m \preceq 2n$ and $2m-1 \preceq 2n-1$ if and only if $m \leq n$ with respect to the usual order in $\mathbb{N}$, together with $2n-1 \preceq 2m$ for any $m, n \in \mathbb{N}$. Thus, for example, $5 \preceq 13$, $6 \preceq 12$ but $81 \preceq 4$.

3. Let X be any non-empty set and let I denote the collection of all subsets of X. Then I is directed with respect to the ordering given by set inclusion, that is, $A \preceq B$ if and only if $A \subseteq B$, for any subsets A, B of X.

4. The collection of all neighbourhoods, $\mathcal{V}_x$, of a given point x in a topological space, is a directed set when equipped with the partial ordering of 'reverse inclusion', that is, $U \preceq V$ if and only if $V \subseteq U$. Indeed, for any $U, V \in \mathcal{V}_x$, set $W = U \cap V$, then $W \in \mathcal{V}_x$ and both $U \preceq W$ and $V \preceq W$.

5. We can generalize the previous example slightly. Let $\mathcal{N}_x$ be any neighbourhood base at the point x in a topological space $(X, \mathcal{T})$. Order $\mathcal{N}_x$ by reverse inclusion. Then for any $U, V \in \mathcal{N}_x$, $U \cap V$ is a neighbourhood of x and so there is some $W \in \mathcal{N}_x$ such that $W \subseteq U \cap V$. Then $W \subseteq U$ and $W \subseteq V$ which is precisely the statement that $U \preceq W$ and $V \preceq W$. Hence $\mathcal{N}_x$ is directed. This example is most relevant for our purposes.

A sequence in a set X is a mapping from $\mathbb{N}$ into X, where it is implicitly understood that $\mathbb{N}$ is directed via its usual order structure. We generalize this to general directed sets.

Definition 2.6 A net in a topological space $(X, \mathcal{T})$ is a mapping from a directed set I into X; denoted $(x_\alpha)_{\alpha \in I}$.

If I is $\mathbb{N}$ with its usual ordering, then we recover the notion of a sequence. In other words, a sequence is a special case of a net.

Definition 2.7 Let $(x_\alpha)_{\alpha \in I}$ be a net in a topological space $(X, \mathcal{T})$. We say that $(x_\alpha)_{\alpha \in I}$ is eventually in the set A if there is $\beta \in I$ such that $x_\alpha \in A$ whenever $\alpha \succeq \beta$.

We say that $(x_\alpha)_{\alpha \in I}$ converges to the point $x \in X$ if, for any neighbourhood U of x, $(x_\alpha)_{\alpha \in I}$ is eventually in U. x is called a limit of $(x_\alpha)_{\alpha \in I}$, and we write $x_\alpha \to x$ (along I).

We shall see that nets serve as a general replacement for sequences, whilst retaining their intuitive appeal.

Proposition 2.8 *Let $\mathcal{N}_x$ be a neighbourhood base at a point x in a topological space $(X, \mathcal{T})$ and suppose that, for each $U \in \mathcal{N}_x$, x_U is a given point in U. Then the net $(x_U)_{\mathcal{N}_x}$ converges to x, where $\mathcal{N}_x$ is partially ordered by reverse inclusion.*

Proof For any given neighbourhood W of x there is $V \in \mathcal{N}_x$ such that $V \subseteq W$. Let $U \succeq V$. Then, by definition, $U \subseteq V$ so that $x_U \in V$. Thus $x_U \in W$ whenever $U \succeq V$ and $x_U \to x$ along $\mathcal{N}_x$. ■

We can characterize the closure of a set A as that set consisting of all the limits of nets in A.

Theorem 2.9 *Let A be a subset of a topological space $(X, \mathcal{T})$. Then $x \in \overline{A}$ if and only if there is a net $(a_\alpha)_{\alpha \in I}$ with $a_\alpha \in A$ such that $a_\alpha \to x$.*

Proof We know that a point $x \in X$ belongs to $\overline{A}$ if and only if every neighbourhood of x meets A. Suppose then that $(a_\alpha)_{\alpha \in I}$ is a net in A such that $a_\alpha \to x$. By definition of convergence, $(a_\alpha)_{\alpha \in I}$ is eventually in every neighbourhood of x, so certainly $x \in \overline{A}$.

Suppose, on the other hand, that $x \in \overline{A}$. Let $\mathcal{V}_x$ be the collection of all neighbourhoods of x ordered by reverse inclusion. Then $\mathcal{V}_x$ is a directed set. We know that for each $U \in \mathcal{V}_x$ the set $U \cap A$ is non-empty so let a_U be any element of $U \cap A$. Then $a_U \to x$. ■

The next result shows that sequences are sufficient provided that there are not too many open sets, that is, if the space is first countable. To see this, we first make an observation. Suppose that $\{A_n : n \in \mathbb{N}\}$ is a countable neighbourhood base at some given point in a topological space. For each $n \in \mathbb{N}$, let $B_n = A_1 \cap A_2 \cap \cdots \cap A_n$. Then $\{B_n : n \in \mathbb{N}\}$ is also a neighbourhood base at the given point, but has the additional property that it is nested; $B_{n+1} \subseteq B_n$, for $n \in \mathbb{N}$.

Theorem 2.10 *Suppose that $(X, \mathcal{T})$ is a first countable topological space and let $A \subseteq X$. Then $x \in \overline{A}$ if and only if there is a sequence $(a_n)_{n \in \mathbb{N}}$ in A such that $a_n \to x$.*

Proof The proof that $a_n \to x$ implies that $x \in \overline{A}$ proceeds exactly as before.

For the converse, suppose that $x \in \overline{A}$. We must exhibit a sequence in A which converges to x. Since $(X, \mathcal{T})$ is first countable, x has a countable neighbourhood base $\{B_n : n \in \mathbb{N}\}$, say. By the remark above, we may assume, without loss of generality, that $B_{n+1} \subseteq B_n$ for all $n \in \mathbb{N}$. Since $x \in \overline{A}$, $B_n \cap A \neq \varnothing$ for all $n \in \mathbb{N}$. So if we let a_n be any element of $B_n \cap A$, we obtain a sequence $(a_n)_{n \in \mathbb{N}}$ in A. We claim that $a_n \to x$ as $n \to \infty$. To see this, let G be any neighbourhood of x. Then there is some member B_N, say, of the neighbourhood base such that $B_N \subseteq G$. If $n \geq N$ then $a_n \in B_n \cap A \subseteq B_N$ and so $a_n \in G$ and it follows that $a_n \to x$ as claimed. ■

Theorem 2.11 *A set F in a topological space $(X, \mathcal{T})$ is closed if and only if no net in F can converge to a point in $X \setminus F$. If $(X, \mathcal{T})$ is first countable we can replace net by sequence in the previous statement.*

Proof Suppose that $z \notin F$ and let (x_α) be any net in F Then $X \setminus F$ is an open neighbourhood of z which contains no members of the net (x_α). Certainly (x_α) cannot be eventually in $X \setminus F$ and so no net in F can converge to z. This argument holds for sequences as well as for nets—irrespective of whether $(X, \mathcal{T})$ is first countable or not.

Now suppose that it is impossible for any net in F to converge to a point not belonging to F, and suppose that F is not closed. This means that $X \setminus F$ is not open and so not all its points are interior points. Thus there is some point $z \in X \setminus F$ such that for no open neighbourhood U of z is it true that $U \subseteq X \setminus F$. That is, $U \cap F \neq \varnothing$ for every neighbourhood U of z. For each such U, let x_U be any point in $U \cap F$. Then $(x_U)_{U \in I}$ is a net in F, where I is the family of neighbourhoods of z ordered by reverse inclusion. Evidently $x_U \to z$ along I, and we have a net in F which converges to a point not in F. This contradiction shows that F must be closed, as required.

Now suppose that $(X, \mathcal{T})$ is first countable, and that no sequence in F can converge to a point not in F. As above, if F is not closed there is some point $z \in X \setminus F$ which is not an interior point of $X \setminus F$. Now, z has a countable neighbourhood base $\{B_n : n \in \mathbb{N}\}$ such that $B_{n+1} \subseteq B_n$. Arguing as before, but with B_n instead of U, we see that there is a sequence $(x_n)_{n \in \mathbb{N}}$ with $x_n \in B_n \cap F$ for each $n \in \mathbb{N}$. Then $x_n \to z$ which is a contradiction since $z \notin F$. ∎

Proposition 2.12 *Suppose that $(X, \mathcal{T})$ is a Hausdorff topological space and that $(x_\alpha)_I$ is a net in X with $x_\alpha \to y$ and $x_\alpha \to z$. Then $y = z$. In other words, in a Hausdorff space, convergent nets have unique limits.*

Proof If $y \neq z$, there are disjoint open neighbourhoods U and V of y and z, respectively. Since $x_\alpha \to y$, there is $\beta \in I$ such that $x_\alpha \in U$ whenever $\alpha \succeq \beta$. Similarly, there is $\beta' \in I$ such that $x_\alpha \in V$ whenever $\alpha \succeq \beta'$. Let $\gamma \in I$ be such that $\beta \preceq \gamma$ and $\beta' \preceq \gamma$. Then $x_\gamma \in U \cap V$, which is impossible since $U \cap V = \varnothing$. It follows that $y = z$. ∎

Now we show that the continuity of mappings between topological spaces can be characterized using nets.

Theorem 2.13 *Let X and Y be topological spaces. A mapping $f : X \to Y$ is continuous if and only if whenever $(x_\alpha)_I$ is a net in X convergent to x then the net $(f(x_\alpha))_I$ converges to $f(x)$.*

Proof Suppose that $f : X \to Y$ is continuous and suppose that $(x_\alpha)_I$ is a net in X such that $x_\alpha \to x$. We wish to show that $(f(x_\alpha))_I$ converges to $f(x)$. To see this, let V be any neighbourhood of $f(x)$ in Y. Since f is continuous, there is a neighbourhood U of x such that $f(U) \subseteq V$. But since (x_α) converges to x, it is eventually in U, and so $(f(x_\alpha))_I$ is eventually in V, that is, $f(x_\alpha) \to f(x)$.

Conversely, suppose that $f(x_\alpha) \to f(x)$ whenever $x_\alpha \to x$. Let V be open in Y. We must show that $f^{-1}(V)$ is open in X. If this is not true, then there is a point $x \in f^{-1}(V)$ which is not an interior point. This means that every open neighbourhood of x meets $X \setminus f^{-1}(V)$, the complement of $f^{-1}(V)$. This is to say that x is a limit point of $X \setminus f^{-1}(V)$ and so there is a net $(x_\alpha)_I$ in $X \setminus f^{-1}(V)$ such that $x_\alpha \to x$. By hypothesis, it follows that $f(x_\alpha) \to f(x)$. In particular, $(f(x_\alpha))_I$ is eventually in V, that is, $(x_\alpha)_I$ is eventually in $f^{-1}(V)$. However, this contradicts the fact that $(x_\alpha)_I$ is a net in the complement of $f^{-1}(V)$. We conclude that $f^{-1}(V)$ is, indeed, open and therefore f is continuous. ■

Nets can sometimes be useful in the study of the $\sigma(X, \mathcal{F})$-topology introduced earlier.

Theorem 2.14 *Let X be a non-empty set and let $\mathcal{F}$ be a collection of maps $f_\lambda : X \to Y_\lambda$ from X into topological spaces $(Y_\lambda, \mathcal{S}_\lambda)$, for $\lambda \in \Lambda$. A net $(x_\alpha)_I$ in X converges to $x \in X$ with respect to the $\sigma(X, \mathcal{F})$-topology if and only if $(f_\lambda(x_\alpha))_I$ converges to $f_\lambda(x)$ in $(Y_\lambda, \mathcal{S}_\lambda)$ for every $\lambda \in \Lambda$.*

Proof Suppose that $x_\alpha \to x$ with respect to the $\sigma(X, \mathcal{F})$-topology on X. Fix $\lambda \in \Lambda$. Let V be any neighbourhood of $f_\lambda(x)$ in Y_λ. Then, by definition of the $\sigma(X, \mathcal{F})$-topology, $f_\lambda^{-1}(V)$ is a neighbourhood of x in X. Hence $(x_\alpha)_I$ is eventually in $f_\lambda^{-1}(V)$, which implies that $(f_\lambda(x_\alpha))_I$ is eventually in V. Thus $f_\lambda(x_\alpha) \to f_\lambda(x)$ along I, as required.

Conversely, suppose that $(x_\alpha)_I$ is a net in X such that $f_\lambda(x_\alpha) \to f_\lambda(x)$, for each $\lambda \in \Lambda$. We want to show that $(x_\alpha)_I$ converges to x with respect to the $\sigma(X, \mathcal{F})$-topology. Let U be any neighbourhood of x with respect to the $\sigma(X, \mathcal{F})$-topology. Then there is $m \in \mathbb{N}$, members $\lambda_1, \dots, \lambda_m$ of Λ, and neighbourhoods $V_{\lambda_1}, \dots, V_{\lambda_m}$ of $f_{\lambda_1}(x), \dots, f_{\lambda_m}(x)$, respectively, such that

$$x \in f_{\lambda_1}^{-1}(V_{\lambda_1}) \cap \dots \cap f_{\lambda_m}^{-1}(V_{\lambda_m}) \subseteq U.$$

By hypothesis, for each $1 \le j \le m$, there is $\beta_j \in I$ such that $f_{\lambda_j}(x_\alpha) \in V_{\lambda_j}$ whenever $\alpha \succeq \beta_j$. Let $\gamma \in I$ be such that $\gamma \succeq \beta_j$ for all $1 \le j \le m$. Then

$$x_\alpha \in f_{\lambda_1}^{-1}(V_{\lambda_1}) \cap \dots \cap f_{\lambda_m}^{-1}(V_{\lambda_m}) \subseteq U$$

whenever $\alpha \succeq \gamma$. It follows that $x_\alpha \to x$ with respect to the $\sigma(X, \mathcal{F})$-topology. ■

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 2.15 A map $F : J \to I$ between directed sets J and I 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 $F : J \to I$ 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 2.16 If we set $I = J = \mathbb{N}$, equipped with the usual ordering, and let $F : J \to I$ be any increasing map, then the subnet $(y_n) = (x_{F(n)})$ is a subsequence of the sequence (x_n).

Example 2.17 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 $F : J \to I$ 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 $x_{3m} = 0$.

Example 2.18 Let $J = \mathbb{R}$, and let $I = \mathbb{N}$, both equipped with their usual ordering, and let $F : \mathbb{R} \to \mathbb{N}$ be any function such that $F(t) \to \infty$ as $t \to \infty$. Then F is cofinal in $\mathbb{N}$ and $(x_{F(t)})_{t \in \mathbb{R}}$ is a subnet of the sequence $(x_n)_{n \in \mathbb{N}}$. This provides a simple example of a subnet of a sequence which is *not* a subsequence.

We need to introduce a little more terminology.

Definition 2.19 The net $(x_\alpha)_{\alpha \in I}$ is said to be frequently in the set A if, for any given $\gamma \in I$, $x_\alpha \in A$ for some $\alpha \in I$ with $\alpha \succeq \gamma$.

A point z is a cluster point of the net $(x_\alpha)_{\alpha \in I}$ if $(x_\alpha)_{\alpha \in I}$ is frequently in any neighbourhood of z.

Note that if $x_\alpha = x$ for every $\alpha \in I$, then x is certainly a cluster point of the net $(x_\alpha)_{\alpha \in I}$. However, x is *not* a limit point of the point set in X determined by the net, namely $\{x_\alpha : \alpha \in I\}$, since this is just the one-point set $\{x\}$, which has no limit points at all.

Proposition 2.20 *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 $F : \mathcal{E} \to I$ 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 2.21 *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 2.22 *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 $1 \le i \le k$, 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 2.23 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 2.24 *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 2.25

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 2.26 *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 $\Phi \neq \varnothing$. 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

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

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. ■

Theorem 2.27 *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 2.20 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 2.28 *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 2.29 *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 $(K, \mathcal{T}_K)$. ■

3. Product Spaces

Let $(X_1, \mathcal{T}_1)$ and $(X_2, \mathcal{T}_2)$ be topological spaces and let Y be their cartesian product

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

We have discussed the product topology on Y in example 2 of Examples 1.2 and in Example 1.38. Those sets of the form $U \times V$, with $U \in \mathcal{T}_1$ and $V \in \mathcal{T}_2$, form a base and, since $U \times V = U \times X_2 \cap X_1 \times V$, 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 product topology.

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

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

Then the product topology is the weakest topology on the cartesian product $X_1 \times X_2$ such that both p_1 and p_2 are continuous—the $\sigma(X_1 \times X_2, \{p_1, p_2\})$-topology.

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-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 $\mathcal{T}_\alpha$. Two possibilities suggest themselves. The first is the weakest topology on X with respect to which all the projection maps $p_\alpha \to X_\alpha$ 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. We take the first of these as our definition of the product topology for arbitrary products.

Definition 3.1 The product topology, denoted $\mathcal{T}_{\text{prod}}$ on the cartesian product of the topological spaces $\{(X_\alpha, \mathcal{T}_\alpha) : \alpha \in I\}$ is the $\sigma(\prod_\alpha X_\alpha, \mathcal{F})$-topology, where $\mathcal{F}$ is the family of projection maps $\{p_\alpha : \alpha \in I\}$.

Remark 3.2 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}, $1 \le i \le n$, such that

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

Hence there are open sets S_α, $\alpha \in I$, such that $\gamma \in \prod_\alpha S_\alpha \subseteq G$ and where $S_\alpha = X_\alpha$ except possibly for at most a finite number of values of α in I. 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 the whole space X_α.

Definition 3.3 The topology on the cartesian product $\prod_\alpha X_\alpha$ constructed in this way is called the box-topology on X and denoted $\mathcal{T}_{\text{box}}$.

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

Proposition 3.4 *A net (x_λ) converges in $(\prod_\alpha X_\alpha, \mathcal{T}_{\text{prod}})$ if and only if $(p_\alpha(x_\lambda))$ converges in $(X_\alpha, \mathcal{T}_\alpha)$ for each $\alpha \in I$.*

Proof This is a direct application of Theorem 2.14. ■

Example 3.5 Let $I = \mathbb{N}$, let X_k be the open interval $(-2, 2)$ for each $k \in \mathbb{N}$, and let $\mathcal{T}_k$ be the usual (Euclidean) topology on X_k. 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 $p_k(x_n) \to 0$ as $n \to \infty$, for each k and so the sequence (x_n) converges to z in $(\prod_k X_k, \mathcal{T}_{\text{prod}})$ 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.

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 3.6 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 disc in X_k with centre t and radius $\frac{1}{k}$,

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

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

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

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

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

Each of the n intervals $I_m(x_j(m))$, for $1 \le j \le n$, 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.

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

Proposition 3.7 *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 $A_1 \cap \dots \cap A_n \in \mathcal{D}$, and*

(ii) *if A is any subset of X such that $A \cap D \neq \varnothing$ 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 i_0 such that $\mathcal{S}_i \subseteq \mathcal{S}_{i_0}$ for all $1 \le i \le n$. Hence $S_1, \dots, S_n \in \mathcal{S}_{i_0}$ and so $S_1 \cap \dots \cap S_n \neq \varnothing$ since $\mathcal{S}_{i_0}$ 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 $A \cap \mathcal{D} \neq \varnothing$ 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 $1 \le i \le m$, then $D_1 \cap \dots \cap D_m \neq \varnothing$ 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, $A \cap (D_1 \cap \dots D_k) \neq \varnothing$. 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 which states that the product of compact topological spaces is compact with respect to the product topology. In fact we shall present three proofs. The first is based on the previous proposition, the second (due to P. Chernoff) uses the idea of partial cluster points together with Zorn's lemma, and the third uses universal nets.

Theorem 3.8 (Tychonov's theorem) *Let $\{(X_\alpha, \mathcal{T}_\alpha) : \alpha \in I\}$ be any given collection of compact topological spaces. Then the cartesian product $(\prod_\alpha X_\alpha, \mathcal{T}_{\mathrm{prod}})$, equipped with the product topology is compact.*

Proof (version 1) Let $\mathcal{F}$ be any family of closed subsets of $\prod_\alpha X_\alpha$ satisfying the fip. We must show that $\bigcap_{F\in\mathcal{F}} F \neq \varnothing$. 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

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

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 $U_\alpha \cap p_\alpha(D) \neq \varnothing$. Hence $p_\alpha^{-1}(U_\alpha) \cap D \neq \varnothing$ 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,

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

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

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

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

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

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

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

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

and therefore $D \cap G \neq \varnothing$. 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

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

since it contains x. The result follows. ∎

Proof (version 2) 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 $((x_n, y_n)) = \big(((-1)^n, (-1)^{n+1})\big)$ 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 $J \subseteq I$, $J \neq \varnothing$, let $\gamma \restriction J$ denote the element of the partial cartesian product $\prod_{j\in J} X_j$ whose j^{th} component is given by $\gamma \restriction J(j) = \gamma(j)$, for $j \in J$. In other words, $\gamma \restriction J$ 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 $(\gamma_\alpha \restriction J)_{\alpha\in A}$ 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 $g \in \cap_{j\in F} p_j^{-1}(U_j) \subseteq G$. 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 $(\gamma_\alpha \restriction J_\lambda)_{\alpha\in A}$. 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 $(\gamma_\alpha \restriction J)_\alpha$. 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 $(\gamma_\alpha \restriction J)_{\alpha\in A}$. 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 $(\gamma_\alpha \restriction J)_{\alpha \in A}$ in $\prod_{j\in J} X_j$, it is the limit of some subnet $(\gamma_{\phi(\beta)} \restriction J)_{\beta\in B}$, 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 $(\gamma_\alpha \restriction J')_{\alpha\in A}$. 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 $(\gamma_\alpha \restriction J')_{\alpha\in A}$, 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 (version 3) 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; $\gamma_\alpha(i) \to x_i$, say, for $i \in I$. Let $\gamma \in X$ be given by $\gamma(i) = x_i$, $i \in I$. Then we have that $p_i(\gamma_\alpha) = \gamma_\alpha(i) \to x_i = \gamma(i)$ for each $i \in I$ and therefore $\gamma_\alpha \to \gamma$ in X. Thus every universal net in X converges, and we conclude that X is compact. ■

4. Separation

The Hausdorff property of a topological space is the ability to separate distinct points. We shall extend this idea and consider the separation of disjoint closed sets.

Definition 4.1 A topological space $(X, \mathcal{T})$ is said to be normal if for any pair of disjoint closed sets A and B there exist disjoint open sets U and V such that $A \subseteq U$ and $B \subseteq V$.

In other words, a topological space is normal if and only if disjoint closed sets can be separated by disjoint open sets. If we extend the use of the word neighbourhood by saying that N is a neighbourhood of a set C if there is some open set G such that $C \subseteq G \subseteq N$, then $(X, \mathcal{T})$ is normal if and only if every pair of disjoint closed sets have disjoint neighbourhoods.

If $(X, \mathcal{T})$ is normal and if every one-point set is closed, then $(X, \mathcal{T})$ is Hausdorff. For this reason, the condition that one-point sets be closed is often taken as part of the definition of a normal topological space.

Example 4.2 Let $X = \{0, 1, 2\}$ with $\mathcal{T} = \big\{\varnothing, X, \{0\}, \{1, 2\}\big\}$. Then $(X, \mathcal{T})$ is normal—every closed set is also open. However, $(X, \mathcal{T})$ is not Hausdorff. In this case, the one-point sets $\{1\}$ and $\{2\}$ are not closed.

Let $B(x; r)$ denote the 'open' ball $B(x; r) = \{x' \in X : d(x, x') < r\}$ in the metric space (X, d).

Proposition 4.3 *Every metrizable topological space is normal.*

Proof Suppose that A and B are non-empty disjoint closed sets in the metrizable space $(X, \mathcal{T})$. Then $X \setminus A$ and $X \setminus B$ are open sets with $B \subseteq X \setminus A$ and $A \subseteq X \setminus B$. Hence, for each $a \in A$, there is some $\varepsilon_a > 0$ such that $B(a; \varepsilon_a) \subseteq X \setminus B$, and similarly, for each $b \in B$, there is some $\varepsilon_b > 0$ such that $B(b; \varepsilon_b) \subseteq X \setminus A$. It follows that $d(a, b) \geq \max\{\varepsilon_a, \varepsilon_b\}$, for any $a \in A$, $b \in B$. For $a \in A$, set $U_a = B(a; \frac{1}{2}\varepsilon_a)$, and for $b \in B$, set $V_b = B(b; \frac{1}{2}\varepsilon_b)$. Put $U = \bigcup_{a \in A} U_a$ and $V = \bigcup_{b \in B} V_b$. Then U and V are both open and $A \subseteq U$ and $B \subseteq V$.

We claim that U and V are disjoint. Indeed, if $x \in U \cap V$, then $x \in U_a \cap V_b$ for some $a \in A$ and $b \in B$. This means that $d(x,a) < \frac{1}{2}\varepsilon_a$ and $d(x,b) < \frac{1}{2}\varepsilon_b$. Hence

$$\begin{aligned} d(a,b) &\le d(x,a) + d(x,b) \\ &< \tfrac{1}{2}(\varepsilon_a + \varepsilon_b) \\ &\le \max\{\varepsilon_a, \varepsilon_b\} \end{aligned}$$

which is a contradiction. We conclude that $U \cap V = \varnothing$, as required. ■

We can give an alternative proof using the continuity of the distance between a point and a set in a metric space. We recall that if x is any point in a metric space (X,d) and A is any non-empty subset of X, the distance between x and A is given by

$$\operatorname{dist}(x,A) = \inf\{d(x,a) : a \in A\}.$$

One shows that $x \mapsto \operatorname{dist}(x,A)$ is a continuous mapping from X into $\mathbb{R}$ and that $x \in \overline{A}$ if and only if $\operatorname{dist}(x,A) = 0$. Now suppose that A and B are disjoint closed non-empty sets in X. Put $U = \{x \in X : \operatorname{dist}(x,B) - \operatorname{dist}(x,A) > 0\}$ and $V = \{x \in X : \operatorname{dist}(x,B) - \operatorname{dist}(x,A) < 0\}$. Then U is the inverse image of the open set $\{t \in \mathbb{R} : t > 0\}$ under the continuous map $x \mapsto \operatorname{dist}(x,B) - \operatorname{dist}(x,A)$ and so is open in X. Similarly, V is open, being the inverse image of the open set $\{t \in \mathbb{R} : t < 0\}$ under the same continuous map. It is clear that $U \cap V = \varnothing$. Now, if $a \in A$, then $\operatorname{dist}(a,A) = 0$. If $\operatorname{dist}(a,B)$ were 0, we would conclude that $a \in \overline{B}$. But B is closed so that $\overline{B} = B$ and we know that $a \notin B$. Hence $\operatorname{dist}(a,B) > 0$ and so $a \in U$. Thus $A \subseteq U$. Similarly, $B \subseteq V$ and the result follows.

Theorem 4.4 *Every compact Hausdorff space is normal.*

Proof Suppose that $(X,\mathcal{T})$ is a compact Hausdorff topological space. We shall first show that if $z \in X$ and F is a closed set not containing z, then z and F have disjoint neighbourhoods.

For any $x \in F$, there are disjoint open sets U_x and V_x such that $z \in U_x$ and $x \in V_x$ (by the Hausdorff property). The sets $\{V_x : x \in F\}$ form an open cover of F. Now, F is a closed set in a compact space and so is compact. Hence there is a finite set J in F such that $F \subseteq \bigcup_{x \in J} V_x$. Put $U = \bigcap_{x \in J} U_x$ and $V = \bigcup_{x \in J} V_x$. Then U and V are both open, $z \in U$, $F \subseteq V$ and for any $x \in J$, $U \subseteq U_x$ so that $U \cap V_x = \varnothing$. Hence $U \cap V = \varnothing$.

Now let A and B be any pair of disjoint non-empty closed sets in X. For each $a \in A$ there are disjoint open sets U_a and V_a such that $a \in U_a$ and $B \subseteq V_a$, by the above argument. The sets $\{U_a : a \in A\}$ form an open cover of the compact set A and so there is a finite set S in A such that $A \subseteq \bigcup_{a \in S} U_a$. Put $U = \bigcup_{a \in S} U_a$ and $V = \bigcap_{a \in S} V_a$. Then U and V are both open and $A \subseteq U$ and $B \subseteq V$. Furthermore, $U_a \cap V \subseteq U_a \cap V_a = \varnothing$ and so U and V are disjoint and the proof is complete. ■

Next we show that normal spaces are characterized by the property that any open neighbourhood of a closed set contains the closure of an open neighbourhood of the closed set.

Theorem 4.5 *The topological space $(X, \mathfrak{T})$ is normal if and only if for any closed set F and open set V with $F \subseteq V$, there is an open set U such that*

$$F \subseteq U \subseteq \overline{U} \subseteq V.$$

Proof For notational convenience, let us denote the complement $X \setminus A$ of any set A by A^c. Suppose that $(X, \mathfrak{T})$ is normal and let F be any closed set and V any open set with $F \subseteq V$. Then $F \cap V^c = \varnothing$. Now, V^c is closed and so there are disjoint open sets U and W such that $F \subseteq U$ and $V^c \subseteq W$. Thus $W^c \subseteq V$. But to say that U and W are disjoint is to say that $U \subseteq W^c$. Since W is open, W^c is closed and therefore $\overline{U} \subseteq W^c$. Piecing all this together, we have

$$F \subseteq U \subseteq \overline{U} \subseteq W^c \subseteq V.$$

In particular, $F \subseteq U \subseteq \overline{U} \subseteq V$.

For the converse, let A and B be any pair of disjoint closed sets. Then B^c is open and $A \subseteq B^c$. By hypothesis, there is an open set U such that $A \subseteq U \subseteq \overline{U} \subseteq B^c$. Put $V = \overline{U}^c$. Then V is open and $\overline{U} \subseteq B^c$ implies that $B \subseteq \overline{U}^c = V$. It is clear that $U \cap V = \varnothing$ and we conclude that $(X, \mathfrak{T})$ is normal. ∎

The next result, Urysohn's lemma, ensures a plentiful supply of continuous functions on a normal space.

Theorem 4.6 (Urysohn's lemma) *For any pair of non-empty disjoint closed sets A and B in a normal space $(X, \mathfrak{T})$ there is a continuous map $f : X \to \mathbb{R}$ with values in $[0, 1]$ such that $f \restriction A = 0$ and $f \restriction B = 1$.*

Proof Suppose that A and B are non-empty disjoint closed sets in the normal space $(X, \mathfrak{T})$. Let $U_1 = X \setminus B$. Then U_1 is open and $A \subseteq U_1$. Hence there is an open set U_0, say, such that

$$A \subseteq U_0 \subseteq \overline{U_0} \subseteq U_1.$$

We shall construct a family $\{U_q : q \in \mathbb{Q}\}$ of open sets, labelled by the rationals, $\mathbb{Q}$, such that $\overline{U_p} \subseteq U_q$ whenever $p < q$.

To do this, we first consider the rational numbers in the interval $[0, 1]$. This is a countable set, so we may list its members in a sequence, and we may begin with the rationals $0, 1$ as the first two terms:

$$\mathbb{Q} \cap [0, 1] = \{p_1 = 0, p_2 = 1, p_3, p_4, \dots\}.$$

We have already set $U_{p_1} = U_0$ and $U_{p_2} = U_1$. We shall construct U_{p_k} recursively. Suppose that we have already constructed open sets $U_{p_1}, \dots, U_{p_n}$ such that $\overline{U_{p_i}} \subseteq U_{p_j}$ whenever $p_i < p_j$, $1 \le i < j \le n$. We now construct $U_{p_{n+1}}$. Let p_i be the largest and p_j the smallest member of the set $\{p_1, \dots, p_n\}$ such that $p_i < p_{n+1} < p_j$. By hypothesis, $\overline{U_{p_i}} \subseteq U_{p_j}$. Let $U_{p_{n+1}}$ be any open set satisfying $\overline{U_{p_i}} \subseteq U_{p_{n+1}} \subseteq \overline{U_{p_{n+1}}} \subseteq U_{p_j}$. Thus, to every rational p in $[0,1]$ we have associated an open set U_p such that $\overline{U_p} \subseteq U_q$ whenever $p < q$. For $p \in \mathbb{Q}$ with $p < 0$, set $U_p = \varnothing$, and for $p \in \mathbb{Q}$ with $p > 1$, set $U_p = X$. Then $\overline{U_p} \subseteq U_q$ whenever $p < q$, for any $p, q \in \mathbb{Q}$.

We use this family of open sets to construct a suitable function on X. To this end, for given $x \in X$, let

$$Q_x = \{p \in \mathbb{Q} : x \in U_p\}.$$

If $p \in Q_x$, then $x \in U_p$. But $U_p \subseteq U_q$ whenever $p \le q$, so that $x \in U_q$ for any $q \ge p$. That is, $q \in Q_x$, whenever $q \ge p$. Furthermore, if $p < 0$, then $U_p = \varnothing$ so that $x \notin U_p$. It follows that $Q_x \subseteq [0, \infty)$. Now, if $p > 1$, then $U_p = X$, so that $x \in U_p$. Hence Q_x contains all rationals greater than 1. It follows that $\inf Q_x \in [0,1]$.

Define the map $f : X \to \mathbb{R}$ by $f : x \mapsto \inf Q_x$, $x \in X$. Then f takes its values in the interval $[0,1]$. We observe that if $x \in A$, then $x \in U_0$ and so $0 \in Q_x$ and therefore $f(x) = 0$. Next, we see that if $x \in B$, then $x \notin U_1 = X \setminus B$, so that $x \notin U_p$ for any $p \le 1$. Hence $Q_x = \{p \in \mathbb{Q} : p > 1\}$ and so $f(x) = 1$. Thus $f \restriction A = 0$ and $f \restriction B = 1$, and all that remains is to prove that f is continuous.

It is sufficient to show that $f^{-1}(G)$ is open in $(X, \mathcal{T})$ for any open set G of the form (a,b) with $a < b$ since such sets form a base for the topology on $\mathbb{R}$. We shall show that $f^{-1}((-\infty, a))$ is open. If $a \le 0$, then $f^{-1}((-\infty, a)) = \varnothing$, and if $a \ge 1$, $f^{-1}((-\infty, a)) = X$, so we need only consider $0 < a < 1$. We have $f^{-1}((-\infty, a)) = \{x \in X : f(x) < a\}$. We claim that this set is equal to $\bigcup_{p<a} U_p$. Indeed, suppose that x is such that $f(x) < a$. Then $\inf Q_x < a$ and so there is some $p \in Q_x$ such that $p < a$. Thus $x \in U_p$. On the other hand, if $p \in \mathbb{Q}$ with $p < a$ and if $x \in U_p$, then $p \in Q_x$ so that $\inf Q_x \le p < a$. Hence

$$\{x \in X : f(x) < a\} = \bigcup_{p<a} U_p$$

as claimed. It follows that $\{x \in X : f(x) < a\}$ is an open set.

Now consider $f^{-1}((b, \infty)) = \{x \in X : f(x) > b\}$. If $b \le 0$ then this set is equal to X, and if $b \ge 1$ it is empty, so let $0 < b < 1$. We claim that this set is equal to $\bigcup_{q>b} X \setminus \overline{U_q}$. To see this, suppose that $f(x) > b$. Then $\inf Q_x > b$ and so there is some rational $p > b$ such that $p \notin Q_x$, that is, $x \notin U_p$. Now let q be any rational such that $b < q < p$. Then $\overline{U_q} \subseteq U_p$ and so $x \notin \overline{U_q}$. Thus, $x \in X \setminus \overline{U_q}$.

Conversely, suppose that q is a rational with $q > b$ and $x \in X \setminus \overline{U_q}$. Then $x \notin \overline{U_q}$ and so certainly $x \notin U_q$. Hence $x \notin U_p$ for any $p \le q$ and we deduce that Q_x contains no rationals less than q. It follows that $\inf Q_x \ge q > b$, and therefore $f(x) > b$. Thus we have shown that

$$\{x \in X : f(x) > b\} = \bigcup_{q>b} X \setminus \overline{U_q}$$

which is an open set in X. It follows that $f^{-1}((a,b)) = f^{-1}((-\infty, a)) \cap f^{-1}((b, \infty))$ is an open set in X and we conclude that f is continuous and the proof is complete. ∎

Remark 4.7 The values 0 and 1 in the theorem are not critical. Indeed, if f is as in the statement of the theorem and if $\alpha < \beta$ is any pair of real numbers, put $g = \alpha + (\beta - \alpha) f$. Then g is a continuous map from X into $\mathbb{R}$ with values in the interval $[\alpha, \beta]$ such that $g \restriction A = \alpha$ and $g \restriction B = \beta$.

Note also that the theorem makes no claim as to the value of f outside the sets A and B, other than it lies in $[0,1]$. It is quite possible for f to assume either of the values 0 or 1 outside A or B. The next result sheds some light on this. A subset of a topological space is said to be a G_δ set if it is equal to a countable intersection of open sets.

Theorem 4.8 *Let A and B be non-empty closed disjoint subsets of a normal space $(X, \mathcal{T})$. There is a continuous map $f : X \to \mathbb{R}$ with values in $[0,1]$ such that*

(i) $A \subseteq \{x \in X : f(x) = 0\}$ *and*

(ii) $B \subseteq \{x \in X : f(x) = 1\}$

with equality in (i) if and only if A is a G_δ set, and equality in (ii) if and only if B is a G_δ set.

Proof If $f : X \to \mathbb{R}$ is continuous with values in $[0,1]$, then the closed set $\{x \in X : f(x) = 0\}$ can be written as

$$\{x \in X : f(x) = 0\} = \bigcap_{n \in \mathbb{N}} \{x \in X : f(x) < \frac{1}{n}\}$$

which is evidently a G_δ set since each term on the right hand side is an open set in X. Similarly,

$$\{x \in X : f(x) = 1\} = \bigcap_{n \in \mathbb{N}} \{x \in X : f(x) > 1 - \frac{1}{n}\}$$

is a G_δ set. Thus equality in (i) or (ii) demands that A or B be a G_δ set, respectively.

Suppose now that A and B are disjoint closed sets and that A is a G_δ set, say, $A = \bigcap_n G_n$, G_n open in X. By replacing $\{G_n\}_{n\in\mathbb{N}}$ by the family $\{G_1 \cap \cdots \cap G_n \cap (X \setminus B)\}_{n\in\mathbb{N}}$, we may assume that $G_{n+1} \subseteq G_n$ and that $G_n \cap B = \varnothing$ for $n \in \mathbb{N}$.

For each $n \in \mathbb{N}$, let $f_n : X \to \mathbb{R}$ be a continuous map with values in $[0,1]$ such that $f_n \restriction A = 0$ and $f_n \restriction X \setminus G_n = 1$. Such f_n exist by Urysohn's lemma. Put $f = \sum_{n=1}^\infty 2^{-n} f_n$. Then one checks that f is continuous on X with values in $[0,1]$ and that $f \restriction B = 1$ (because $\sum_{n=1}^\infty 2^{-n} = 1$). We wish to show that $A = \{x \in X : f(x) = 0\}$. Certainly f vanishes on A since each f_n does. Suppose that $f(x) = 0$. Then $f_n(x) = 0$ for each $n \in \mathbb{N}$. This means that $x \in G_n$ for every n and so $x \in \bigcap_{n\in\mathbb{N}} G_n = A$. Hence $A = \{x \in X : f(x) = 0\}$.

Suppose now that B is a G_δ set. Construct a continuous map $f : X \to \mathbb{R}$ as above but with $B = \{x \in X : f(x) = 0\}$ and $f \restriction A = 1$. Set $g = 1 - f$. Then $g : X \to \mathbb{R}$ is continuous, has values in $[0,1]$ and $g \restriction A = 0$ and $B = \{x \in X : g(x) = 1\}$.

Now suppose that both A and B are G_δ sets. Let f and g be as above and set $h = \frac{1}{2}(f + g)$. Then $h : X \to \mathbb{R}$ is continuous, takes values in $[0,1]$ and $A = \{x \in X : h(x) = 0\}$ and $B = \{x \in X : h(x) = 1\}$. ■

The next result we shall discuss is the Tietze extension theorem which asserts that a partially defined continuous map on a normal space can be extended to the whole space without spoiling its continuity. We first recall that if $(X, \mathcal{T})$ is a topological space then $C_b(X, \mathbb{R})$, the linear space of continuous real-valued maps on X, equipped with the norm $\|f\|_\infty = \sup\{|f(x)| : x \in X\}$, is a complete normed space. Indeed, $C_b(X, \mathbb{R})$ is a closed subspace of $\ell^\infty(X, \mathbb{R})$, the Banach space of all bounded real-valued maps on X, equipped with the sup-norm.

Theorem 4.9 (Tietze extension theorem) *Suppose that A is a closed subset of a normal topological space $(X, \mathcal{T})$.*

(i) *Any continuous map f from A into $\mathbb{R}$ with values in the interval $[a,b]$ can be extended to a continuous map of X into $\mathbb{R}$ also with values in $[a,b]$; that is, there is a continuous map g from X into $\mathbb{R}$ such that $g(x) \in [a,b]$ for each $x \in X$ and such that $g(x) = f(x)$ for $x \in A$.*

(ii) *Any continuous map from A into $\mathbb{R}$ may be extended to a continuous map from X into $\mathbb{R}$.*

Proof (i) Without loss of generality, we may suppose that $[a,b]$ is the interval $[-1,1]$. Suppose, then, that $f : A \to \mathbb{R}$ is continuous and that f has values in the interval $[-1,1]$. Let $A_0 = \{x : f(x) \le \frac{1}{3}\}$ and $B_0 = \{x : f(x) \ge \frac{1}{3}\}$. Evidently, $A_0 \cap B_0 = \varnothing$, and the continuity of f implies that A_0 and B_0 are closed sets in

A. Since A is closed in X, both A_0 and B_0 are closed in X. By Urysohn's lemma, there is a continuous function $g_0 : X \to \mathbb{R}$ with values in the interval $[-\frac{1}{3}, \frac{1}{3}]$ such that $g_0(x) = -\frac{1}{3}$ for $x \in A_0$ and $g_0(x) = \frac{1}{3}$ for $x \in B_0$. Then $|(f - g_0)(x)| \le \frac{2}{3}$ for each $x \in A$, so that $f - g_0$ is a continuous real-valued function on A with values in the interval $[-\frac{2}{3}, \frac{2}{3}]$. Set $f_1 = \frac{3}{2}(f - g_0)$. Then f_1 is a continuous map from A into $\mathbb{R}$ with values in $[-1, 1]$.

Set $f_0 = f$, and suppose that $n \ge 0$ and that $f_n : A \to \mathbb{R}$ has been constructed such that f_n is continuous and takes its values in the interval $[-1, 1]$. The above argument can be applied to f_n instead of f to yield a continuous map $g_n : X \to \mathbb{R}$ with values in $[-\frac{1}{3}, \frac{1}{3}]$ such that $f_n - g_n$ has values in $[-\frac{2}{3}, \frac{2}{3}]$. We then put $f_{n+1} = \frac{3}{2}(f_n - g_n)$, so that $f_{n+1} : A \to \mathbb{R}$ is continuous and has values in $[-1, 1]$. Thus f_n and g_n are defined recursively for all $n \ge 0$.

By construction,

$$\begin{aligned} f = f_0 &= g_0 + \tfrac{2}{3} f_1 \\ &= g_0 + \tfrac{2}{3} g_1 + \left(\tfrac{2}{3}\right)^2 f_2 \\ &= g_0 + \tfrac{2}{3} g_1 + \left(\tfrac{2}{3}\right)^2 g_2 + \left(\tfrac{2}{3}\right)^3 f_3 \\ &= g_0 + \tfrac{2}{3} g_1 + \cdots + \left(\tfrac{2}{3}\right)^n g_n + \left(\tfrac{2}{3}\right)^{n+1} f_{n+1}. \end{aligned}$$

Put $s_n = g_0 + \frac{2}{3} g_1 + \cdots + \left(\frac{2}{3}\right)^n g_n$ on X. Then it is clear that (s_n) is a Cauchy sequence in $C_b(X, \mathbb{R})$ so that there is g such that $s_n \to g$ in $C_b(X, \mathbb{R})$. Since $|g_i(x)| \le \frac{1}{3}$ for each $i \ge 0$, we see that $s_n(x) \le 1$ and so it follows that $|g(x)| \le 1$ for all $x \in X$. But we also have that, for any $x \in A$, $\left(\frac{2}{3}\right)^{n+1} f_{n+1}(x) \to 0$ in $\mathbb{R}$ as $n \to \infty$. It follows that, for $x \in A$,

$$\begin{aligned} f(x) &= s_n(x) + \left(\tfrac{2}{3}\right)^{n+1} f_{n+1}(x) \\ &\to g(x) + 0 \end{aligned}$$

and so $g = f$ on A, and $g : X \to \mathbb{R}$ is a continuous extension of f, as required.

(ii) Suppose now that $f : A \to \mathbb{R}$ is continuous and that f takes values in the open interval $(-1, 1)$. We shall show that f can be extended to a continuous function with values also in $(-1, 1)$. Certainly, f takes its values in $[-1, 1]$ and so, as above, there is a continuous map $g : X \to \mathbb{R}$ with values in $[-1, 1]$ such that $g = f$ on A. We must modify g to remove the possibility of the values ± 1. Set $D = g^{-1}(\{-1, 1\})$. Then D is a closed subset of X, since g is continuous. Furthermore, on A, g agrees with f which does not assume either of the values 1 or -1, so $A \cap D = \varnothing$. By Urysohn's lemma, there is a continuous map $\varphi : X \to \mathbb{R}$ such that φ takes its values in $[0, 1]$ and such that φ vanishes on D and is equal to 1 on A. Put $h(x) = \varphi(x) g(x)$ for $x \in X$. Then $h : X \to \mathbb{R}$ is continuous, since it is the product of continuous maps. Moreover, for any $x \in A$, $h(x) = \varphi(x) g(x) =$

$1\,g(x) = f(x)$. Finally, we note that for $x \in D$, $h(x) = \varphi(x)g(x) = 0$, and, for $x \notin D$, $|h(x)| = |\varphi(x)|\,|g(x)| < 1$ since $|\varphi(x)| \leq 1$, for all $x \in X$, and $|g(x)| < 1$ for $x \notin D$. Hence h has values in the interval $(-1, 1)$.

For the general case, $f : A \to \mathbb{R}$, put $F(x) = \psi(f(x))$, for $x \in A$, where $\psi : \mathbb{R} \to (-1, 1)$ is the map $x \mapsto \psi(x) = x/(1 + |x|)$. Then ψ is a continuous map with a continuous inverse. By the above, there is a continuous map $G : X \to \mathbb{R}$ with $G \restriction A = F$. Setting $g(x) = \psi^{-1}(G(x))$, for $x \in X$, gives a continuous extension of f as required. ■

5. Vector Spaces

We shall collect together here some of the basic algebraic results we will need concerning vector spaces. We wish to consider vector spaces either over $\mathbb{R}$, the field of real numbers, or over $\mathbb{C}$, the complex numbers, so for notational convenience, we use the symbol $\mathbb{K}$ to stand for either $\mathbb{R}$ or $\mathbb{C}$. It is often immaterial which of these fields of scalars is used, but when it is important we will indicate explicitly which is meant.

Definition 5.1 A finite set of elements $x_1, \dots, x_n$ in a vector space X over $\mathbb{K}$ is said to be linearly independent if and only if

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

with $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ implies that $\alpha_1 = \cdots = \alpha_n = 0$. 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

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

for some $m \in \mathbb{N}$, non-zero $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ 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 $x = \alpha_1 u_1 + \cdots + \alpha_m u_m$, as above, then this representation is unique. To see this, suppose that we also have that $x = \beta_1 v_1 + \cdots + \beta_k v_k$, 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 get

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

Suppose that $m \le k$. 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, $\beta_1, \dots, \beta_m$ is the same permutation of $\alpha_1, \dots, \alpha_m$ and so the representation of x as a finite linear combination of elements of A (with non-zero coefficients) is unique.

In order to establish the existence of a Hamel basis, we shall need to use Zorn's lemma, which shall discuss next. First we need the concept of maximal member in a partially ordered set.

Definition 5.3 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 5.4 Let P be the half-plane in $\mathbb{R}^2$ given by $P = \{(x, y) : x + y \leq 0\}$, equipped with the partial ordering $(a, b) \preceq (c, d)$ if and only if both $a \leq c$ and $b \leq d$ in $\mathbb{R}$. 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 5.5 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 subset C of a partially ordered set $(P, \preceq)$ is said to be a totally ordered subset (or a linearly ordered subset or a chain) in P if, for any pair $c', c'' \in C$, either $c' \preceq c''$ or $c'' \preceq c'$. In other words, C is totally ordered if any two elements in C are comparable.

We now have sufficient terminology to state Zorn's lemma which we shall accept without proof.

Zorn's lemma Let P be a non-empty partially ordered set. If every chain in P has an upper bound, then P possesses at least one maximal element.

Remark 5.6 As stated, the underlying intuition 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 equivalent to Hausdorff's maximality principle and the axiom of choice. These are the following.

Hausdorff's maximality principle Any non-empty partially ordered set contains a maximal chain, i.e., a totally ordered subset maximal with respect to being totally ordered.

Axiom of Choice Let $\{A_\alpha : \alpha \in J\}$ be a family of (pairwise disjoint) non-empty sets, indexed by the non-empty set J. Then there is a mapping $\varphi : J \to \bigcup_\alpha A_\alpha$ 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. (The requirement that the $\{A_\alpha\}$ be pairwise disjoint is not essential and can easily be removed—by replacing A_α by $B_\alpha = \{(\alpha, a) : a \in A_\alpha\}$.) As a consequence, this axiom gives substance to the cartesian product $\prod_\alpha A_\alpha$.

We are now in a position to attack the existence problem of a Hamel basis.

Theorem 5.7 *Every vector space X ($\neq \{0\}$) 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 $\lambda_1, \dots, \lambda_m \in \mathbb{K}$. 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 $\lambda_1, \dots, \lambda_k \in \mathbb{K}$. Then, for any distinct $u_1, \dots, u_k \in M$, an equality of the form

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

must entail $\alpha = 0$. This means that $\lambda_1 = \cdots = \lambda_k = 0$, by independence, and so $x, u_1, \dots, u_k$ 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

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

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

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

Theorem 5.8 *Let A be a linearly independent subset of a vector space X. Then there is a Hamel basis of X containing A; that is, any linearly independent subset of a vector 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. ■

Now we turn to the relationship between subspaces of a vector space and linear functionals. First we need a definition.

Definition 5.9 A subspace V in a vector space X is said to be a maximal proper subspace if it is not contained in any other proper subspace.

Proposition 5.10 *Let V be a linear subspace of a vector space X. The following statements are equivalent.*

(i) *V is a maximal proper subspace of X.*

(ii) *There is some $z \in X$, with $z \neq 0$, such that $X = \{v + tz : v \in V, t \in \mathbb{K}\}$.*

(iii) *V has codimension one, i.e., X/V has dimension one.*

Proof Suppose that V is a maximal proper subspace of X. Then there is some $z \in X$ with $z \notin V$. The set $\{v + tz : v \in V, t \in \mathbb{K}\}$ is a linear subspace of X containing V as a proper subset. By the maximality of V, this set must be the whole of X, thus, (ii) follows from (i).

Suppose that (ii) holds and let $[x]$ be any member of X/V, i.e., $[x]$ is the equivalence class in X/V which contains the element x. Then, writing x as $x = v + tz$, for some $v \in V$ and $t \in \mathbb{K}$, we see that x is equivalent to tz, i.e., $[x] = [tz] = t[z]$. It follows that X/V is one-dimensional (with basis element given by $[z]$), which proves (iii).

Finally, suppose that X/V is one-dimensional. For any $x \in X$, we have $[x] = t[z]$ in X/V, for some $t \in \mathbb{K}$ and some $z \notin V$. Hence $x = v + tz$ for some $v \in V$. In particular, any linear subspace containing V as a proper subset must contain z. The only such subspace is X itself, and we deduce that V is a maximal proper subspace. ■

Definition 5.11 Let X and Y be vector spaces over $\mathbb{K}$, i.e., both over $\mathbb{R}$ or both over $\mathbb{C}$. A map $T : X \to Y$ is said to be a linear mapping if

$$T(\alpha x' + x'') = \alpha\, Tx' + Tx'', \qquad \text{for } \alpha \in \mathbb{K} \text{ and } x', x'' \in X.$$

A map $T : X \to Y$, with $\mathbb{K} = \mathbb{C}$, is called conjugate linear if

$$T(\alpha x' + x'') = \overline{\alpha}\, Tx' + Tx'', \qquad \text{for } \alpha \in \mathbb{C} \text{ and } x', x'' \in X.$$

A linear map $\lambda : X \to \mathbb{K}$ is called a linear functional or linear form. If $\mathbb{K} = \mathbb{R}$, then λ is called a real linear functional, whereas if $\mathbb{K} = \mathbb{C}$, λ is called a complex linear functional. The algebraic dual of X (often denoted X') is the vector space (over $\mathbb{K}$) of all linear functionals on X equipped with the obvious operations of addition and scalar multiplication.

Example 5.12 Let $X = \mathbb{C}^n$. For any given $u = (u_1, \dots, u_n) \in \mathbb{C}^n$, the map $z = (z_1, \dots, z_n) \mapsto \sum_{i=1}^n u_i z_i$ is a (complex) linear functional, whereas the map $z \mapsto \sum_{i=1}^n u_i \overline{z_i}$ is a conjugate linear functional. In fact, every linear (respectively, conjugate linear) functional on $\mathbb{C}^n$ has this form. We can see this by induction. Indeed, if $\lambda : \mathbb{C} \to \mathbb{C}$ is a linear functional on $\mathbb{C}$, then $\lambda(z_1) = u_1 z_1$ for any $z_1 \in \mathbb{C}$, where $u_1 = \lambda(1)$, and so the statement is true for $n = 1$.

Suppose now that is true for $n = k$ and let $\lambda : \mathbb{C}^{k+1} \to \mathbb{C}$ be a linear functional. Let e_i, $i = 1, \dots, k+1$ be the standard basis vectors for $\mathbb{C}^{k+1}$, that is, $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$ etc. The map $(z_1, \dots, z_k) \mapsto \lambda((z_1, \dots, z_k, 0))$ is a linear functional on $\mathbb{C}^k$ and so, by the induction hypothesis, there are elements $u_1, \dots, u_k \in \mathbb{C}$ such that

$$\lambda((z_1, \dots, z_k, 0)) = \sum_{i=1}^k u_i z_i.$$

Hence, for any $(z_1, \dots, z_{k+1}) \in \mathbb{C}^{k+1}$, we have

$$\begin{aligned} \lambda((z_1, \dots, z_{k+1})) &= \lambda((z_1, \dots, z_k, 0)) + \lambda((0, \dots, 0, z_{k+1})) \\ &= \sum_{i=1}^k u_i z_i + z_{k+1} \lambda((0, \dots, 0, 1)) \\ &= \sum_{i=1}^{k+1} u_i z_i \end{aligned}$$

where $u_{k+1} = \lambda((0, \dots, 0, 1))$ and the result follows.

The case of a conjugate linear functional can be proved similarly. Alternatively, it can be deduced from the linear case by noticing that if $\ell : \mathbb{C}^n \to \mathbb{C}$ is conjugate linear, then the map $z \mapsto \lambda(z) = \overline{\ell(z)}$ is linear and so has the above form, $\lambda(z) = \sum_{i=1}^n u_i z_i$. But then $\ell(z) = \overline{\lambda(z)}$ is of the form $\ell(z) = \sum_{i=1}^n v_i \overline{z_i}$ with $v_i = \overline{u_i}$, $1 \le i \le n$, as required.

Suppose that X is a vector space and that $z \in X$, with $z \neq 0$. Let B be a Hamel basis for X containing z. Any non-zero element $x \in X$ can be written uniquely as $x = \alpha_1 x_1 + \dots \alpha_n x_n$ for elements $x_1, \dots, x_n$ in B and non-zero $\alpha_1, \dots, \alpha_n$ in $\mathbb{K}$. It follows that the assignment $\lambda_z(z) = 1$ and $\lambda_z(x) = 0$ for all $x \in B$, $x \neq z$, defines an element, λ_z, of the algebraic dual of X. In particular, if $x \neq y$ are any two elements of X, then putting $z = x - y$ we see that there is an element λ in the algebraic dual of X such that $\lambda(x) \neq \lambda(y)$. We shall see later that when X is furnished with a vector space topology, then, under suitable circumstances, λ may be chosen to be continuous. We note also that if $x_1, \dots, x_n$ are linearly independent elements of X, then the linear functionals $\lambda_{x_1}, \dots, \lambda_{x_n}$ are linearly independent in X'. In particular, it easy to see that if X is finite dimensional then so is X', both having the same dimension, and that if X is infinite dimensional, so is X'.

A complex vector space can be regarded as a real vector space by simply restricting the field of scalars to be $\mathbb{R}$. This permits a natural correspondence between real and complex linear functionals on a complex vector space as we now show. Let X be a complex vector space, and $\lambda : X \to \mathbb{C}$ a (complex) linear functional on X. Define $\ell : X \to \mathbb{R}$ 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 vector space. Substituting ix for x, one readily checks that

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

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

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

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

Definition 5.13 The kernel of the linear functional $\ell : X \to \mathbb{K}$ is the set $\ker \ell = \{x \in X : \ell(x) = 0\}$, the null space of ℓ.

Proposition 5.14 *Suppose that ℓ is a linear functional on the vector space X such that ℓ does not vanish on the whole of X. Then $\ker \ell$ is a maximal proper subspace. Conversely, any maximal proper subspace of X is the kernel of a linear functional.*

Proof It is clear that the kernel of any linear functional ℓ is a linear subspace of X. Moreover, $\ker \ell$ is a proper subspace if and only if ℓ is not identically zero. Suppose that $\ell(z) \neq 0$ for some $z \in X$. For any $x \in X$, we have

$$x = \frac{\ell(x)}{\ell(z)}\, z + \Big(x - \frac{\ell(x)}{\ell(z)}\, z \Big).$$

But $x - (\ell(x)/\ell(z))\, z \in \ker \ell$, so it follows that $\ker \ell$ is maximal.

Now suppose that V is a maximal proper subspace of X, and let z be any element of X such that $z \notin V$. Since V is maximal, any $x \in X$ can be written as $x = \alpha z + v$, for suitable $\alpha \in \mathbb{K}$ and $v \in V$. Moreover, this decomposition of x is unique; if also $x = \alpha' z + v'$ for some $\alpha' \in \mathbb{K}$ and $v' \in V$, then $0 = (\alpha - \alpha')z + (v - v')$. Since $z \notin V$, we must have $\alpha = \alpha'$ and therefore $v = v'$. Define $\ell : X \to \mathbb{K}$ by $\ell(x) = \alpha$, where $x = \alpha z + v$, as above, Evidently, ℓ is well-defined and is linear. If $x = \alpha z + v$ and $\ell(x) = 0$, we have $\alpha = 0$ and so $x = v \in V$. Clearly, $\ell(v) = 0$ for $v \in V$ and so $\ker \ell = V$. ■

Proposition 5.15 *Let ℓ_1 and ℓ_2 be linear functionals on the vector space X, neither being identically zero. Then ℓ_1 and ℓ_2 have the same kernel if and only if they are proportional.*

Proof Suppose that $\ker \ell_1 = \ker \ell_2$. Let $z \in X$ with $\ell_1(z) \neq 0$. For any $x \in X$, $x - (\ell_1(x)/\ell_1(z))\, z \in \ker \ell_1 = \ker \ell_2$ so that $\ell_2(x) = \ell_1(x)\ell_2(z)/\ell_1(z)$. Thus ℓ_1 and ℓ_2 are proportional.

Clearly, if ℓ_1 and ℓ_2 are proportional, then they have the same kernel since the constant of proportionality is not zero, by hypothesis. ■

Proposition 5.16 *Let ℓ and $\ell_1, \dots \ell_n$ be linear functionals on the vector space X. Then either*

(i) *ℓ is a linear combination of $\ell_1, \dots \ell_n$, or*

(ii) *there is $z \in X$ such that $\ell(z) = 1$ and $\ell_1(z) = \cdots = \ell_n(z) = 0$.*

Proof By considering a subset of the ℓ_i if necessary, we may assume, without loss of generality, that $\ell_1, \dots \ell_n$ are linearly independent. Define the map $\gamma : X \to \mathbb{K}^{n+1}$ by

$$\gamma(x) = (\ell(x), \ell_1(x), \dots, \ell_n(x)), \quad x \in X.$$

It is clear that γ is linear and so $\gamma(X)$, the range of γ, is a linear subspace of $\mathbb{K}^{n+1}$. Suppose that $\gamma(x) = \mathbb{K}^{n+1}$. Then, in particular, there is $z \in X$ such that $\gamma(z) = (1, 0, 0, \dots, 0)$, that is, $\ell(z) = 1$, $\ell_1(z) = \cdots = \ell_n(z) = 0$ and so (ii) holds.

If $\gamma(X) \neq \mathbb{K}^{n+1}$, there is some $a = (a_0, a_1, \dots, a_n) \in \mathbb{K}^{n+1}$ orthogonal to $\gamma(X)$; that is,

$$a_0\ell(x) + a_1\ell_1(x) + \cdots + a_n\ell_n(x) = 0$$

for all $x \in X$. (For example, let $\eta_1, \dots \eta_m$ be a basis for $\gamma(X)$ and let η be any element of $\mathbb{K}^{n+1}$ not in $\gamma(X)$. Applying the Gram-Schmidt orthogonalisation procedure will yield a suitable element a orthogonal to $\gamma(X)$ in $\mathbb{K}^{n+1}$.) By hypothesis, $\ell_1, \dots, \ell_n$ are linearly independent and so a_0 cannot be zero. Putting $b_i = -a_i/a_0 \in \mathbb{K}$, $1 \leq i \leq n$, gives

$$\ell = b_1\ell_1 + \cdots + b_n\ell_n$$

which is (i). ■

Corollary 5.17 *Suppose that $\ell, \ell_1, \dots, \ell_n$ are linear functionals on the vector space X such that*

$$\bigcap_{i=1}^{n} \ker \ell_i \subseteq \ker \ell.$$

Then ℓ is a linear combination of $\ell_1, \dots, \ell_n$.

Proof If $\ell_i(x) = 0$ for all $1 \leq i \leq n$, then $\ell(x) = 0$, so that (ii) of the proposition is impossible. Hence (i) holds. ■

Corollary 5.18 *Suppose that $\ell_1, \dots, \ell_n$ are linearly independent linear functionals on the vector space X. There exist elements $z_1, \dots, z_n \in X$ such that $\ell_i(z_j) = \delta_{ij}$, $1 \leq i, j \leq n$.*

Proof ℓ_1 is not a linear combination of $\ell_2, \dots, \ell_n$, by hypothesis, and so (ii) of the proposition holds; there is $z_1 \in X$ such that $\ell_1(z_1) = 1$, $\ell_2(z_1) = \cdots = \ell_n(z_1) = 0$. Repeating this argument for each ℓ_i in turn gives the desired conclusion. ■

Next we consider the first of several versions of the Hahn-Banach theorem. This one is for real vector spaces and uses the order structure in $\mathbb{R}$. Some terminology is needed.

Definition 5.19 We say that a map $p : X \to \mathbb{R}$ on a real vector space X is subadditive if

$$p(x + y) \leq p(x) + p(y), \quad \text{for } x, y \in X$$

and that p is positively homogeneous if

$$p(tx) = t\,p(x), \quad \text{for any } x \in X \text{ and } t > 0.$$

Note that if p is positively homogeneous, then, with $x = 0$ and $t = 2$, we see that $p(0) = 2p(0)$ so that $p(0) = 0$. Hence $p(tx) = tp(x)$ for all $t \geq 0$. If p is also subadditive, then, setting $y = -x$, we get $0 = p(0) \leq p(x) + p(-x)$, that is, $-p(-x) \leq p(x)$, for $x \in X$.

Theorem 5.20 (Hahn-Banach) *Let X be a real vector space and suppose that $f : M \to \mathbb{R}$ is a linear mapping defined on a linear subspace M of X such that $f(y) \leq p(y)$, for all $y \in M$, for some subadditive and positively homogeneous map $p : X \to \mathbb{R}$. Then there is a linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda(x) = f(x)$ for $x \in M$ and*

$$-p(-x) \leq \Lambda(x) \leq p(x) \quad \text{for } x \in X.$$

(In other words, f can be extended to X whilst retaining the same bound.)

Proof We first show how we can extend the definition of f by one extra dimension. If $M \neq X$, let $x_1 \in X$ with $x_1 \notin M$. Define $M_1 = \{x + tx_1 : x \in M, t \in \mathbb{R}\}$. Then M_1 is a linear subspace of X—the subspace of X spanned by $\{x_1\} \cup M$. Any $z \in M_1$ can be written uniquely as $z = x + tx_1$ for $x \in M$ and $t \in \mathbb{R}$. Indeed, if $x + tx_1 = x' + t'x_1$ for $x, x' \in M$ and $t, t' \in \mathbb{R}$, then $0 = (x - x') + (t - t')x_1$. Since $x_1 \notin M$, this implies that $t = t'$ and therefore $x = x'$.

The implications of the existence of a suitable extension to f will provide the idea of how to actually construct such an extension. So suppose for the moment that f_1 is an extension of f to M_1, satisfying the stated bound in terms of p. Let $x + tx_1 \in M_1$, $x \in M$, $t \in \mathbb{R}$. Then $f_1(x + tx_1) = f(x) + tf_1(x_1) = f(x) + t\mu$, where we have set $\mu = f_1(x_1)$. The given bound demands that $f(x) + t\mu \leq p(x + tx_1)$ for all $t \in \mathbb{R}$ and $x \in M$. For $t = 0$, this is nothing other than the hypothesis on f. Replacing x by tx, we get $f(tx) + t\mu \leq p(tx + tx_1)$. Setting $t = 1$ and $t = -1$, we obtain

$$f(x) + \mu \leq p(x + x_1) \quad \text{for } x \in M \text{ (taking } t = 1)$$

and

$$-f(x) - \mu \le p(-x - x_1) \quad \text{for } x \in M, \text{ (taking } t = -1).$$

Replacing x by $-y$ and $y \in M$ in this last inequality we see that μ must satisfy

$$\mu \le p(x + x_1) - f(x) \quad x \in M$$

and

$$f(y) - p(y - x_1) \le \mu \quad y \in M.$$

The idea of the proof is to show that such a μ exists, and then to work backwards to show that f_1, as defined above, does satisfy the boundedness requirement. Let $x, y \in M$. Then

$$\begin{aligned} f(x) + f(y) = f(x + y) &\le p(x + y) \quad \text{by hypothesis} \\ &= p(x - x_1 + y + x_1) \\ &\le p(x - x_1) + p(y + x_1). \end{aligned}$$

Hence

$$f(x) - p(x - x_1) \le p(y + x_1) - f(y), \quad \text{for } x, y \in M.$$

Thus the set $\{f(x) - p(x - x_1) : x \in M\}$ is bounded from above in $\mathbb{R}$. Let μ denote its least upper bound. Then we have

$$f(x) - p(x - x_1) \le \mu \le p(y + x_1) - f(y), \quad \text{for } x, y \in M.$$

Hence

(1) $$f(x) - \mu \le p(x - x_1) \quad \text{for } x \in M$$

and

(2) $$f(y) + \mu \le p(y + x_1) \quad \text{for } y \in M.$$

Define f_1 on M_1 by $f_1(x + tx_1) = f(x) + t\mu$, $x \in M$, $t \in \mathbb{R}$. Then f_1 is linear on M_1 and $f_1 = f$ on M. To verify that f_1 satisfies the required bound, we use the inequalities (1) and (2) and reverse the argument of the preamble. For $t > 0$, replacing x by $t^{-1}x$ in (1) gives

$$f(t^{-1}x) - \mu \le p(t^{-1}x - x_1)$$

and so, multiplying by t and using the linearity of f and the positive homogeneity of p, we get

$$f(x) - t\mu \le p(x - tx_1) \quad \text{for } x \in M \text{ and } t > 0.$$

Thus,

$$f_1(x + sx_1) \le p(x + sx_1) \quad \text{for } x \in M \text{ and any } s < 0.$$

From (2), replacing y by $t^{-1}y$, with $t > 0$, we get

$$f(t^{-1}y) + \mu \le p(t^{-1}y + x_1), \quad y \in M.$$

Thus, for any $t > 0$,

$$f(y) + t\mu \le p(y + tx_1), \quad y \in M,$$

i.e.,

$$f_1(y + tx_1) \le p(y + tx_1), \quad \text{for any } y \in M, \text{ and } t > 0.$$

This inequality also holds for $t = 0$, by hypothesis on f. Combining these inequalities shows that

$$f_1 \le p \quad \text{on } M_1.$$

We have shown that f can be extended from M to a subspace with one extra dimension, whilst retaining the bound in terms of p. To obtain an extension to the whole of X, we shall use Zorn's lemma. (An idea would be to apply the above procedure again and again, but one then has the problem of showing that this would eventually exhaust the whole of X. Use of Zorn's lemma is a way around this difficulty.) Let $\mathcal{E}$ denote the family of ordered pairs (N, h), where N is a linear subspace of X containing M and h is a real linear functional on N such that $h \restriction M = f$ and $h(x) \le p(x)$ for all $x \in N$. The pair (M, f) (and also (M_1, f_1) constructed above) is an element of $\mathcal{E}$, so that $\mathcal{E}$ is not empty. Moreover, $\mathcal{E}$ is partially ordered by extension, that is, we define $(N, h) \preceq (N', h')$ if $N \subseteq N'$ and $h' \restriction N = h$. Let $\mathcal{C}$ be any totally ordered subset of $\mathcal{E}$ and set $N' = \bigcup_{(N,h)\in\mathcal{C}} N$. Let $x \in N'$. Suppose that $x \in N_1 \cap N_2$, where (N_1, h_1) and (N_2, h_2) are members of $\mathcal{C}$. Then either $(N_1, h_1) \succeq (N_2, h_2)$ or $(N_2, h_2) \succeq (N_1, h_1)$. In any event, $h_1(x) = h_2(x)$. Hence we may define h' on N' by the assignment $h'(x) = h(x)$ if $x \in N$ with $(N, h) \in \mathcal{C}$. Then h' is a linear functional on N' such that $h' \restriction M = f$ and $h'(x) \le p(x)$ for all $x \in N'$. This means that (N', h') is an upper bound for $\mathcal{C}$ in $\mathcal{E}$. By Zorn's lemma, $\mathcal{E}$ contains a maximal element (N, Λ), say. If $N \ne X$, then we could construct an extension of (N, Λ), as earlier, which would contradict maximality. We conclude that $N = X$ and that Λ is a linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda \restriction M = f$, and $\Lambda(x) \le p(x)$ for all $x \in X$.

Replacing x by $-x$ gives $\Lambda(-x) \le p(-x)$, i.e., $-\Lambda(x) \le p(-x)$ and so $-p(-x) \le \Lambda(x)$ for $x \in X$. ∎

We can extend this result to the complex case, but of course one has to take the modulus of the various quantities appearing in the inequalities. We also need to know that p behaves well with regard to the extraction of complex scalars from its argument. The appropriate notion turns out to be that of a seminorm—this will play a major rôle in the study of topological vector spaces.

Definition 5.21 Let X be a vector space over $\mathbb{K}$. A mapping $p : X \to \mathbb{R}$ is said to be a seminorm if
(1) $p(x) \ge 0$ for all $x \in X$,
(2) $p(x+y) \le p(x) + p(y)$ for all $x, y \in X$,
(3) $p(\lambda x) = |\lambda|\, p(x)$ for all $x \in X$ and all $\lambda \in \mathbb{K}$.
In words, p is positive, subadditive and absolutely homogeneous. If p has the additional property that $p(x) = 0$ implies that $x = 0$, then p is a norm on X.

The following complex version of the Hahn-Banach theorem will be seen to be a consequence of the real version by exploiting the relationship between a complex linear functional and its real part.

Theorem 5.22 (Hahn-Banach theorem (complex version)) *Suppose that M is a linear subspace of a vector space X over $\mathbb{K}$, p is a seminorm on X, and f is a linear functional on M such that*

$$|f(x)| \le p(x) \quad \text{for } x \in M.$$

Then there is a linear functional Λ on X that satisfies $\Lambda \restriction M = f$ and

$$|\Lambda(x)| \le p(x) \quad \text{for all } x \in X.$$

Proof If $\mathbb{K} = \mathbb{R}$, then, by the previous theorem, there is a suitable Λ with

$$-p(-x) \le \Lambda(x) \le p(x) \quad \text{for } x \in X.$$

However, $p(-x) = p(x)$ since p is a seminorm, and so $|\Lambda(x)| \le p(x)$ for $x \in X$, as required.

Now suppose that $\mathbb{K} = \mathbb{C}$. Consider X as a real vector space, and note that $\operatorname{Re} f$ is a real linear functional on X satisfying $\operatorname{Re} f(x) \le |f(x)| \le p(x)$, for $x \in M$. By the earlier version of the Hahn-Banach theorem, there is a real linear functional $g : X \to \mathbb{R}$ on X such that $g \restriction M = \operatorname{Re} f$ and $g(x) \le p(x)$ for $x \in X$. Set $\Lambda(x) = g(x) - ig(ix)$ for $x \in X$. Then $\Lambda : X \to \mathbb{C}$ is a complex linear functional. Moreover, for any $x \in M$, we have

$$\Lambda(x) = g(x) - ig(ix) = \operatorname{Re} f(x) - i \operatorname{Re} f(ix) = f(x)$$

so that $\Lambda \upharpoonright M = f$. Furthermore, for any given $x \in X$, let $\alpha \in \mathbb{C}$ be such that $|\alpha| = 1$ and $\alpha\Lambda(x) = |\Lambda(x)|$. Then

$$\begin{aligned} |\Lambda(x)| &= \alpha\Lambda(x) = \Lambda(\alpha x) \\ &= \operatorname{Re}\Lambda(\alpha x), \quad \text{since lhs is real,} \\ &= g(\alpha x) \\ &\le p(\alpha x) \\ &= p(x), \quad \text{for } x \in X. \end{aligned}$$

■

For a normed space, with norm $\|\cdot\|$, the corresponding result is as follows.

Corollary 5.23 *Let M be a linear subspace of a normed space X over $\mathbb{K}$ and let f be a linear functional on M such that $|f(x)| \le C\|x\|$ for some constant $C > 0$ and for all $x \in M$. Then there is a linear functional Λ on X such that $\Lambda \upharpoonright M = f$ and $|\Lambda(x)| \le C\|x\|$ for all $x \in X$.*

Proof This follows immediately from the observation that the mapping $x \mapsto p(x) = C\|x\|$ is a seminorm on X. ■

Corollary 5.24 *Suppose that X is a normed space over $\mathbb{K}$ and that $x_0 \in X$. There is a linear functional Λ on X such that $\Lambda(x_0) = \|x_0\|$ and $|\Lambda(x)| \le \|x\|$ for all $x \in X$. In particular, for any pair of distinct points $x \ne y$ in X, there is a bounded linear functional Λ on X such that $\Lambda(x) \ne \Lambda(y)$.*

Proof If $x_0 = 0$, take $\Lambda = 0$ on X. If $x_0 \ne 0$, let M be the one-dimensional linear subspace of X spanned by x_0. Set $p(x) = \|x\|$, for $x \in X$, and define $f : M \to \mathbb{K}$ by $f(\alpha x_0) = \alpha\|x_0\|$, $\alpha \in \mathbb{K}$. Then, for $x = \alpha x_0 \in M$,

$$|f(x)| = |\alpha|\|x_0\| = \|\alpha x_0\| = \|x\|.$$

By the last corollary, there is an extension Λ of f such that $|\Lambda(x)| \le \|x\|$ for all $x \in X$.

Finally, for $x \ne y$, set $x_0 = x - y$ and let Λ be as above. Then $\Lambda(x) - \Lambda(y) = \Lambda(x - y) = \Lambda(x_0) = \|x_0\| = \|x - y\| \ne 0$. ■

6. Topological Vector Spaces

There are many situations in which we encounter a natural linear structure as well as a topological one; $\mathbb{R}^n$ and $C([0,1])$ being obvious examples. The harmonious coalescence of linearity and topological constructs is realized in the concept of topological vector space.

Definition 6.1 A topological vector space over $\mathbb{K}$ is a vector space X over $\mathbb{K}$ furnished with a topology $\mathcal{T}$ such that

(i) the map $(x, y) \mapsto x + y$ is continuous from $X \times X$ into X (where $X \times X$ is given the product topology);

(ii) the map $(t, x) \mapsto tx$ is continuous from $\mathbb{K} \times X$ into X (where $\mathbb{K}$ has its usual topology, and $\mathbb{K} \times X$ the product topology).

One says that $\mathcal{T}$ is a vector topology on the vector space X, or that $\mathcal{T}$ is compatible with the linear structure of X.
We say that a topological vector space $(X, \mathcal{T})$ is separated if the topology $\mathcal{T}$ is a Hausdorff topology.

In words, a topological vector space is a vector space which is at the same time a topological space such that addition and scalar multiplication are continuous. We will see later that, as a consequence of the compatibility between the topological and linear structure, a topological vector space is Hausdorff if and only every one-point set is closed. Sometimes the requirement that the topology be Hausdorff is taken as part of the definition of a topological vector space.

Example 6.2 Any real or complex normed space is a topological vector space when equipped with the topology induced by the norm.
Any vector space is a topological vector space when equipped with the indiscrete topology. Of course, this will fail to be a separated topological vector space unless it is the zero-dimensional space $\{0\}$.

It is convenient to introduce some notation at this point. Basically, we simply wish to extend the vector space notation for addition and scalar multiplication to the obvious thing for subsets. Let X be a vector space over $\mathbb{K}$ and let A and B be subsets of X. We shall use the following notation:

$$A + B = \{x \in X : x = a + b,\ a \in A,\ b \in B\}$$
$$t\,A = \{x \in X : x = ta,\ a \in A\}, \quad \text{for } t \in \mathbb{K}$$
$$A + z = A + \{z\} = \{x \in X : x = a + z,\ a \in A\}, \quad \text{for } z \in X.$$

It should be noted that $(s + t)A \subseteq s\,A + t\,A$ but equality need not hold. For example, if $X = \mathbb{C}$ and $A = \{1, i\}$, then $2A = \{2, 2i\}$ whereas $A+A = \{2, 2i, 1+i\}$.

Remark 6.3 Suppose that X is a topological vector space. Let $\Phi : X \times X \to X$ and $\Psi : \mathbb{K} \times X \to X$ denote the mappings $\Phi((x, y)) = x + y$ and $\Psi((t, x)) = tx$, $x, y \in X$, $t \in \mathbb{K}$. By definition, these maps are continuous. In particular, for any $x, y \in X$, and any neighbourhood W of $\Phi(x, y)$, there are neighbourhoods U of x and V of y such that $\Phi(U \times V) \subseteq W$. That is to say, for any $x, y \in X$, and any neighbourhood W of $x + y$, there are neighbourhoods U of x and V of y such that $U + V \subseteq W$.

Similarly, the continuity of scalar multiplication implies that for given $s \in K$, $x \in X$ and any neighbourhood W of $\Psi(s, x)$, there is a neighbourhood V of s in $\mathbb{K}$ and a neighbourhood U of x in X such that $\Psi(V \times U) \subseteq W$. Now, any such V contains a set of the form $\{t \in \mathbb{K} : |t - s| < \delta\}$, for some $\delta > 0$. Hence we may say that for any neighbourhood W of sx there is a neighbourhood U of x and $\delta > 0$ such that $tU \subseteq W$ for all $t \in \mathbb{K}$ with $|t - s| < \delta$.

Example 6.4 We know that any non-empty set can be equipped with a Hausdorff topology (for example, the discrete topology), so one might ask whether every non-trivial vector space (i.e., not equal to $\{0\}$) can be given a Hausdorff vector topology. The discrete topology is too fine for this. Indeed, each point is a neighbourhood of itself in the discrete topology, and so addition is certainly continuous—the neighbourhood $\{x\} \times \{y\}$ of $(x, y) \in X \times X$ is mapped into the neighbourhood $\{x+y\}$ of $x+y$ under addition. (We are taking $U = \{x\}$, $V = \{y\}$ and $W = \{x+y\}$ in the preceding discussion.) However, scalar multiplication is not continuous—for example, if $W = \{sx_0\}$, a neighbourhood of sx_0, then tx_0 belongs to W only if $t = s$. Thus there is no neighbourhood V of s in $\mathbb{K}$ such that $tx_0 \in W$ for all $t \in V$.

However, every vector space X can be normed. To see this, let B be a Hamel basis for X. Any $x \in X$, with $x \neq 0$, can be written uniquely as

$$x = t_1 u_1 + \cdots + t_n u_n$$

for suitable $u_1, \dots, u_n$ in B and non-zero $t_1, \dots, t_n$ in $\mathbb{K}$. Set

$$\|x\| = |t_1| + \cdots + |t_n|$$

and $\|0\| = 0$. It is clear that $\|\cdot\|$ is, indeed, a norm on X. The topology induced by a norm is Hausdorff so that X furnished with this norm (or, indeed, any other) becomes a topological vector space.

Proposition 6.5 *Let X be a topological vector space. For given $a \in X$ and $s \in \mathbb{K}$, with $s \neq 0$, the translation map $T_a : x \mapsto x + a$ and the multiplication map $M_s : x \mapsto sx$, $x \in X$, are homeomorphisms of X onto itself.*

Proof Let $\Phi : X \times X \to X$ and $\Psi : \mathbb{K} \times X \to X$ be the mappings introduced above. Then, by continuity, if $((x_\alpha, y_\alpha))$ is a net which converges to (x, y) in $X \times X$, we have

$$x_\alpha + y_\alpha = \Phi((x_\alpha, y_\alpha)) \to \Phi((x, y)) = x + y.$$

Similarly, if $(s_\alpha, x_\alpha) \to (s, x)$ in $\mathbb{K} \times X$, then

$$s_\alpha x_\alpha = \Psi((s_\alpha, x_\alpha)) \to \Psi((s, x)) = sx.$$

Hence, for any net (x_α) in X with $x_\alpha \to x$, set $y_\alpha = a$ for all α and $s_\alpha = s$ for all α, where $s \neq 0$. Then $(x_\alpha, y_\alpha) \to (x, a)$ and $(s_\alpha, x_\alpha) \to (s, x)$ and we conclude that

$$T_a(x_\alpha) = x_\alpha + a \to x + a = T_a(x)$$

and

$$M_s(x) = sx_\alpha \to sx = M_s(x).$$

Thus both T_a and M_s are continuous maps from X into X. Furthermore, T_a has inverse T_{-a} and M_s has inverse $M_{s^{-1}}$, which are also continuous by the same reasoning. Therefore T_a and M_s are homeomorphisms of X onto itself. ■

Corollary 6.6 *For any open set G in a topological vector space X, the sets $a + G$, $A + G$ and sG are open, for any $a \in X$, $A \subseteq X$, and $s \in \mathbb{K}$ with $s \neq 0$. In particular, if U is a neighbourhood of 0, then so is tU for any $t \in \mathbb{K}$ with $t \neq 0$.*

Proof Suppose that G is open in X. For any $a \in X$ and $s \in \mathbb{K}$, we have $a + G = T_a(G)$ and $sG = M_s(G)$. By the proposition, T_a and M_s, for $s \neq 0$, are homeomorphisms and so $T_a(G)$ and $M_s(G)$ are open sets. Now we simply observe that $A + G = \bigcup_{a \in A}(a + G)$, which is a union of open sets and so is itself open.

If U is a neighbourhood of 0, there is an open set G with $0 \in G \subseteq U$. Hence $0 \in tG \subseteq tU$ and tG is open so that tU is a neighbourhood of 0. ■

Remark 6.7 The continuity of the map $x \mapsto x + a$ in a topological vector space X has the following fundamental consequence. If V is a neighbourhood of a, then $b - a + V$ is a neighbourhood of b, for any $a, b \in X$. Suppose that $\mathcal{N}_a$ is a local neighbourhood base at a in X. Let $b \in X$, and let G be any neighbourhood of b. Then $T_{a-b}(G) = a-b+G$ is a neighbourhood of a, since T_{a-b} is a homeomorphism, and so there is an element U in $\mathcal{N}_a$ such that $U \subseteq a - b + G$. Hence $b - a + U \subseteq G$ and $b - a + U$ is a neighbourhood of b. It follows that $\{b - a + U : U \in \mathcal{N}_a\}$ is a local neighbourhood base at b. In particular, given any local neighbourhood base $\mathcal{N}_0$ at 0, $\{a + U : U \in \mathcal{N}_0\}$ is a local neighbourhood base at any point $a \in X$. Now, in any topological space, if $\mathcal{N}_x$ is any open neighbourhood base at x, then $\bigcup_{x \in X} \mathcal{N}_x$ is a base for the topology. It follows that in a topological vector space the family of translates $\{x + \mathcal{N}_0 : x \in X\}$ is a base for the topology, for any open local neighbourhood base $\mathcal{N}_0$ at 0. Thus, the topological structure of a topological vector space is determined by any local open neighbourhood base at 0.

Proposition 6.8 *For any neighbourhood U of 0 there is a neighbourhood V of 0 such that $V + V \subseteq U$. Moreover, V can be chosen to be symmetric, i.e., such that $V = -V$.*

Proof As discussed above, for any neighbourhood U of 0, there are neighbourhoods V_1 and V_2 of 0 such that $V_1 + V_2 \subseteq U$. The result follows by setting $V = V_1 \cap V_2 \cap (-V_1) \cap (-V_2)$. ■

This proposition offers a substitute for the triangle inequality or $\varepsilon/2$-arguments which are available in a normed space but not in a general topological vector space. Note that, by induction, for any neighbourhood U of 0 and any $n \in \mathbb{N}$, there is a (symmetric) neighbourhood V of 0 such that $V + \cdots + V \subseteq U$, where the left hand side consists of n terms.

Corollary 6.9 *For any topological vector space X and any $n \in \mathbb{N}$, the map $((t_1, \dots, t_n), (x_1, \dots, x_n)) \mapsto t_1 x_1 + \cdots + t_n x_n$ is continuous from $\mathbb{K}^n \times X^n$ to X.*

Proof Let $((s_1, \dots, s_n), (x_1, \dots, x_n)) \in \mathbb{K}^n \times X^n$, and let W be any neighbourhood of $z = s_1 x_1 + \cdots + s_n x_n$ in X. Then $-z + W$ is a neighbourhood of 0 and so, by Proposition 6.8, there is a neighbourhood V of 0 such that

$$\underbrace{V + \cdots + V}_{n \text{ terms}} \subseteq -z + W.$$

Now, for each $1 \le i \le n$, $s_i x_i + V$ is a neighbourhood of $s_i x_i$ and therefore there is $\delta_i > 0$ and a neighbourhood U_i of x_i such that $tU_i \subseteq s_i x_i + V$ whenever $|t - s_i| < \delta_i$.

Set $\delta = \min\{\delta_1, \dots, \delta_n\}$. Then, for any $x'_i \in U_i$, $1 \le i \le n$,

$$\begin{aligned} t_1 x'_1 + \cdots + t_n x'_n &\in s_1 x_1 + V + \cdots + s_n x_n + V \\ &\subseteq s_1 x_1 + \cdots + s_n x_n + (-z + W) \\ &= W \end{aligned}$$

whenever $|t_i - s_i| < \delta$ and the result follows. ■

Next, we shall show that the continuity of the linear operations implies that the space is Hausdorff provided every point is closed.

Proposition 6.10 *A topological vector space $(X, \mathcal{T})$ is separated if and only if every one-point set is closed.*

Proof Of course, if $\mathcal{T}$ is a Hausdorff topology, then all one-point sets are closed. For the converse, let x and y be distinct points in a topological vector space X, and let $w = x - y$, so that $w \neq 0$. Now, if $\{w\}$ is closed, $X \setminus \{w\}$ is open. Indeed, $X \setminus \{w\}$ is an open neighbourhood of 0. The continuity of addition (at 0) implies that there are neighbourhoods U and V of 0 such that

$$U + V \subseteq X \setminus \{w\}.$$

Since V is a neighbourhood of 0, so is $-V$ (because the scalar multiplication M_{-1} is a homeomorphism) and therefore $w - V$ is a neighbourhood of w (because the translation T_w is a homeomorphism). Any $z \in w - V$ has the form $z = w - v$ with $v \in V$, and so $w = z + v$. Since $w \notin U + V$, we must have that $z \notin U$. Therefore $U \cap (w - V) = \varnothing$, and so, translating by y, $(y + U) \cap (y + w - V) = \varnothing$, that is, $(y + U) \cap (x - V) = \varnothing$. Hence $x - V$ and $y + U$ are disjoint neighbourhoods of x and y, respectively, and X is Hausdorff. ■

Definition 6.11 A subset A in a vector space X is said to be absorbing if for any $x \in X$ there is $\mu > 0$ such that $x \in tA$ whenever $|t| > \mu$, $t \in \mathbb{K}$. Setting $x = 0$, it follows that any absorbing set must contain 0. An equivalent requirement for A to be absorbing is that for any $x \in X$ there is some $\lambda > 0$ such that $sx \in A$ for all $s \in \mathbb{K}$ with $|s| < \lambda$.

Example 6.12 It is easy to see that a finite intersection of absorbing sets is absorbing. However, it may happen that an infinite intersection of absorbing sets fails to be absorbing. For example, let X be the linear space ℓ^∞ of all bounded complex sequences. For each $m \in \mathbb{N}$, let A_m be the subset consisting of those (a_n) in ℓ^∞ such that $|a_m| < \frac{1}{m}$. It is clear that each A_m is absorbing but their intersection $\bigcap_m A_m$, however, is not. Indeed, if $x = (x_n)$ is the sequence with $x_n = 1$ for each $n \in \mathbb{N}$, then sx is not in $\bigcap_m A_m$ for any $s \neq 0$.

Proposition 6.13 *Any neighbourhood of* 0 *in a topological vector space* X *is absorbing.*

Proof For any $x \in X$ and neighbourhood V of 0, the continuity of scalar multiplication (at $(0, x)$) ensures the existence of $\delta > 0$ such that $tx \in V$ whenever $|t| < \delta$. ■

In the vector space $\mathbb{R}^3$, we can readily visualise the proper subspaces (lines or planes) and we see that these can have no interior points, since any open set must contain a ball. In particular, no such subspace can be open. This result is quite general as we now show.

Proposition 6.14 *The only open linear subspace in a topological vector space* X *is* X *itself.*

Proof Suppose that M is an open linear subspace in a topological vector space X. Then 0 belongs to M so there is a neighbourhood V of 0 in X such that $V \subseteq M$. Let $x \in X$. Since V is absorbing, there is $\delta > 0$ such that $tx \in V$ whenever $|t| < \delta$. In particular, $tx \in M$ for some (suitably small) t with $t \neq 0$. But M is linear, so it follows that $x = t^{-1}tx \in M$, and we have $M = X$. ■

In fact, a proper linear subspace M of a topological vector space X can have no interior points at all. To see this, suppose that x is an interior point of M. Then there is a neighbourhood U of x with $U \subseteq M$. But then $-x + U$ belongs to M and is a neighbourhood of zero. We now argue as before to conclude that $M = X$.

Definition 6.15 A subset non-empty subset B in a vector space X is said to be balanced if $tB \subseteq B$ for all $t \in \mathbb{K}$ with $|t| \leq 1$.

Note that any balanced set contains 0 and that if B is balanced, then $B \subseteq rB$ for all r with $|r| \geq 1$.

Example 6.16 The balanced subsets of $\mathbb{C}$, considered as a complex vector space, are all the discs (either with or without their boundary) with centre at 0—together with $\mathbb{C}$ itself. If $\mathbb{C}$ is considered as a vector space over $\mathbb{R}$, then the balanced sets are given by unions of line segments symmetrical about and passing through 0. For example, the rectangle $\{z \in \mathbb{C} : |\operatorname{Re} z| < 1 \text{ and } |\operatorname{Im} z| \leq 1\}$ is a balanced set in $\mathbb{C}$ when treated as a real vector space.

Proposition 6.17 *Let A be a balanced subset of a topological vector space. Then $\overline{A}$, the closure of A, is balanced and, if 0 is an interior point of A, the interior of A is balanced.*

Proof For any $t \in \mathbb{K}$ with $0 < |t| \leq 1$, we have

$$t\,\overline{A} = M_t(\overline{A}) = \overline{M_t(A)} = \overline{tA} \subseteq \overline{A},$$

since M_t is a homeomorphism. This clearly also holds for $t = 0$ and so $\overline{A}$ is balanced.

Suppose that $0 \in \operatorname{Int} A$. Then, as before, for $0 < |t| \leq 1$,

$$t \operatorname{Int} A = \operatorname{Int}(tA) \subseteq tA \subseteq A.$$

But $t \operatorname{Int} A$ is open and so $t \operatorname{Int} A \subseteq \operatorname{Int} A$. This is also valid for $t = 0$ and it follows that $\operatorname{Int} A$ is balanced. ■

Proposition 6.18 *Any neighbourhood of 0 in a topological vector space contains a balanced neighbourhood of 0.*

Proof Let V be a given neighbourhood of 0 in a topological vector space X. Since scalar multiplication is continuous (at $(0,0)$), there is $\delta > 0$ and a neighbourhood U of 0 such that if $|\beta| < \delta$ and $x \in U$, then $\beta x \in V$, that is, $\beta U \subseteq V$ if $|\beta| < \delta$. Put $W = \bigcup_{|\beta|<\delta} \beta U$. Then $0 \in W \subseteq V$ and W is a neighbourhood of 0 since $\frac{\delta}{2} U \subseteq W$ and $\frac{\delta}{2} U$ is a neighbourhood of 0, by Corollary 6.6. Let $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$ and let $x \in W$. Then $x \in \beta U$ for some $\beta \in \mathbb{K}$ with $|\beta| < \delta$. Hence $\alpha x \in \alpha\beta U$. But $|\alpha\beta| \leq |\beta| < \delta$, so that $\alpha x \in W$ and therefore W is balanced. ■

Proposition 6.19 *Let X and Y be topological vector spaces and suppose that $T : X \to Y$ is a linear map. Then T is continuous if and only if T is continuous at 0.*

Proof Suppose that T is continuous at 0. Let $x \in X$ and suppose that (x_ν) is a net in X such that $x_\nu \to x$. Then $x_\nu - x \to x - x = 0$. By hypothesis, $T(x_\nu - x) \to T0 = 0$. That is, $Tx_\nu - Tx \to 0$, or $Tx_\nu - Tx + Tx \to Tx$. Thus $Tx_\nu \to Tx$ in Y and hence T is continuous at $x \in X$. The converse is clear. ■

Corollary 6.20 *Suppose that X and Y are normed spaces and $T : X \to Y$ is a linear map. The following statements are equivalent.*

(i) *T is continuous at 0.*

(ii) *T is continuous.*

(iii) *There is $C > 0$ such that $\|Tx\| \le C\|x\|$, for all $x \in X$.*

Proof By the proposition, (i) and (ii) are equivalent, and certainly (iii) implies (i). We shall show that (i) implies (iii). Let V be the neighbourhood of 0 in Y given by $V = \{y \in Y : \|y\| \le 1\}$. The continuity of T at 0 implies that there is a neighbourhood U of 0 in X such that $T(U) \subseteq V$. But U contains a ball of radius r, for some $r > 0$, so that $\|x\| < r$ implies that $Tx \in V$, that is, $\|Tx\| \le 1$. Hence, for any $x \in X$, with $x \ne 0$, we have that $\|T(rx/2\|x\|)\| \le 1$. It follows that

$$\|Tx\| \le \frac{2}{r}\,\|x\| \qquad x \in X,\ x \ne 0,$$

and this inequality persists for $x = 0$. ∎

Theorem 6.21 *Suppose that Λ is a linear functional on a topological vector space X, and suppose that $\Lambda(x) \ne 0$ for some $x \in X$. The following statements are equivalent.*

(i) *Λ is continuous.*

(ii) *$\ker \Lambda$ is closed.*

(iii) *Λ is bounded on some neighbourhood of 0.*

Proof Since $\ker \Lambda = \Lambda^{-1}(\{0\})$ and $\{0\}$ is closed in $\mathbb{K}$, it is clear that (i) implies (ii).

Suppose that $\ker \Lambda$ is closed. By hypothesis, $\ker \Lambda \ne X$, so the complement of $\ker \Lambda$ is a non-empty open set in X. Hence there is a point $x \notin \ker \Lambda$ and $V = X \setminus \ker \Lambda$ is a (open) neighbourhood of x. Let $U = V - x$. Then U is an open neighbourhood of 0 and so contains a balanced neighbourhood W, say, of 0. Thus $x + W \subseteq V$ and so $(x + W) \cap \ker \Lambda = \varnothing$. The linearity of Λ implies that the set $\Lambda(W)$ is a balanced subset of $\mathbb{K}$ and so is either bounded or equal to the whole of $\mathbb{K}$. If $\Lambda(W) = \mathbb{K}$, there is some $w \in W$ such that $\Lambda(w) = -\Lambda(x)$. This gives $\Lambda(x + w) = 0$, which is impossible since $x + W$ contains no points of the kernel of Λ. It follows that $\Lambda(W)$ is bounded, that is, there is some $K > 0$ such that $|\Lambda(w)| < K$ for all $w \in W$. Hence (ii) implies (iii).

If (iii) holds, there is some neighbourhood V of 0 and some $K > 0$ such that $|\Lambda(v)| < K$ for all $v \in V$. Let $\varepsilon > 0$ be given. Then $\frac{\varepsilon}{K} V$ is a neighbourhood of 0, and if $x \in \frac{\varepsilon}{K} V$ we have $\frac{K}{\varepsilon} x \in V$ so that $|\Lambda(\frac{K}{\varepsilon} x)| < K$, that is, $|\Lambda(x)| < \varepsilon$. It follows that Λ is continuous at 0, and hence Λ is continuous on X, by Proposition 6.19. ∎

Remark 6.22 In a normed space, any neighbourhood of 0 contains a ball $\{x \in X : \|x\| < r\}$ for some $r > 0$. Thus property (iii) of the theorem is equivalent to the property that for some $K > 0$, there is $r > 0$ such that

$$|\Lambda(x)| < K \quad \text{for } x \in X \text{ with } \|x\| < r.$$

If $x \neq 0$, then $y = rx/2\|x\|$ has norm less than r so that $|\Lambda(y)| < K$, i.e., $|\Lambda(x)| < \frac{2K}{r}\|x\|$. Setting $\frac{2K}{r} = M$ and allowing for $x = 0$, we see that (iii) is equivalent to

$$|\Lambda(x)| \le M\|x\|, \quad x \in X.$$

This is the property that Λ be a bounded linear functional—in which case we have $\|\Lambda\| \le M$.

Theorem 6.23 *Every finite dimensional Hausdorff topological vector space $(X, \mathfrak{T})$ over $\mathbb{K}$ with dimension n is linearly homeomorphic to $\mathbb{K}^n$. In particular, any two norms on a finite dimensional vector space are equivalent.*

Proof Suppose that $\dim X = n$ and that $x_1, \dots, x_n$ is a basis for X. Let $e_1, \dots, e_n$ be the standard basis for $\mathbb{K}^n$, i.e., $e_i = (\delta_{ij})_{j=1}^n \in \mathbb{K}^n$, and let $\varphi : \mathbb{K}^n \to X$ be the linear map determined by the assignment $\varphi(e_i) = x_i$, so that

$$\varphi((t_1, \dots, t_n)) = \varphi(t_1 e_1 + \cdots + t_n e_n) = t_1 x_1 + \cdots + t_n x_n.$$

It is clear that $\varphi : \mathbb{K}^n \to X$ is a (linear) isomorphism between $\mathbb{K}^n$ and X. Furthermore, φ is continuous, by Corollary 6.9. We shall show that the linear map $\varphi^{-1} : X \to \mathbb{K}^n$ is continuous.

Let K be the surface of the unit ball in $\mathbb{K}^n$,

$$K = \{\zeta \in \mathbb{K}^n : \|\zeta\| = 1\}.$$

Then K is compact in $\mathbb{K}^n$ and so $\varphi(K)$ is compact in X, because φ is continuous. Since X is Hausdorff, $\varphi(K)$ is closed in X. Now, $0 \notin K$ so that $0 = \varphi(0)$ is not an element of $\varphi(K)$. In other words, 0 is an element of the open set $X \setminus \varphi(K)$. By Proposition 6.18, there is a balanced neighbourhood W of 0 with

$$W \subseteq X \setminus \varphi(K).$$

In particular, $W \cap \varphi(K) = \varnothing$. Since φ^{-1} is linear, it follows that $\varphi^{-1}(W)$ is a balanced set in $\mathbb{K}^n$ such that $\varphi^{-1}(W) \cap K = \varnothing$. We claim that $\|\zeta\| < 1$ for every $\zeta \in \varphi^{-1}(W)$. Indeed, if $\zeta \in \varphi^{-1}(W)$ and if $\|\zeta\| \ge 1$, then $\|\zeta\|^{-1}\zeta \in K$ and also belongs to the balanced set $\varphi^{-1}(W)$. This contradicts $\varphi^{-1}(W) \cap K = \varnothing$ and we conclude that every $\zeta \in \varphi^{-1}(W)$ satisfies $\|\zeta\| < 1$, as claimed.

For each $1 \le i \le n$, let $\ell_i : \mathbb{K}^n \to \mathbb{K}$ be the projection map $\ell_i((t_1, \dots, t_n)) = t_i$. Then $\ell_i \circ \varphi^{-1} : X \to \mathbb{K}$ is a linear functional such that $|\ell_i \circ \varphi^{-1}(x)| < 1$ for all $x \in W$. By Theorem 6.21, $\ell_i \circ \varphi^{-1}$ is continuous, for each $1 \le i \le n$, and therefore

$$\varphi^{-1} : x \mapsto (\ell_1 \circ \varphi^{-1}(x), \dots, \ell_n \circ \varphi^{-1}(x))$$

is continuous. We conclude that $\varphi : \mathbb{K}^n \to X$ is a linear homeomorphism.

Now suppose that $|||\cdot|||$ is a norm on the finite dimensional vector space X. From the above, there is a linear homeomorphism $\psi : X \to \mathbb{K}^n$, where n is the dimension of X, and so there are positive constants m and M such that, for any $x \in X$,

$$\|\psi(x)\| \le M|||x||| \quad \text{since } \psi \text{ is continuous}$$

and

$$|||\psi^{-1}(x)||| \le m\|x\| \quad \text{since } \psi^{-1} \text{ is continuous}$$

where $\|\cdot\|$ is the usual Euclidean norm on $\mathbb{K}^n$. Hence

$$\frac{1}{m}|||x||| \le \|\psi(x)\| \le M|||x|||$$

which means that $|||\cdot|||$ is equivalent to the norm $x \mapsto \|\psi(x)\|$, $x \in X$. The result follows since equivalence of norms is an equivalence relation. ■

Corollary 6.24 *Every linear map from a finite dimensional Hausdorff topological vector space into a topological vector space is continuous.*

Proof Suppose that X and Y are topological vector spaces, with X finite dimensional and Hausdorff, and let $T : X \to Y$ be a linear map. Let $\{x_1, \dots, x_n\}$ be a basis for X. Then T is the composition of the maps

$$x = t_1x_1 + \cdots + t_nx_n (\in X) \mapsto (t_1, \dots, t_n)(\in \mathbb{K}^n) \mapsto t_1Tx_1 + \cdots + t_nTx_n \in Y.$$

The second of these is continuous, and by Theorem 6.23, so is the first and so, therefore, is T. ■

Proposition 6.25 *For any non-empty subset A in a topological vector space X and any $x \in \overline{A}$, there is a net (a_U) in A, indexed by $\mathcal{V}_0$, the set of neighbourhoods of 0 (partially ordered by reverse inclusion), such that $a_U \to x$.*

Proof For any $U \in \mathcal{V}_0$, $x+U$ is a neighbourhood of x. Since $x \in \overline{A}$, $(x+U) \cap A \ne \varnothing$. Let a_U be any element of $(x+U) \cap A$. Then (a_U) is a net in A such that $a_U \to x$. Indeed, for each $U \in \mathcal{V}_0$, $a_U - x \in U$ and so $a_U - x \to 0$ along $\mathcal{V}_0$, that is, $a_U \to x$. ■

Proposition 6.26 *For any non-empty sets A and B in a topological vector space X,*

(i) $\overline{A} + \overline{B} \subseteq \overline{A + B}$, *and*

(ii) $\overline{A} = \bigcap_{U \in \mathcal{N}_0}(A + U)$, *where $\mathcal{N}_0$ is any neighbourhood base at 0.*

Proof Let $x \in \overline{A}$ and $y \in \overline{B}$. Then there are nets (a_ν) and (b_ν), indexed by $\mathcal{V}_0$, such that $a_\nu \to x$, and $b_\nu \to y$. By continuity of addition, it follows that $a_\nu + b_\nu \to x + y$. Hence $x + y \in \overline{A + B}$, which proves (i).

To prove (ii), suppose that $x \in \overline{A}$ and let $\mathcal{N}_0$ be any neighbourhood base at 0. For any $U \in \mathcal{N}_0$, $-U$ is a neighbourhood of 0 and so $x - U$ is a neighbourhood of x. Since $x \in \overline{A}$ it follows that $(x - U) \cap A \neq \varnothing$, that is, $x \in A + U$, and so $\overline{A} \subseteq \bigcap_{U \in \mathcal{N}_0}(A + U)$. On the other hand, suppose that $x \in \bigcap_{U \in \mathcal{N}_0}(A + U)$ and that V is any neighbourhood of x. Then $-x + V$ is a neighbourhood of 0 and so also is $-(-x + V) = x - V$. Therefore there is some $U \in \mathcal{N}_0$ such that $U \subseteq x - V$. Now, $x \in A + U$, by hypothesis, and so $x \in A + U \subseteq A + x - V$. Hence there is $a \in A$ and $v \in V$ such that $x = a + x - v$, i.e., $a = v$. It follows that every neighbourhood of x contains some element of A and therefore $x \in \overline{A}$. ∎

Corollary 6.27 *The closed balanced neighbourhoods of* 0 *form a local neighbourhood base at* 0.

Proof Let U be any neighbourhood of 0. By Proposition 6.8, there is a neighbourhood V of 0 such that $V + V \subseteq U$. Also, by Proposition 6.18, there is a balanced neighbourhood of 0, W, say, with $W \subseteq V$, so that $W + W \subseteq U$. By Proposition 6.26, $\overline{W} \subseteq W + W \subseteq U$, and $\overline{W}$ is a neighbourhood of 0. We claim that $\overline{W}$ is balanced. To see this, let $x \in \overline{W}$ and let $t \in \mathbb{K}$ with $|t| \le 1$. There is a net (w_ν) in W such that $w_\nu \to x$. Then $tw_\nu \to tx$, since scalar multiplication is continuous. But $tw_\nu \in W$, since W is balanced, and so $tx \in \overline{W}$. It follows that $\overline{W}$ is balanced, as claimed. ∎

Proposition 6.28 *If M is a linear subspace of a topological vector space X, then so is its closure $\overline{M}$. In particular, any maximal proper subspace is either dense or closed.*

Proof Let M be a linear subspace of the topological vector space X. We must show that if $x, y \in \overline{M}$ and $t \in \mathbb{K}$, then $tx + y \in \overline{M}$. There are nets (x_ν) and (y_ν) in M, indexed by the base of all neighbourhoods of 0, such that $x_\nu \to x$ and $y_\nu \to y$. It follows that $tx_\nu \to tx$ and $tx_\nu + y_\nu \to tx + y$ and we conclude that $tx + y \in \overline{M}$, as required.

If M is a maximal proper subspace, the inclusion $M \subseteq \overline{M}$ implies that either $M = \overline{M}$, in which case M is closed, or $\overline{M} = X$, in which case M is dense in X. ∎

Corollary 6.29 *Let $\ell : X \to \mathbb{K}$ be a linear functional on a topological vector space X. Then either ℓ is continuous or $\ker \ell$ is a dense proper subspace of X.*

Proof If ℓ is zero, it is continuous and its kernel is the whole of X. Otherwise, $\ker \ell$ is a maximal proper linear subspace of X which is either closed or dense, by Proposition 6.28. However, ℓ is continuous if and only if its kernel is closed, so if ℓ is not continuous its kernel is a proper dense subspace. ■

It is convenient to introduce here the notion of boundedness in a topological vector space.

Definition 6.30 A subset B of a topological vector space is said to be bounded if for each neighbourhood U of 0 there is $s > 0$ such that $B \subseteq tU$ for all $t > s$.

Example 6.31 Let B be a subset of a normed space X, and, for $r > 0$, let B_r denote the ball $B_r = \{x : \|x\| < r\}$ in X. Any neighbourhood U of 0 contains some ball $B_{r'}$ and so $B_r \subseteq tB_{r'} \subseteq tU$ for all $t > r/r'$, that is, each B_r is bounded according to the definition above. On the other hand, if B is any bounded set, then there is $s > 0$ such that $B \subseteq tB_1$ for all $t > s$, that is, $\|x\| \le s$ for all $x \in B$. Therefore the notion of boundedness given by the above definition coincides with the usual notion of boundedness in a normed space.

Proposition 6.32 *If $(X, \mathfrak{T})$ is a topological vector space and if $z \in X$ and A and B are bounded subsets of X, then the sets $\{z\}$, $A \cup B$ and $A + B$ are bounded.*

Proof Let U be any neighbourhood of 0. Then U is absorbing and so $z \in tU$ for all sufficiently large $t > 0$, that is, $\{z\}$ is bounded.

By Proposition 6.8, there is a neighbourhood V of 0 such that $V + V \subseteq U$. Given that A and B are bounded, there is $s_1 > 0$ such that $A \subseteq tV$ for $t > s_1$ and there is $s_2 > 0$ such that $B \subseteq tV$ for $t > s_2$. Therefore

$$A \cup B \subseteq tV$$

and

$$A + B \subseteq tV + tV \subseteq tU$$

for all $t > \max\{s_1, s_2\}$. ■

Remark 6.33 It follows, by induction, that any finite set in a topological vector space is bounded. Also, taking $A = \{z\}$, we see that any translate of a bounded set is bounded.

Proposition 6.34 *Any convergent sequence in topological vector space is bounded.*

Proof Suppose that (x_n) is a sequence in a topological vector space $(X, \mathcal{T})$ such that $x_n \to z$. For each $n \in \mathbb{N}$, set $y_n = x_n - z$, so that $y_n \to 0$. Let U be any neighbourhood of 0. Let V be any balanced neighbourhood of 0 such that $V \subseteq U$. Then $V \subseteq sV$ for all s with $|s| \geq 1$. Since $y_n \to 0$, there is $N \in \mathbb{N}$ such $y_n \in V$ whenever $n > N$. Hence $y_n \in V \subseteq tV \subseteq tU$ whenever $n > N$ and $t \geq 1$. Set $A = \{y_1, \ldots, y_N\}$ and $B = \{y_n : n > N\}$. Then A is a finite set so is bounded and therefore $A \subseteq tU$ for all sufficiently large t. But then it follows that $A \cup B \subseteq tU$ for sufficiently large t, that is, $\{y_n : n \in \mathbb{N}\}$ is bounded and so is $\{x_n : n \in \mathbb{N}\} = z + (A \cup B)$. ■

Remark 6.35 A convergent *net* in a topological vector space need not be bounded. For example, let I be $\mathbb{R}$ equipped with its usual order and let $x_\alpha \in \mathbb{R}$ be given by $x_\alpha = e^{-\alpha}$. Then $(x_\alpha)_{\alpha \in I}$ is an unbounded but convergent net (with limit 0) in the real normed space $\mathbb{R}$.

7. Locally Convex Topological Vector Spaces

We have seen that a normed space is an example of a topological vector space. A further class of examples is got courtesy of seminorms rather than norms. Let $\mathcal{P} = \{p_\alpha : \alpha \in I\}$ be some family of seminorms on a vector space X. As a possible generalization of the idea of an open ball in a normed space, we consider the set

$$V(x_0, p_1, p_2, \dots, p_n; r) = \{x \in X : p_i(x - x_0) < r,\ 1 \le i \le n\}$$

where $x_0 \in X$, $r > 0$ and $p_1, p_2, \dots, p_n$ is a finite collection of seminorms in $\mathcal{P}$. We will use such sets to construct a topology on X in much the same way that open balls are used to determine a topology in a normed space. Indeed, if the family of seminorms contains just one member, which happens to be a norm, then our construction will give precisely the usual norm topology on X. Notice that

$$V(x_0, p_1, p_2, \dots, p_n; r) = x_0 + V(0, p_1, p_2, \dots, p_n; r).$$

The idea is to define local neighbourhood bases at each point of X via these sets. Thus we are just using translates of the collection centred at the origin and so we might hope that, indeed, we end up with a vector topology.

Theorem 7.1 *Let X be a vector space over $\mathbb{K}$ and let $\mathcal{P}$ be a family of seminorms on X. For each $x \in X$, let $\mathcal{N}_x$ denote the collection of all subsets of X of the form $V(x, p_1, p_2, \dots, p_n; r)$, with $n \in \mathbb{N}$, $p_1, \dots, p_n \in \mathcal{P}$ and $r > 0$. Let $\mathcal{T}$ be the collection of subsets of X consisting of $\varnothing$ together with all those subsets G of X such that for any $x \in G$ there is some $U \in \mathcal{N}_x$ such that $U \subseteq G$. Then $\mathcal{T}$ is a topology on X compatible with the vector space structure and the sets $\mathcal{N}_x$ form an open local neighbourhood base at x. Furthermore, each seminorm $p \in \mathcal{P}$ is continuous. $(X, \mathcal{T})$ is Hausdorff if and only if the family $\mathcal{P}$ is separating, that is, for any $x \in X$ with $x \ne 0$, there is some $p \in \mathcal{P}$ such that $p(x) \ne 0$.*

Proof Evidently $X \in \mathcal{T}$, and it is clear that the union of any family of elements of $\mathcal{T}$ is also a member of $\mathcal{T}$. We shall show that if $A, B \in \mathcal{T}$ then $A \cap B \in \mathcal{T}$.

If this intersection is empty, there is no more to be done, so suppose that $x \in A \cap B$. Then $x \in A$ and $x \in B$ and so there exist $U, V \in \mathcal{N}_x$ such that $U \subseteq A$ and $V \subseteq B$. Suppose that $U = V(x, p_1, \dots, p_m; r)$ and $V = V(x, q_1, \dots, q_n; s)$. Put $W = V(x, p_1, \dots, p_m, q_1, \dots, q_n; t)$ where $t = \min\{r, s\}$. Then $W \in \mathcal{N}_x$ and $W \subseteq U \cap V \subseteq A \cap B$. It follows that $\mathcal{T}$ is a topology on X.

Let $x \in X$ and let $U \in \mathcal{N}_x$. We shall show that U is open. (The method is the analogue of showing that for any point in an open disc D in $\mathbb{R}^2$, there is a (possibly very small) disc centred on the given point and wholly contained in D.) Suppose that $U = V(x, p_1, \dots, p_n; r)$ and that $z \in U$. Then $p_i(z - x) < r$ for $1 \le i \le n$. Let $\delta > 0$ be such that $\delta < r - p_i(z - x)$ for $1 \le i \le n$. For any $1 \le i \le n$ and any $y \in X$ with $p_i(y - z) < \delta$, we have

$$\begin{aligned} p_i(y - x) &\le p_i(y - z) + p_i(z - x) \\ &< \delta + p_i(z - x) \\ &< r. \end{aligned}$$

Hence $V(z, p_1, \dots, p_n; \delta) \subseteq V(x, p_1, \dots, p_n; r) = U$ and so $U \in \mathcal{T}$. Thus $\mathcal{N}_x$ is a neighbourhood base at x, consisting of open sets.

Next we shall show that $\mathcal{T}$ is a compatible topology. To this end, suppose that $(x_\nu, y_\nu) \to (x, y)$ in $X \times X$. We want to show that $x_\nu + y_\nu \to x + y$. Let $V(x + y, p_1, \dots, p_n; r)$ be a given basic neighbourhood of $x + y$. Since $(x_\nu, y_\nu) \to (x, y)$, there is ν_0 such that $(x_\nu, y_\nu) \in V(x, p_1, \dots, p_n; \frac{r}{2}) \times V(y, p_1, \dots, p_n; \frac{r}{2})$ whenever $\nu \succeq \nu_0$. For any $1 \le i \le n$ and $\nu \succeq \nu_0$

$$\begin{aligned} p_i(x + y - (x_\nu + y_\nu)) &\le p_i(x - x_\nu) + p_i(y - y_\nu) \\ &< \frac{r}{2} + \frac{r}{2} \end{aligned}$$

and so $x_\nu + y_\nu \in V(x + y, p_1, \dots, p_n; r)$. Hence $x_\nu + y_\nu \to x + y$, as required.

Now suppose that $(t_\nu, x_\nu) \to (t, x)$ in $\mathbb{K} \times X$. Let $V(tx, p_1, \dots, p_k; r)$ be a basic neighbourhood of tx. For any given $\varepsilon > 0$ and $s > 0$, there is ν_0 such that

$$(t_\nu, x_\nu) \in \{\zeta \in \mathbb{K} : |\zeta - t| < \varepsilon\} \times V(x, p_1, \dots, p_k; s).$$

Hence, for all $1 \le i \le k$ and $\nu \succeq \nu_0$,

$$\begin{aligned} p_i(tx - t_\nu x_\nu) &\le p_i(tx - t_\nu x) + p_i(t_\nu x - t_\nu x_\nu) \\ &\le |t - t_\nu| p_i(x) + |t_\nu| p_i(x - x_\nu) \\ &< \varepsilon p_i(x) + (|t| + \varepsilon) s \\ &< r \end{aligned}$$

if we choose $\varepsilon > 0$ such that $\varepsilon p_i(x) < \frac{r}{2}$, $1 \le i \le k$, and then $s > 0$ such that $(|t| + \varepsilon)s < \frac{r}{2}$. Therefore $t_\nu x_\nu \in V(tx, p_1, \dots, p_k; r)$ whenever $\nu \succeq \nu_0$. We conclude that $(t, x) \mapsto tx$ is continuous and therefore $\mathcal{T}$ is a vector space topology on X.

To show that each $p \in \mathcal{P}$ is continuous, let $\varepsilon > 0$ and suppose that $x_\nu \to x$ in $(X, \mathcal{T})$. Then there is ν_0 such that $x_\nu \in V(x, p; \varepsilon)$ whenever $\nu \succeq \nu_0$. Hence

$$|p(x) - p(x_\nu)| \le p(x - x_\nu) < \varepsilon$$

whenever $\nu \succeq \nu_0$ and it follows that $p : X \to \mathbb{R}$ is continuous.

It remains to show that $(X, \mathcal{T})$ is Hausdorff if and only if $\mathcal{P}$ is a separating family. Suppose that $\mathcal{P}$ is separating, and let $x, y \in X$ with $x \neq y$. Then there is some $p \in \mathcal{P}$ such that $\delta = p(x - y) > 0$. The sets $V(x, p; \frac{\delta}{2})$ and $V(y, p; \frac{\delta}{2})$ are disjoint neighbourhoods of x and y so that $(X, \mathcal{T})$ is Hausdorff.

Conversely, suppose that $(X, \mathcal{T})$ is Hausdorff (which follows, as we have seen, from the assumption that each one-point set be closed). For any given $x \in X$, with $x \neq 0$, there is a neighbourhood of 0 not containing x. In particular, there is $p_1, \dots, p_m \in \mathcal{P}$ and $r > 0$ such that $x \notin V(0, p_1, \dots, p_m; r)$. It follows that $p_i(x - 0) = p_i(x) \ge r$, for some $1 \le i \le m$, and so certainly $p_i(x) > 0$ and we see that $\mathcal{P}$ is a separating family of seminorms on X. ■

Definition 7.2 The topology $\mathcal{T}$ on a vector space X over $\mathbb{K}$ constructed above is called the (vector space) topology determined by the given family of seminorms $\mathcal{P}$.

Theorem 7.3 *Let $\mathcal{T}$ be the vector space topology on a vector space X determined by a family $\mathcal{P}$ of seminorms. A net (x_ν) converges to 0 in $(X, \mathcal{T})$ if and only if $p(x_\nu) \to 0$ for each $p \in \mathcal{P}$.*

Proof Suppose that $x_\nu \to 0$ in $(X, \mathcal{T})$. Then $p(x_\nu) \to p(0) = 0$ for each $p \in \mathcal{P}$, since each such p is continuous.

Conversely, suppose that $p(x_\nu) \to 0$ for each $p \in \mathcal{P}$. Let $p_1, \dots, p_m \in \mathcal{P}$ and let $r > 0$. Then there is ν_0 such that $p_i(x_\nu) < r$ whenever $\nu \succeq \nu_0$, $1 \le i \le m$. Hence $x_\nu \in V(0, p_1, \dots, p_m; r)$ whenever $\nu \succeq \nu_0$. It follows that $x_\nu \to 0$. ■

Remark 7.4 The convergence of a net (x_ν) to x is not necessarily implied by the convergence of $p(x_\nu) \to p(x)$ in $\mathbb{R}$ for each $p \in \mathcal{P}$. Indeed, for any $x \neq 0$ and any $p \in \mathcal{P}$, $p((-1)^n x) \to p(x)$, as $n \to \infty$, but it is not true that $(-1)^n x \to x$ if $(X, \mathcal{T})$ is separated. We will come back to this observation later.

Theorem 7.5 *The topology $\mathcal{T}$ on a vector space X over $\mathbb{K}$ given by a family of seminorms $\mathcal{P}$ is the weakest compatible topology on X such that each member of $\mathcal{P}$ is continuous at* 0.

Proof Let $\mathcal{T}'$ be the weakest vector space topology on X with respect to which each member of $\mathcal{P}$ is continuous at 0. Certainly $\mathcal{T}' \subseteq \mathcal{T}$. Let $p_1, \dots, p_n \in \mathcal{P}$ and let $r > 0$. Then

$$V(0, p_1, \dots, p_n; r) = \bigcap_{i=1}^{n} \{x \in X : p_i(x) < r\}$$

belongs to $\mathcal{T}'$ since each $\{x \in X : p_i(x) < r\} = p_i^{-1}((-\infty, r))$ is in $\mathcal{T}'$, $1 \le i \le n$. By hypothesis, translations are homeomorphisms and so $\mathcal{T}'$ contains all sets of the form $V(x, p_1, \dots, p_n; r)$ for $x \in X$. It follows that $\mathcal{T} \subseteq \mathcal{T}'$ and therefore we have equality $\mathcal{T} = \mathcal{T}'$. ■

Note that we did not use the *joint* continuity of addition $(x, y) \mapsto x + y$ nor that of scalar multiplication $(t, x) \mapsto tx$. This observation leads to the following.

Corollary 7.6 *The topology $\mathcal{T}$ determined by a given family $\mathcal{P}$ of seminorms on a vector space X is the weakest topology such that each member of $\mathcal{P}$ is continuous at* 0 *and such that for each fixed $x_0 \in X$ the translation $x \mapsto x + x_0$ is continuous.*

Proof The proof is exactly as above—the fact that each translation map T_{x_0}, $x_0 \in X$, is a homeomorphism follows from the continuity of the map $x \mapsto x + x_0$, $x \in X$. ■

Remark 7.7 It should be noted that the vector space topology on a vector space X determined by a family $\mathcal{P}$ of seminorms is not the same as the $\sigma(X, \mathcal{P})$-topology, the weakest topology on X making each member of $\mathcal{P}$ continuous. For example, let X be the real vector space $\mathbb{R}$ and let $\mathcal{P}$ be the family with a single member p given by $p(x) = |x|$, $x \in \mathbb{R}$. We observe, incidentally, that this family is separating. For any $s, t \in \mathbb{R}$, $s < t$, $p^{-1}((s, t))$ is equal to the symmetric set $I \cup (-I)$, where $I = [0, \infty) \cap (s, t)$. The weakest topology on $\mathbb{R}$ making p continuous is that consisting of all those subsets which are open in the usual sense and symmetrical about the origin. The sequences $(a_n)_{n \in \mathbb{N}} = ((-1)^n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}} = ((-1)^{n+1})_{n \in \mathbb{N}}$ both converge to 1 since every neighbourhood containing the point 1 also contains the point -1 (so this topology is not Hausdorff.) However, $a_n + b_n = 0$ for all n and so the sequence $(a_n + b_n)$ does not converge to $1 + 1 = 2$. We see that this topology is not compatible with the vector space structure of $\mathbb{R}$. Notice, however, that this topology and the usual (vector space) topology on $\mathbb{R}$ do share a neighbourhood base at 0; for example, the collection $\{(-r, r) : r > 0\}$. The point is that the vector topology on $\mathbb{R}$ is determined by this neighbourhood base at 0 *together with all its translates.* This is, in fact, strictly finer than the $\sigma(\mathbb{R}, \{p\})$-topology.

This observation is valid in general. If $\mathcal{P}$ is any family of seminorms on a vector space X ($\neq \{0\}$), then the $\sigma(X, \mathcal{P})$-topology has subbase given by the sets of the form $\{x : p(x) \in (s,t)\}$ for $p \in \mathcal{P}$ and $s < t$. Each of these sets is symmetrical about 0 in X, and so this property persists for all non-empty sets open with respect to the $\sigma(X, \mathcal{P})$-topology. This topology is never Hausdorff—the points x and $-x$ cannot be separated, for any $x \neq 0$. However, if $p \in \mathcal{P}$ and if $p(x) = r \neq 0$, then $V(x, p; r)$ and $V(-x, p; r)$ are disjoint open neighbourhoods of the points x and $-x$, respectively, with respect to the vector topology determined by $\mathcal{P}$. (Of course, if $\mathcal{P}$ is not separating, it will not be possible to separate x and $-x$ for *all* points x in X. In fact, if $p(x) = 0$ for some $x \neq 0$ in X and all $p \in \mathcal{P}$, we see that every $V(x, p; r)$ contains $-x$.)

We can rephrase this in terms of the continuity of the seminorms. We know that a net x_ν converges to x in X with respect to the $\sigma(X, \mathcal{P})$-topology if and only if $p(x_\nu) \to p(x)$ for each $p \in \mathcal{P}$. As we have noted earlier, the convergence of $p(x_\nu)$, for every $p \in \mathcal{P}$, does not imply the convergence of x_ν with respect to the vector topology determined by the family $\mathcal{P}$.

The sequence $p((-1)^n x)$ converges to $p(x) = p(-x)$ for any $p \in \mathcal{P}$ and any $x \in X$, and so it follows that $(-1)^n x$ converges both to x and to $-x$ with respect to the $\sigma(X, \mathcal{P})$-topology. This cannot happen in the vector topology for any pair x, p with $p(x) \neq 0$—because, as noted above, in this case x and $-x$ can be separated.

Theorem 7.8 *Suppose that ρ is a seminorm on a topological vector space $(X, \mathcal{T})$. The following statements are equivalent.*

(i) *ρ is continuous at 0.*

(ii) *ρ is continuous.*

(iii) *ρ is bounded on some neighbourhood of 0.*

If the topology $\mathcal{T}$ on X is the vector topology determined by a family $\mathcal{P}$ of seminorms, then (i), (ii) and (iii) are each equivalent to the following.

(iv) *There is a finite set of seminorms $p_1, \dots, p_m$ in $\mathcal{P}$ and a constant $C > 0$ such that*

$$\rho(x) \leq C(\, p_1(x) + \cdots + p_m(x)\,) \quad \text{for } x \in X.$$

Proof The inequality $|\rho(x) - \rho(y)| \leq \rho(x - y)$ implies that (ii) follows from (i)—if $x_\nu \to x$ in X, then $x_\nu - x \to 0$ so that $|\rho(x_\nu) - \rho(x)| \leq \rho(x_\nu - x) \to 0$, i.e., ρ is continuous at x.

Clearly (ii) implies (i), and (iv) implies (i)—if x_ν is any net converging to 0, then $p(x_\nu) \to 0$ for each p in $\mathcal{P}$, so that $\rho(x_\nu) \to 0 = \rho(0)$ by (iv).

Next, we observe that (i) implies that there is some neighbourhood U of 0 such that $\rho(U) \subseteq \{t \in \mathbb{K} : |t| < 1\}$, i.e., $|\rho(x)| < 1$ whenever $x \in U$ and so (iii) holds. Conversely, if (iii) holds, there is some neighbourhood V of 0 and some constant $C > 0$ such that $|\rho(x)| < C$ for any $x \in V$. It follows that for any $\varepsilon > 0$, $|\rho(x)| < \varepsilon$ whenever $x \in \frac{C}{\varepsilon} V$ and so ρ is continuous at 0, which is (i).

Finally, we show that (i) implies (iv). So suppose that ρ is continuous at 0. Then there is a neighbourhood U of 0 such that $\rho(U) \subseteq \{t \in \mathbb{R} : |t| < 1\}$, that is, $\rho(x) < 1$ whenever $x \in U$. But there are $p_1, \dots, p_m$ in $\mathcal{P}$ and $r > 0$ such that $V(0, p_1, \dots, p_m; r) \subseteq U$, and therefore $\rho(x) < 1$ for all $x \in V(0, p_1, \dots, p_m; r)$. Let $s(x) = p_1(x) + \cdots + p_m(x)$ and suppose x is such that $s(x) \neq 0$. Then $rx/s(x) \in V(0, p_1, \dots, p_m; r)$ and so $\rho(rx/s(x)) < 1$, that is,

$$\rho(x) < \frac{1}{r} \big(p_1(x) + \cdots + p_m(x) \big).$$

On the other hand, if $s(x) = 0$, then $s(nx) = 0$ for any $n \in \mathbb{N}$. Hence $nx \in V(0, p_1, \dots, p_m; r)$ and so $\rho(nx) = n\rho(x) < 1$ for all $n \in \mathbb{N}$. This forces $\rho(x) = 0$ and we conclude that, in any event,

$$\rho(x) \leq \frac{1}{r} \big(p_1(x) + \cdots + p_m(x) \big)$$

for all $x \in X$. ■

Corollary 7.9 *Suppose that $\Lambda : X \to \mathbb{K}$ is a linear functional on a topological vector space $(X, \mathcal{T})$, where $\mathcal{T}$ is the vector space topology determined by a family $\mathcal{P}$ of seminorms on X. The following statements are equivalent.*

(i) *Λ is continuous at 0.*

(ii) *Λ is continuous.*

(iii) *There is a finite set of seminorms $p_1, \dots, p_m$ in $\mathcal{P}$ and a constant $C > 0$ such that*
$$|\Lambda(x)| \leq C(\, p_1(x) + \cdots + p_m(x)\,) \quad \text{for } x \in X.$$

If $\mathcal{F}$ is a family of linear functionals on X such that $\mathcal{P}$ is the collection $\mathcal{P} = \{|\,\ell\,| : \ell \in \mathcal{F}\}$, then (i), (ii) and (iii) are equivalent to the following.

(iv) *There is a finite set $\ell_1, \dots, \ell_m$ of members of $\mathcal{F}$ and $s_1, \dots, s_m \in \mathbb{K}$ such that*
$$\Lambda(x) = s_1 \ell_1(x) + \cdots + s_m \ell_m(x) \quad \text{for } x \in X.$$

Proof We know already that (i) and (ii) are equivalent, by Proposition 6.19. For $x \in X$, set $\rho(x) = |\Lambda(x)|$. Then ρ is a seminorm on X. To say that ρ is continuous at 0 is to say that Λ is continuous at 0. Thus, by Theorem 7.8, (i) and (iii) are equivalent.

Next we show that (iv) follows from (iii). Indeed, if (iii) holds, and if $\ell_1, \dots, \ell_m$ are elements of $\mathcal{F}$ such that $p_i(x) = |\ell_i(x)|$ for $x \in X$ and $1 \le i \le m$, then it is clear that

$$\bigcap_{i=1}^{m} \ker \ell_i \subseteq \ker \Lambda \,.$$

The statement (iv) now follows, by Corollary 5.17. Finally, it is clear that (iv) implies (i), since each $|\ell_i|$, and hence each ℓ_i, $1 \le i \le m$, is continuous at 0, by Theorem 7.1 ■

Theorem 7.10 *Let X and Y be topological vector spaces with topologies determined by families $\mathcal{P}$ and $\mathcal{Q}$ of seminorms, respectively. For a linear map $T : X \to Y$, the following statements are equivalent.*

(i) *T is continuous at 0 in X.*

(ii) *T is continuous.*

(iii) *For any given seminorm q in $\mathcal{Q}$, there is a finite set of seminorms $p_1, \dots, p_m$ in $\mathcal{P}$ and a constant $C > 0$, possibly depending on q, such that*

$$q(Tx) \le C(p_1(x) + \cdots + p_m(x)) \quad \text{for } x \in X.$$

If, in particular, X and Y are normed spaces, then (i), (ii) and (iii) are equivalent to the following.

(iv) *There is a constant $C > 0$ such that*

$$\|Tx\| \le C\,\|x\| \quad \text{for } x \in X.$$

Proof We have already shown that (i) and (ii) are equivalent—as a direct consequence of the linearity of T (Proposition 6.19). The map $x \mapsto q(Tx)$ is a seminorm on X and so (ii) implies (iii), by Theorem 7.8. Suppose that (iii) holds. Let x_ν be any net converging to 0 in X. Then $p(x_\nu) \to 0$, for any p in $\mathcal{P}$, so that $q(Tx_\nu) \to 0$, by (iii), for any q in $\mathcal{Q}$. Hence $Tx_\nu \to 0$ in Y and (i) holds.

If X and Y are normed, then the families $\mathcal{P}$ and $\mathcal{Q}$ can be taken to be singleton sets consisting of just the norm. Clearly, (iv) follows from (iii) and (i) follows from (iv). ■

Definition 7.11 A subset A of a vector space over $\mathbb{K}$ is said to be convex if $sx + (1-s)y \in A$ whenever $x, y \in A$ and $0 \le s \le 1$. In other words, A is convex if the line segment between any two points in A also lies in A.

Suppose that p is a seminorm on a vector space X, $r > 0$ and let $A = \{x \in X : p(x) < r\}$. Then A is balanced and convex. Indeed, for any $x \in A$, if $|t| \le 1$, then $p(tx) = |t|\,p(x) < r$. Also, if $0 < s < 1$ and $x, y \in A$, then

$$\begin{aligned} p(sx + (1-s)y) &\le p(sx) + p((1-s)y) \\ &= s\,p(x) + (1-s)\,p(y) \\ &< sr + (1-s)r = r \end{aligned}$$

so that $s\,p(x) + (1-s)\,p(y) \in A$. Furthermore, for any $x \in X$ and $t > 0$, $p(tx) = t\,p(x) < r$ for all sufficiently small t. It follows that A is absorbing. To summarise, we have shown that A is a convex, balanced absorbing set. Since these properties are each preserved under (finite) intersections, it follows that the basic neighbourhoods $V(0, p_1, \ldots, p_n; r)$ of 0 determined by a family of seminorms on X are convex, balanced and absorbing.

Definition 7.12 A topological vector space is said to be locally convex if there is a neighbourhood base at 0 consisting of convex sets.

Thus, a topological vector space given by a family of seminorms is a locally convex topological vector space. In fact, we shall show that the converse holds, that is, any locally convex topological vector space is a topological vector space whose topology is determined by a family of seminorms. Such a family is given by certain seminorms associated with the collection of convex and balanced neighbourhoods of 0.

Proposition 7.13 *Any convex neighbourhood of* 0 *contains a balanced convex neighbourhood of* 0*. In particular, in any locally convex topological vector space, there is a neighbourhood base at* 0 *consisting of convex, balanced (and absorbing) sets.*

Proof Let U be any convex neighbourhood of 0. Then U contains a balanced neighbourhood W, say. For any $s \in \mathbb{K}$ with $|s| = 1$, we have $sW \subseteq W$ and $s^{-1}W \subseteq W$, so that $sW = W = s^{-1}W$. In particular, $s^{-1}W = W \subseteq U$ so that $W \subseteq sU$. Set $V = \bigcap_{|s|=1} sU$. Then $V \subseteq U$ and $W \subseteq V$ so that V is a neighbourhood of 0 contained in U. Furthermore, each sU is convex, and so, therefore, is their intersection, V. We claim that V is balanced. To see this, let $t \in \mathbb{K}$ with $|t| \le 1$. For any $v \in V$, we have $v \in sU$ for all s with $|s| = 1$. Hence, for any s with $|s| = 1$, $tv \in s|t|U \subseteq sU$, since U is convex and contains 0. It follows that $tv \in V$ and so V is balanced.

For the last part, we note that, by definition, in a locally convex topological vector space, 0 has a neighbourhood base of convex sets. By the above, each of these contains a balanced, convex neighbourhood of 0. The proof is completed by observing that any neighbourhood of 0 is absorbing. ■

Example 7.14 We have discussed continuous linear maps so it is natural to consider the possibility of discontinuous ones. In this connection, the existence of a Hamel basis proves useful in the construction of such various 'pathological' examples. We first consider the existence of discontinuous or, equivalently, by Theorem 7.10, unbounded linear functionals on a normed space. In some cases such functionals are not difficult to find. 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 $\phi : X \to \mathbb{C}$ 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 make use of a Hamel basis. We shall consider a somewhat more general setting. Suppose, then, that $(X, \mathcal{T})$ is an infinite dimensional topological vector space whose topology $\mathcal{T}$ is determined by a countable family $\mathcal{P}$ of seminorms and let Y be any topological vector space possessing at least one continuous seminorm q, say, such that $q(y) \neq 0$ for some $y \in Y$. Then there is a linear map $T : X \to Y$ such that T is not continuous at any point of X.

To construct such a map, let M be a Hamel basis of X. Then there is a sequence $\{u_n : n \in \mathbb{N}\}$ of distinct elements in M, since X is infinite dimensional. Let $\{y_n : n \in \mathbb{N}\}$ be *any* sequence in Y, with $q(y_n) \neq 0$. Define T on the u_n by

$$T(u_n) = n\left(p_1(u_n) + \cdots + p_n(u_n)\right)(1/q(y_n))\,y_n$$

and set $Tv = 0$ for $v \in M$ with $v \neq u_k$ for any $k \in \mathbb{N}$. The map T is extended to the whole of X by linearity, thus giving a linear map from X into Y. Now, evidently, $q(T(u_n)) = n(p_1(u_n) + \cdots + p_n(u_n))$ and so, by Theorem 7.10, T fails to be continuous, even at 0. T is then discontinuous at every point since continuity at some $z \in X$ is easily seen to imply continuity at 0 (and hence at every point, again by Theorem 7.10). If we set $Y = \mathbb{K}$, then T is an unbounded linear functional.

One might then ask whether the existence of such discontinuous linear functionals also holds for any infinite dimensional topological vector space. It turns out that discontinuous linear functionals *cannot* always be found. If the topology of a topological vector space is sufficiently strong, then *all* linear functionals are continuous.

This would certainly be the case if we could equip a vector space with the discrete topology. However, this is not a vector topology so we must look further. In fact, we will show that *any* vector space X can be equipped with a vector topology $\mathcal{T}$ such that every linear functional on X is continuous. To construct such a topology, let $\mathcal{P}$ denote the collection of *all* seminorms on X. $\mathcal{P}$ is not empty because X can be equipped with a norm, and so this will appear in the family $\mathcal{P}$. Let $\mathcal{T}$ be the vector topology on X determined by $\mathcal{P}$. We note that $\mathcal{T}$ is Hausdorff because $\mathcal{P}$ is separating—after all, it contains a norm. By Theorem 7.1, every member of $\mathcal{P}$ is continuous. Let $\ell : X \to \mathbb{K}$ be *any* linear functional on X. Then $|\,\ell\,|$ is a seminorm on X and therefore $|\,\ell\,|$ belongs to $\mathcal{P}$. It follows that $|\,\ell\,|$ is continuous at 0 and hence ℓ is continuous at 0. By Corollary 7.9, we conclude that ℓ is continuous.

Proposition 7.15 *For any seminorm p on a vector space X over $\mathbb{K}$, the sets $\{x \in X : p(x) < 1\}$ and $\{x \in X : p(x) \le 1\}$ are convex, absorbing and balanced.*

Proof Let $x, y \in \{x : p(x) < 1\}$ and let $0 < s < 1$. Then $p(sx + (1-s)y) \le p(sx) + p((1-s)y) = sp(x) + (1-s)p(y) < 1$ which shows that $\{x : p(x) < 1\}$ is convex. For any $z \in X$, we have $p(tz) = |t|p(z) < 1$ for all $t \in \mathbb{K}$ with $|t|$ sufficiently large, so that $\{x : p(x) < 1\}$ is absorbing. Now let $t \in \mathbb{K}$ with $|t| \le 1$. Then, $p(tx) = |t|p(x) \le p(x) < 1$ and we see that $\{x : p(x) < 1\}$ is balanced. An almost identical argument applies to the set $\{x : p(x) \le 1\}$. ■

This rather easy result tells us how to get convex, absorbing and balanced sets from seminorms. It turns out that one can go in the other direction—from convex, absorbing and balanced sets we can construct seminorms, as we now show.

Theorem 7.16 *Suppose that C is a convex absorbing set in a vector space X. For each $x \in X$, let $A_x = \{s > 0 : x \in sC\}$. Then A_x enjoys the following properties.*

(i) *$A_x \ne \varnothing$ for $x \in X$, $A_0 = (0, \infty)$, and $A_{tx} = t\,A_x$ for any $t > 0$.*

(ii) *For any $x, y \in X$, $A_x + A_y \subseteq A_{x+y}$.*

(iii) *If, in addition, C is balanced, then $A_{tx} = |t|\,A_x$ for any $t \in \mathbb{K}$ with $|t| > 0$.*

Proof A_x is non-empty because C is absorbing—in fact, A_x contains a set of the form (s, ∞) for some suitably large s. The set C contains 0 and so certainly $t0 = 0 \in C$ for all $t > 0$, i.e., $A_0 = (0, \infty)$. For any $t > 0$, the following statements

are equivalent; $s \in A_x$, $x \in sC$, $tx \in tsC$, $ts \in A_{tx}$, which completes the proof of (i).
To prove (ii), let x and y be given elements of X and let $a \in A_x$ and $b \in A_y$. Then we have that $x \in aC$ and $y \in bC$, and so $x + y \in aC + bC \subseteq (a + b)C$, by the convexity of C. That is, $a + b \in A_{x+y}$.

Now assume that C is also balanced and let $t \in \mathbb{K}$ with $|t| > 0$. Write $t = \alpha|t|$, where $|\alpha| = 1$. Then $\alpha C \subseteq C$ and also $\alpha^{-1}C \subseteq C$, so that $\alpha C = C$. Hence, the following statements are equivalent; $s \in A_{|t|x}$, $|t|x \in sC$, $\alpha|t|x \in \alpha sC$, $tx \in sC$, $s \in A_{tx}$. We conclude that $A_{|t|x} = A_{tx}$. ∎

Theorem 7.17 *For any convex absorbing set C in a vector space X the map $p_C : x \mapsto p_C(x) = \inf\{s > 0 : x \in sC\}$, $x \in X$, is positively homogeneous and subadditive on X.*

If, in addition, C is balanced, then p_C is a seminorm on X. Furthermore, $\{x : p_C(x) < 1\} \subseteq C \subseteq \{x : p_C(x) \le 1\}$, and if B is any convex, absorbing and balanced set such that $\{x : p_C(x) < 1\} \subseteq B \subseteq \{x : p_C(x) \le 1\}$ then $p_B = p_C$.

Proof We use the notation of the previous theorem, Theorem 7.16. By part (i) of Theorem 7.16, for any $x \in X$ and $t > 0$, we have that $p_C(tx) = \inf A_{tx} = \inf tA_x = t \inf A_x = tp_C(x)$ which shows that p_C is positively homogeneous.

Also, for any $x, y \in X$, we have (by part (ii) of Theorem 7.16), $p_C(x + y) = \inf A_{x+y} \le a + b$ for any $a \in A_x$, $b \in A_y$. Hence $p_C(x + y) - b \le \inf A_x$, $b \in A_y$, and so $p_C(x + y) - \inf A_x \le \inf A_y$ which gives $p_C(x + y) \le p_C(x) + p_C(y)$. That is, p_C is subadditive.

Now suppose that C is also balanced. We wish to show that p_C is a seminorm. Clearly $p_C(x) \ge 0$ and $p_C(0) = \inf A_0 = 0$. Now let $t \in \mathbb{K}$. Then $p_C(tx) = \inf A_{tx} = \inf A_{|t|x} = \inf |t|A_x = |t|p_C(x)$ and it follows that p_C is a seminorm on X, as required.

If $z \in \{x : p_C(x) < 1\}$, then $\inf A_z < 1$ and so there is some $s \in A_z$ with $s < 1$. Thus $z \in sC$ and so $(1/s)z \in C$. It follows that $z \in C$ since C is convex and contains 0. If $x \in C$, then $1 \in A_x$ and so $\inf A_x \le 1$, i.e., $p_C(x) \le 1$.

Suppose that D, D' are any two convex, absorbing, balanced sets such that $D \subseteq D'$. Then, for any $x \in X$, $\{s : sx \in D\} \subseteq \{s : sx \in D'\}$. Taking the infima, it follows that $p_{D'}(x) \le p_D(x)$. It follows that if B is any convex, absorbing and balanced set such that $D \subseteq B \subseteq D'$ then $p_{D'} \le p_B \le p_D$.

In particular, this holds with $D = \{x : p_C(x) < 1\}$ and $D' = \{x : p_C(x) \le 1\}$. However, $p_D(x) = \inf\{s > 0 : x \in sD\} = \inf\{s > 0 : x/s \in D\} = \inf\{s > 0 : p_C(x) < s\} = p_C(x)$. Similarly, $p_{D'}(x) = \inf\{s > 0 : x \in sD'\} = \inf\{s > 0 : x/s \in D'\} = \inf\{s > 0 : p_C(x) \le s\} = p_C(x)$ and we have shown that $p_D = p_{D'} = p_C$. We conclude that $p_C = p_{D'} \le p_B \le p_D = p_C$ and the proof is complete. ∎

Definition 7.18 The map p_C constructed in Theorem 7.17 is called the Minkowski gauge or functional associated with the convex absorbing set C in the vector space X.

Example 7.19 Let X be a normed space, and let C be the ball $C = \{x : \|x\| < r\}$, with $r > 0$. Then $p_C(x) = \inf\{s > 0 : x \in sC\} = \inf\{s > 0 : \|x/s\| < r\} = \inf\{s > 0 : \|x\| < sr\} = \|x\|/r$.

Example 7.20 Let S be the strip $S = \{z \in \mathbb{C} : |\operatorname{Im} z| < 1\}$ in the one-dimensional complex vector space $\mathbb{C}$. Then S is convex and absorbing but not balanced—for example, $2 \in S$ but $2i \notin S$. Given $s > 0$, we see that $z \in sS$ if and only if $|\operatorname{Im} z| < s$, so that $p_S(z) = |\operatorname{Im} z|$.

We see that $p_S(2) = 0 \neq 2 = p_S(2i)$, so that p_S fails to be absolutely homogeneous and so is not a seminorm. Notice that p_S is positively homogeneous and subadditive, as it should be.

Example 7.21 Let S be the strip $S = \{(x, y) \in \mathbb{R}^2 : |y| < 1\}$ in the two-dimensional real vector space $\mathbb{R}^2$. We see that S is convex, absorbing and balanced. For given $s > 0$, $(x, y) \in sS$ if and only if $|y| < s$, which implies that $p_S((x, y)) = |y|$. In this case, p_S *is* a seminorm; clearly $p_S((x, y)) \geq 0$, $p_S(t(x, y)) = |t|\,|y| = |t|p_(S(x, y))$, and

$$p_S((x, y) + (x', y')) = |y + y'| \leq |y| + |y'| = p_S((x, y)) + p_S((x', y')).$$

Notice that S is an unbounded set in $\mathbb{R}^2$ and that p_S is not a norm. Indeed, $p_S((x, 0)) = 0$ for any $x \in \mathbb{R}$.

Proposition 7.22 *The Minkowski functional p_V associated with a bounded, balanced, convex neighbourhood of* 0 *in a topological vector space X is a norm.*

Proof We recall that any neighbourhood of 0 is absorbing and so p_V is well-defined. Let $x \in X$ with $x \neq 0$. Then there is a neighbourhood W of 0 such that $x \notin W$. However, since V is bounded, $V \subseteq sW$ for all sufficiently large s. But $x \notin W$ implies that $sx \notin sW$, so that $sx \notin V$ for some (sufficiently large) s. By Theorem 7.17, it is false that $p_V(sx) < 1$, and so we must have $p_V(sx) \geq 1$. In particular, $p_V(x) = (1/s)p_V(sx) \neq 0$, and we conclude that the seminorm p_V is in fact a norm. ■

Example 7.23 Let X be the real vector space of bounded, continuously differentiable real-valued functions on $(-1, 1)$ which vanish at the origin and which have bounded derivatives. Let p_1 and p_2 be the norms on X given by $p_1(f) = \sup\{|f(x)| : x \in (-1, 1)\}$ and $p_2(f) = \sup\{|f'(x)| : x \in (-1, 1)\}$ for $f \in X$. Equip X with the vector space topology $\mathcal{T}$ given by the pair $\mathcal{P} = \{p_1, p_2\}$ and let

$C = \{f \in X : p_1(f) < 1\}$. Then we see that the Minkowski functional p_C is equal to the norm p_1. In particular, p_C is a norm. However, we claim that C is not bounded. To see this, let U be the neighbourhood of 0 given by

$$U = \{f \in X : p_1(f) < 1 \text{ and } p_2(f) < 1\}.$$

If C *were* bounded, then $C \subseteq sU$ for all sufficiently large s. However, for any $s > 0$, the function $g(x) = \frac{1}{2}\sin 4sx$ belongs to C but not to sU (since $p_2(g) = 2s > s$). It follows that C is not bounded, as claimed.

Theorem 7.24 *The topology on a separated topological vector space $(X, \mathcal{T})$ is a norm topology if and only if* 0 *possesses a bounded convex neighbourhood.*

Proof Suppose that the topology on the topological vector space X is given by a norm $\|\cdot\|$. Then the set $\{x : \|x\| < 1\}$ is an open bounded convex neighbourhood of 0.

Conversely, suppose that U is a bounded convex neighbourhood of 0. Then, by Proposition 7.13, there is a balanced, convex neighbourhood V of 0 with $V \subseteq U$. Since U is bounded, so is V. By Proposition 7.22, μ_V, the Minkowski functional associated with V is a norm.

We shall show that $\|\cdot\| = \mu_V(\cdot)$ induces the topology $\mathcal{T}$ on X. Since $\mu_V(x) \le 1$, for $x \in V$, by Theorem 7.17, it follows from Theorem 7.8 that μ_V is continuous. On the other hand, suppose that $\|x_\nu\| \to 0$. For any neighbourhood G of 0 there is $s > 0$ such that $V \subseteq sG$, since V is bounded. Now, there is ν_0 such that $\|x_\nu\| < 1/s$ for all $\nu \succeq \nu_0$. For such ν, we have, $\|sx_\nu\| < 1$ and so $sx_\nu \in V \subseteq sG$, by Theorem 7.17. Hence $x_\nu \in G$ whenever $\nu \succeq \nu_0$. It follows that $x_\nu \to 0$ with respect to $\mathcal{T}$. Hence, we have the following equivalent statements;

$$\begin{aligned} & x_\nu \to x \text{ with respect to } \mathcal{T}, \\ & \qquad \iff x_\nu - x \to 0 \text{ with respect to } \mathcal{T}, \\ & \qquad \iff x_\nu - x \to 0 \text{ with respect to } \|\cdot\|, \\ & \qquad \iff x_\nu \to x \text{ with respect to } \|\cdot\|. \end{aligned}$$

It follows that $\mathcal{T}$ and the topology given by the norm $\|\cdot\|$ have the same convergent nets, and hence the same closed sets and therefore the same open sets, i.e., they are equal. ∎

Theorem 7.25 *The topology of any locally convex topological vector space is determined by a family of seminorms. Indeed, this family may be taken to be the family $\mathcal{P}$ of Minkowski functionals associated with all convex, absorbing, balanced neighbourhoods of* 0.

Proof We have seen that in every locally convex topological vector space $(X, \mathcal{T})$, 0 possesses a neighbourhood base of convex, balanced (absorbing) neighbourhoods. Let $\mathcal{P}$ be the family of associated Minkowski functionals. Denote by $\mathcal{T}_{\mathcal{P}}$ the vector topology on X determined by $\mathcal{P}$. We must show that $\mathcal{T}_{\mathcal{P}} = \mathcal{T}$. Let $p \in \mathcal{P}$. Then, by definition, p is bounded on some neighbourhood of 0. Indeed, $p = p_C$ for some (convex, balanced and absorbing) neighbourhood C of 0 and so $C \subseteq \{x \in X : p_C(x) \le 1\}$. That is to say, if $x \in C$, then $p(x) = p_C(x) \le 1$. Thus p is bounded on the neighbourhood C of 0 and so is continuous, by Theorem 7.8. Since $\mathcal{T}_{\mathcal{P}}$ is the weakest vector topology on X such that each member of $\mathcal{P}$ is continuous, we deduce that $\mathcal{T}_{\mathcal{P}} \subseteq \mathcal{T}$.

On the other hand, suppose that U is a convex, balanced neighbourhood of 0 with respect to the topology $\mathcal{T}$. Then the interior of U is a convex, balanced neighbourhood of 0, so we may assume that U is open. Let $x \in U$, so that U is a neighbourhood of x. By continuity of scalar multiplication, there is $\delta > 0$ such that $rx \in U$ whenever $|r - 1| < \delta$. In particular, there is $r > 1$ such that $rx \in U$, i.e., $x \in \frac{1}{r}U$. Hence $p_U(x) = \inf\{s > 0 : x \in sU\} < 1$ for any $x \in U$. It follows from Theorem 7.17 that $U = \{x : p_U(x) < 1\}$. This means that $U = p_U^{-1}((-\infty, 1))$ is in $\mathcal{T}_{\mathcal{P}}$. Now, any non-empty $\mathcal{T}$-open set is a union of translations of open convex, balanced neighbourhoods of 0, so we conclude that any such set is also $\mathcal{T}_{\mathcal{P}}$-open, i.e., $\mathcal{T} \subseteq \mathcal{T}_{\mathcal{P}}$, and we have equality, as required. ■

Proposition 7.26 *A subset B of a locally convex topological vector space $(X, \mathcal{T})$ is bounded if and only if $p(B)$ is a bounded subset of $\mathbb{R}$ for each continuous seminorm p on X.*

Proof Suppose that B is bounded and p is a continuous seminorm on X. Then $U = \{x : p(x) < 1\}$ is a neighbourhood of 0 in X. Since B is bounded, there is $s > 0$ such that $B \subseteq sU$. But then if $x \in B$, we have $(1/s)x \in U$ and therefore $p((1/s)x) < 1$. Thus $p(x) < s$ for all $x \in B$ and $p(B)$ is bounded.

Conversely, suppose that $p(B)$ is bounded for each continuous seminorm p on X and let U be a neighbourhood of 0. Since $(X, \mathcal{T})$ is locally convex, there is a family of seminorms $\mathcal{P}$ on X which determine the topology $\mathcal{T}$, by Theorem 7.25. Thus there is some $r > 0$ and seminorms $p_1, \dots, p_n \in \mathcal{P}$ such that $V(0, p_1, \dots, p_n; r) \subseteq U$. Now, each member of $\mathcal{P}$ is continuous and so, by hypothesis, each $p_i(B)$ is bounded, so there is some $d > 0$ such that $p_i(x) \le d$ for $x \in B$ and $1 \le i \le n$.

Hence, for $x \in B$,

$$p_i(x/s) \le \frac{d}{s} < r \quad \text{whenever } s > d/r,$$

i.e., $x/s \in V(0, p_1, \dots, p_n; r) \subseteq U$, whenever $s > d/r$. It follows that

$$B \subseteq sV(0, p_1, \dots, p_n; r) \subseteq sU \quad \text{for } s > d/r$$

and we conclude that B is bounded, as required. ■

Next we consider a further version of the Hahn-Banach theorem.

Theorem 7.27 (Hahn-Banach theorem) *Let X be a locally convex topological vector space determined by the family $\mathcal{P}$ of seminorms. Suppose that M is a linear subspace of X and that $\lambda : M \to \mathbb{K}$ is a continuous linear functional on M. Then there is a constant $C > 0$ and a finite set of elements $p_1, \dots, p_n$ in $\mathcal{P}$, such that*

$$|\lambda(x)| \le C(\, p_1(x) + \cdots + p_n(x)\,) \quad \text{for } x \in M.$$

Furthermore, there is a continuous linear functional Λ on X such that

$$|\Lambda(x)| \le C(\, p_1(x) + \cdots + p_n(x)\,) \quad \text{for } x \in X,$$

and $\Lambda \upharpoonright M = \lambda$.

Proof The relative topology on M is determined by the restrictions of the seminorms in $\mathcal{P}$ to M. Indeed, a neighbourhood base at 0 in M is given by the sets

$$\begin{aligned} V(0, p_1, \dots, p_k; r) \cap M &= \{x \in M : p_i(x) < r,\ 1 \le i \le k\} \\ &= \{x \in M : (p_i \upharpoonright M)(x) < r,\ 1 \le i \le k\} \\ &= V(0, p_1 \upharpoonright M, \dots, p_k \upharpoonright M; r). \end{aligned}$$

Let $\mathcal{P}_M$ denote the collection of restrictions $\mathcal{P}_M = \{p \upharpoonright M : p \in \mathcal{P}\}$. Since $\lambda : M \to \mathbb{K}$ is a continuous linear functional on M, it follows that there is $C > 0$ and $p_1 \upharpoonright M, \dots, p_n \upharpoonright M$ in $\mathcal{P}_M$ such that

$$|\lambda(x)| \le C\,(\, p_1 \upharpoonright M(x) + \cdots + p_n \upharpoonright M(x)\,), \quad x \in M,$$

which is precisely the statement that

$$|\lambda(x)| \le C\,(\, p_1(x) + \cdots + p_n(x)\,), \quad x \in M.$$

For $x \in X$, set $q(x) = C(p_1(x) + \cdots + p_n(x))$. Then q is a seminorm on X and we have $|\lambda(x)| \le q(x)$ on M. By our earlier version of the Hahn-Banach theorem, Theorem 5.22, there is a linear functional Λ on X such that $\Lambda \upharpoonright M = \lambda$ and

$$|\Lambda(x)| \le q(x) \quad \text{for } x \in X.$$

This bound implies that Λ is continuous, by Corollary 7.9. ■

Corollary 7.28 *The space of continuous linear functionals on a separated locally convex topological vector space separates the points of X, that is, for any $x_1 \neq x_2$ in X there is a continuous linear functional Λ on X such that $\Lambda(x_1) \neq \Lambda(x_2)$.*

Proof Given $x_1 \neq x_2$ in X, set $z = x_1 - x_2$. Then $z \neq 0$. Let M be the linear subspace $M = \{tz : t \in \mathbb{K}\}$ and define $\lambda : M \to \mathbb{K}$ by $\lambda(tz) = t$. Then λ is a linear functional on M with $\ker \lambda = \{0\}$ which is closed in M, since X and therefore M is Hausdorff, by hypothesis. It follows that λ is continuous on M. Hence, by Theorem 7.27, there is a continuous linear functional $\Lambda : X \to \mathbb{K}$ such that $\Lambda \upharpoonright M = \lambda$. Furthermore, $\Lambda(x_1) - \Lambda(x_2) = \Lambda(z) = \lambda(z) = 1 \neq 0$. ■

Remark 7.29 It is worthwhile explicitly highlighting the following statement which has more or less been proved in the course of the above argument. For any non-zero element z in a separated topological vector space X, the map $tz \mapsto t$ is a homeomorphism between M, the one-dimensional subspace spanned by z, and $\mathbb{K}$. It is clear that this map is one-one and onto. Indeed, its inverse is $t \mapsto tz$, which is continuous due to the continuity of scalar multiplication. The somewhat less trivial part is in showing that $tz \mapsto t$ is continuous—which as we saw above, follows from the fact that a linear functional is continuous if (and only if) its kernel is closed.

Corollary 7.30 *Let M be a linear subspace of a locally convex topological vector space X and suppose that x_0 is an element of X not belonging to the closure of M. Then there is a continuous linear functional Λ on X such that $\Lambda(x_0) = 1$ and $\Lambda(x) = 0$ for all $x \in M$.*

Proof Let B be a Hamel basis for $\overline{M}$, the closure of M. Since $\overline{M}$ is a linear subspace and $x_0 \notin \overline{M}$, it follows that $\{x_0\} \cup B$ is a linearly independent set. Let M_1 be the linear span of $\overline{M}$ and x_0, that is, $M_1 = \{tx_0 + u : t \in \mathbb{K}, u \in \overline{M}\}$. Any $x \in M_1$ can be written as $x = tx_0 + u$ for unique $t \in \mathbb{K}$ and $u \in \overline{M}$. Thus we may define $\lambda : M_1 \to \mathbb{K}$ by

$$\lambda(x) = t, \quad \text{for } x = tx_0 + u \in M_1.$$

Now, $\overline{M}$ is closed and $x_0 \notin \overline{M}$ so there is a neighbourhood of x_0 which does not meet $\overline{M}$. Thus,

$$V(x_0, p_1, \dots, p_n; r) \cap \overline{M} = \varnothing$$

for suitable seminorms $p_1, \dots, p_n$ and $r > 0$. Let $s(x) = p_1(x) + \cdots + p_n(x)$, $x \in X$. Then, for any $u \in \overline{M}$, $p_i(x_0 - u) \geq r$ for some $1 \leq i \leq n$ so that $s(x_0 - u) \geq r$. In particular, for any $t \neq 0$ and $u \in \overline{M}$,

$$s(tx_0 + u) = |t|\, s\big(x_0 + \frac{u}{t}\big) \geq |t|\, r.$$

Thus, for $t \in \mathbb{K}$ and $u \in \overline{M}$,

$$|\lambda(tx_0 + u)| = |t| \le \frac{1}{r}\, s(tx_0 + u).$$

That is,

$$|\lambda(x)| \le \frac{1}{r}\, s(x) \quad \text{for } x \in M_1$$

which shows that $\lambda : M_1 \to \mathbb{K}$ is continuous. By the Hahn-Banach theorem, Theorem 7.27, it follows that there is a continuous linear functional Λ on X which extends λ (and obeys the same bound). In particular, $\Lambda(x_0) = 1$ and $\Lambda(x) = 0$ for all $x \in M$. ■

Theorem 7.31 *Let X be a vector space over $\mathbb{K}$ and suppose that $\mathcal{F}$ is a family of linear functionals on X. Then the $\sigma(X, \mathcal{F})$-topology coincides with the locally convex topology $\mathcal{T}_{\mathcal{P}}$ determined by the family $\mathcal{P} = \{|\,\ell\,| : \ell \in \mathcal{F}\}$ of seminorms on X. This latter topology is Hausdorff if and only if $\mathcal{F}$ is a separating family, i.e., for any $x \in X$ with $x \ne 0$ there is some $\ell \in \mathcal{F}$ such that $\ell(x) \ne 0$.*

Proof First we observe that if ℓ is a linear functional on X then $|\,\ell\,|$ is a seminorm, and the condition that $\mathcal{F}$ be a separating family of linear functionals is equivalent to $\mathcal{P}$ being a separating family of seminorms.

To show that the vector space topology determined by $\mathcal{P}$ is the same as the $\sigma(X, \mathcal{F})$-topology, we shall show that they have the same convergent nets. To this end, let (x_ν) be a net in X. We have

$$\begin{aligned}
&x_\nu \to x \text{ with respect to the } \sigma(X,\mathcal{F})\text{-topology,}\\
&\quad\iff \ell(x_\nu) \to \ell(x), \text{ for each } \ell \in \mathcal{F},\\
&\quad\iff \ell(x_\nu - x) \to 0, \text{ for each } \ell \in \mathcal{F},\\
&\quad\iff |\ell(x_\nu - x)| \to 0, \text{ for each } \ell \in \mathcal{F},\\
&\quad\iff x_\nu - x \to 0 \text{ with respect to } \mathcal{T}_{\mathcal{P}},\\
&\quad\iff x_\nu \to x \text{ with respect to } \mathcal{T}_{\mathcal{P}}.
\end{aligned}$$

Hence these two topologies have the same convergent nets, the same closed sets and therefore the same open sets. ■

Examples 7.32

1. Equip $C(\mathbb{R})$, the linear space of complex-valued continuous maps on $\mathbb{R}$, with the separating family of seminorms $\mathcal{P} = \{p_K : K \text{ compact in } \mathbb{R}\}$, where p_K is given by $p_K(f) = \sup\{|f(x)| : x \in K\}$ for $f \in C(\mathbb{R})$.

Convergence in $(C(\mathbb{R}), \mathcal{T}_{\mathcal{P}})$ is uniform convergence on compact sets.

2. Denote by (s) the linear space of all complex sequences (x_n) equipped with the family $\mathcal{P}$ of seminorms $\mathcal{P} = \{p_k : k \in \mathbb{N}\}$, where $p_k((x_n)) = |x_k|$. Convergence in $\mathcal{T}_{\mathcal{P}}$ is componentwise convergence.

For each k, let $\ell_k : (s) \to \mathbb{C}$ be the map $\ell_k((x_n)) = x_k$. Then ℓ_k is a linear functional and $\mathcal{T}_{\mathcal{P}}$ is equal to the $\sigma((s), \mathcal{F})$-topology, where $\mathcal{F} = \{\ell_k : k \in \mathbb{N}\}$.

3. We can extend the previous example somewhat. Let X be the linear space of maps f from a given set Ω into $\mathbb{C}$. For each $\omega \in \Omega$, let $p_\omega(f) = f(\omega)$ and let $\mathcal{P} = \{p_\omega : \omega \in \Omega\}$. The topology $\mathcal{T}_{\mathcal{P}}$ on X is the topology of pointwise convergence on Ω. If $\Omega = \mathbb{N}$, then this example reduces to the previous one.

4. The space $\mathcal{S}(\mathbb{R})$ is the space of infinitely differentiable complex-valued functions f on $\mathbb{R}$ such that for each $m, n = 0, 1, 2, \ldots$, the set $\{|x|^m |\, D^n f(x)\,| : x \in \mathbb{R}\}$ is bounded. For such f, set $p_{m,n}(f) = \sup(1 + |x|^m)|D^n f(x)| : x \in \mathbb{R}\}$. Then $\mathcal{P} = \{p_{m,n} : m, n \in \mathbb{N}\}$ is a separating family of seminorms. $\mathcal{S}(\mathbb{R})$ is the Schwartz space of smooth functions of rapid decrease. In a similar way, one defines $\mathcal{S}(\mathbb{R}^n)$, the space of smooth functions on $\mathbb{R}^n$ with rapid decrease. These spaces play an important rôle in the general theory of partial differential equations and also in quantum field theory.

5. Let $X = \mathcal{B}(\mathcal{H})$, the linear space of bounded operators on a Hilbert space $\mathcal{H}$. For each $x \in \mathcal{H}$, let p_x be the seminorm $p_x(T) = \|Tx\|$, for $T \in \mathcal{B}(\mathcal{H})$, and set $\mathcal{P} = \{p_x : x \in \mathcal{H}\}$. Then $\mathcal{P}$ is a separating family and $\mathcal{T}_{\mathcal{P}}$ is the topology of strong operator convergence, $T_\nu \to T$ if and only if $\|T_\nu x - Tx\| \to 0$, for each $x \in \mathcal{H}$.

6. Let $X = \mathcal{B}(\mathcal{H})$, as above. For each $x, y \in \mathcal{H}$, let $\ell_{x,y}$ be the linear functional on X given by $\ell_{x,y}(T) = \langle Tx, y\rangle$ and let $p_{x,y}$ be the seminorm $p_{x,y}(T) = |\ell_{x,y}(T)|$. Let $\mathcal{P}$ be the collection of all such $p_{x,y}$ with $x, y \in \mathcal{H}$ and let $\mathcal{F}$ be the collection of the linear functionals $\ell_{x,y}$. The topology $\mathcal{T}_{\mathcal{P}}$ is equal to the $\sigma(\mathcal{B}(\mathcal{H}), \mathcal{F})$-topology—it is the weak operator topology. A net T_ν converges to T with respect to this topology if and only if $\langle T_\nu x, y\rangle \to \langle Tx, y\rangle$, for each pair $x, y \in \mathcal{H}$.

8. Banach Spaces

In this chapter, we introduce Banach spaces and spaces of linear operators. Recall that if X is a normed space with norm $\|\cdot\|$, then 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 8.1 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 8.2

1. If $\mathbb{R}$ is equipped with the norm $\|\lambda\| = |\lambda|$, $\lambda \in \mathbb{R}$, then it becomes a real Banach 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, with this norm, $\mathbb{R}^n$ becomes a real Banach space (when equipped 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. These norms are the Euclidean norms on $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively.

2. Equip $C([0,1])$, 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 $C([0,1])$ 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 $C([0,1])$ with the norm

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

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

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

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

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

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

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

as $n \to \infty$. Therefore h is equal to 1 on any interval of the form $[\frac{1}{2} + \varepsilon, 1]$, for $0 < \varepsilon < \frac{1}{2}$. 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 $C([0,1])$ 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

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

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.

We have seen, in Theorem 6.23, that there is only one Hausdorff vector space topology on a finite dimensional topological vector space and so, in particular, all norms on such a space are equivalent. We can prove this last part directly as follows.

Theorem 8.3 *Any two norms on a finite dimensional vector space over $\mathbb{K}$ are equivalent.*

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

$$T(t_1, \dots, t_n) = \|t_1 e_1 + \dots + t_n e_n\|.$$

The inequality

$$\big|\, \|t_1 e_1 + \dots + t_n e_n\| - \|s_1 e_1 + \dots + s_n e_n\| \,\big| \le \|(t_1 - s_1)e_1 + \dots + (t_n - s_n)e_n\|$$

shows that T is continuous on $\mathbb{K}^n$. Now, $T(t_1, \dots, t_n) = 0$ only if every $t_i = 0$. In particular, T does not vanish on the unit sphere, $\{z : \|z\| = 1\}$, in $\mathbb{K}^n$. 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

$$m\sqrt{\sum_{k=1}^{n} |t_k|^2} \;\le\; \|t_1 e_1 + \dots + t_n e_n\| \;\le\; M\sqrt{\sum_{k=1}^{n} |t_k|^2}\,.$$

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

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

for every $x \in X$. ∎

Proposition 8.4 *The normed space 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 $\sum_{n=1}^{\infty} \|x_n\| < \infty$. In other words, a normed space is complete if and only if every absolutely convergent series is convergent.*

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 $n > m$,

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

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 $\sum_{k=1}^{m} x_k \to x$ as $m \to \infty$, that is,

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

Hence $y_{n_m} \to x + y_{n_1}$ in X as $m \to \infty$. 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. X/M denotes the set of equivalence classes. The definitions $[x] + [y] = [x + y]$ and $t[x] = [tx]$, for $t \in \mathbb{K}$ 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

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

where the last equality follows because M is a subspace. Thus $\|[x]\|$ 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 itself and zero; $\|z\| = \|z - 0\|$. This suggests that the definition of $\|[x]\|$ above is perhaps a reasonable one.

To see whether this does give a norm or not we consider the various requirements. First, suppose that $t \in \mathbb{K}$, $t \neq 0$, and consider

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

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

Next, we consider the triangle inequality;

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

Clearly, $\|[x]\| \ge 0$ 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, in general, the answer is no, but if M is closed the answer is yes, as the following argument shows.

Proposition 8.5 *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 $-z_n \to x$ in X as $n \to \infty$. Since M is a closed subspace, it follows that $x \in M$ and hence $[x] = 0$ in X/M, as required. ∎

Proposition 8.6 *Let M be a closed linear subspace of a normed space X and let $\pi : X \to X/M$ be the canonical map $\pi(x) = [x]$, $x \in X$. Then π is continuous.*

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

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

and the result follows. ∎

Proposition 8.7 *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 $\sum_{n=1}^k [x_n] \to [y]$ as $k \to \infty$.

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

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

by definition of the infimum. Hence

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

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

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

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

Example 8.8 Let X be the linear space $C([0,1])$ 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 $\phi : X/M \to \mathbb{C}$ 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

$$\begin{aligned}\phi(t[f] + s[g]) &= \phi([tf + sg]) \\ &= tf(0) + sg(0) \\ &= t\phi([f]) + s\phi([g])\end{aligned}$$

for any $t, s \in \mathbb{C}$, and $f, g \in X$. Hence $\phi : X/M \to \mathbb{C}$ 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 $s \in \mathbb{C}$, there is $f \in X$ with $f(0) = s$ and so $\phi([f]) = s$. 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

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

Then $g_n \in M$, for each $n \in \mathbb{N}$, and $g_n \to 1$ with respect to $\|\cdot\|_1$, but $1 \notin M$. The 'quotient norm' is *not* a norm in this case. Indeed, $\|[f]\| = 0$ 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

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

which implies that

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

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

Next we discuss linear mappings between normed spaces. These are also called linear operators.

Definition 8.9 The linear operator $T : X \to Y$ from a normed space X into a normed space Y is said to be bounded if there is some $k > 0$ such that

$$\|Tx\| \le k\|x\|$$

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

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

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

Proposition 8.10 *Suppose that $T : X \to Y$ is a bounded linear operator. Then we have*

$$\begin{aligned}\|T\| &= \sup\{\|Tx\| : \|x\| \le 1\} \\ &= \sup\{\|Tx\| : \|x\| = 1\} \\ &= \sup\Big\{\frac{\|Tx\|}{\|x\|} : x \ne 0\Big\}.\end{aligned}$$

Proof One uses the facts that if $\|x\| \le 1$ (and $x \ne 0$), then $\|Tx\| \le \|Tx\|/\|x\| = \|Tx/\|x\|\|$. ■

Note that if T is bounded, then, by the very definition of $\|T\|$, we have $\|Tx\| \le \|T\|\|x\|$, 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 the ball of radius $\|T\|$ in Y. We have seen in Corollary 6.20 that if T is a continuous linear operator then it is bounded. Hence

$$\|Tx' - Tx\| = \|T(x' - x)\| \le \|T\|\|x' - x\|$$

for any $x, x' \in X$. 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.

Definition 8.11 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 8.12 *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 $s, t \in \mathbb{K}$, the linear operator $sS + tT$ is defined by $(sS + tT)x = sSx + tTx$ for $x \in X$. Furthermore, for any $x \in X$,

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

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 $\|T\| \ge 0$ and $\|T\| = 0$ if $T = 0$. On the other hand, if $\|T\| = 0$, then

$$0 = \|T\| = \sup\Big\{\frac{\|Tx\|}{\|x\|} : x \ne 0\Big\}$$

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

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

$$\begin{aligned}\|sT\| &= \sup\{\|sTx\| : \|x\| \le 1\} \\ &= \sup\{|s|\,\|Tx\| : \|x\| \le 1\} \\ &= |s| \sup\{\|Tx\| : \|x\| \le 1\} \\ &= |s|\,\|T\|\,.\end{aligned}$$

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

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

and the proof is complete. ■

Proposition 8.13 *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

$$\|A_n - A_m\| = \sup\{\|A_n x - A_m x\|/\|x\| : x \ne 0\} \to 0\,,$$

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

$$\begin{aligned}A(sx+x') &= \lim_n A_n(sx+x') \\ &= \lim_n (sA_n x + A_n x') \\ &= s\lim_n A_n x + \lim_n A_n x' \\ &= sAx + Ax'\,,\end{aligned}$$

for any $x, x' \in X$ and $s \in \mathbb{K}$. It follows that $A : X \to Y$ 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 $x \in X$,

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

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

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

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

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

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

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

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

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

Remark 8.14 If S and T belong to $\mathcal{B}(X)$, then $ST : X \to X$ is defined by $STx = S(Tx)$, for any $x \in X$. Clearly ST is a linear operator. Also, $\|STx\| \le \|S\| \|Tx\| \le \|S\| \|T\| \|x\|$, which implies that ST is bounded and $\|ST\| \le \|S\| \|T\|$. 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 8.15

1. Let $A = (a_{ij})$ be any $n \times n$ complex matrix. 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:

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

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

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

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

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

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

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

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

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

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

$$\|Tf\|_\infty \le M\,\|f\|_1$$

and that

$$\|Tf\|_1 \le M\,\|f\|_\infty$$

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

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 $\|Tx\|_1 \le \|x\|_1$. Thus T is a bounded linear operator from $\ell^1 \to \ell^1$, with $\|T\| \le 1$. In fact, $\|T\| = 1$ (take $x = (0,1,0,0,\dots)$). T is called the left shift on ℓ^1.

Similarly, one sees that $T : \ell^\infty \to \ell^\infty$ 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 $\ell^1 \to \ell^1$, 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 8.16 *Suppose that X is a normed space and Y is a Banach space, and suppose that $T : X \to Y$ 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, $\|Tx\| \le \|T\|\|x\|$, for all $x \in D(T)$, where $\|T\|$ is the norm of T as a map $D(T) \to Y$. Let $x \in X$. Since $D(T)$ is dense in X, there is a sequence (ξ_n) in $D(T)$ such that $\xi_n \to x$, in X, as $n \to \infty$. In particular, (ξ_n) is a Cauchy sequence in X. But

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

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 $T\xi_n \to y$ in Y. We would like to construct an extension $\widehat{T}$ of T by defining $\widehat{T}x$ 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 $\eta_n \to x$ in X. Then, as before, we deduce that there is y', say, in Y, such that $T\eta_n \to y'$. 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 $\widehat{T} : X \to Y$ by the prescription $\widehat{T}x = y$, 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 $\widehat{T}x = Tx$, and hence that $\widehat{T}$ is an extension of T. We show that $\widehat{T}$ is a bounded linear operator from X to Y.

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

$$\begin{aligned}
\widehat{T}(sx + x') &= \lim_n T(s\xi_n + {\xi_n}')\,, \qquad \text{using the linearity of } D(T), \\
&= \lim_n sT\xi_n + T{\xi_n}' \\
&= s\widehat{T}x + \widehat{T}x'\,.
\end{aligned}$$

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 $\xi_n \to x$, then, by construction, $T\xi_n \to \widehat{T}x$, and so $\|T\xi_n\| \to \|\widehat{T}x\|$. Hence, the inequalities $\|T\xi_n\| \le \|T\|\,\|\xi_n\|$, for $n \in \mathbb{N}$, imply (—by taking the limit) that

$\|\widehat{T}x\| \le \|T\|\,\|x\|$. Therefore $\|\widehat{T}\| \le \|T\|$. But since $\widehat{T}$ is an extension of T we have that

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

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 8.17 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.

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.

Example 8.18 Let $X = \ell^1$ and $Y = \ell^2$ and, for $k \in \mathbb{N}$, let e_k be the element $e_k = (\delta_{km})_{m\in\mathbb{N}}$ in ℓ^1 and let f_k denote the corresponding element in ℓ^2. For any t with $0 < t < 1$, let $b_t = (t, t^2, t^3, \dots)$. Then $\{e_k : k \in \mathbb{N}\} \cup \{b_t : 0 < t < 1\}$ form an independent set in ℓ^1, and $\{f_k : k \in \mathbb{N}\} \cup \{b_t : 0 < t < 1\}$ form an independent set in ℓ^2. These sets can be extended to Hamel bases B_1 and B_2 of ℓ^1 and ℓ^2, respectively. Both B_1 and B_2 contain a subset of cardinality $2^{\aleph_0}$. On the other hand, $X \subseteq \mathbb{C}^{\mathbb{N}}$ and $Y \subseteq \mathbb{C}^{\mathbb{N}}$ so we deduce that B_1 and B_2 both have cardinality equal to $2^{\aleph_0}$. In particular, there is an isomorphism φ from B_1 onto B_2 which sends e_k into f_k, for each $k \in \mathbb{N}$. By linearity, φ defines a linear isomorphism from ℓ^1 onto 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 $\|b_n\|_2 = 1/\sqrt{n}$.

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 $\|\varphi(x_n) - y\|_2 \to 0$. 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 $|||x_n - x||| \to 0$ as $n \to \infty$. 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 $X = \ell^1$. Indeed, we have that $\varphi(a_n) = b_n$ and so $|||a_n||| = \|b_n\|_2 = 1/\sqrt{n} \to 0$ as $n \to \infty$. However, $\|a_n\|_1 = 1$ for all n.

9. The Dual Space of a Normed Space

Let X be a normed space over $\mathbb{K}$. The space of all bounded linear functionals on X, $\mathcal{B}(X, \mathbb{K})$, is denoted by X^* and called the dual space of X. Since $\mathbb{K}$ 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 from X^* into the field of scalars $\mathbb{K}$. Moreover,

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

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

Theorem 9.1 *Let X be a normed space, and for $x \in X$, let $\varphi_x : X^* \to \mathbb{K}$ be the evaluation map $\varphi_x(\ell) = \ell(x)$, $\ell \in X^*$. Then $x \mapsto \varphi_x$ is an isometric linear mapping 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. We have

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

for all $x, y \in X$, and $t \in \mathbb{K}$. Also

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

and so $\|\varphi_x\| \le \|x\|$. However, by the Hahn-Banach theorem, Corollary 5.24, for any given $x \in X$, $x \ne 0$, 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. ■

Thus, we may consider X as a subspace of X^{**} via the linear isometric embedding $x \mapsto \varphi_x$. This leads to the following observation that any normed space has a completion.

Proposition 9.2 *Any normed space is linearly isometrically isomorphic to a dense subspace of a Banach space.*

Proof The space X^{**} is a Banach space and, as above, X is linearly isometrically isomorphic to the subspace $\{\varphi_x : x \in X\}$. The closure of this subspace is the required Banach space. ■

Definition 9.3 A Banach space X is said to be reflexive if $X = X^{**}$ via the above embedding.

Note that X^{**} is a Banach space, so X must be a Banach space if it is to be reflexive.

Theorem 9.4 *A Banach space X is reflexive if and only if X^* is reflexive.*

Proof If $X = X^{**}$, then $X^* = X^{***}$ via the appropriate embedding. We see this 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^{***},

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

We have $X^* \subseteq X^{***}$ via $\ell \mapsto \psi_\ell$. Let $\lambda \in X^{***}$. Any $z \in X^{**}$ has the form φ_x, for suitable $x \in X$, and so $\lambda(z) = \lambda(\varphi_x)$. Define $\ell : X \to \mathbb{K}$ by $\ell(x) = \lambda(\varphi_x)$. Then

$$\begin{aligned} |\ell(x)| &= |\lambda(\varphi_x)| \\ &\le \|\lambda\| \, \|\varphi_x\| \\ &= \|\lambda\| \, \|x\|. \end{aligned}$$

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

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

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

Now suppose that $X^* = X^{***}$ and suppose that $X \neq X^{**}$. Then, by Corollary 7.30, there is $\lambda \in X^{***}$ such that $\lambda \neq 0$ but λ vanishes on X in X^{**}, i.e., $\lambda(\varphi_x) = 0$ 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

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

which gives $0 = \lambda(\varphi_x) = \ell(x)$ 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 and we conclude that $X = X^{**}$. ∎

Corollary 9.5 *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 Let $X_0 = X$ and, for $n \in \mathbb{N}$, set $X_n = X^*_{n-1}$. If $X_n = X_{n+2}$, then it follows that $X_{n-1} = X_{n+1}$. Repeating this argument, we deduce that $X_0 = X_2$, which is to say that X is reflexive. Note that the equalities here are to be understood as the linear isometric isomorphisms as introduced above. ∎

We shall now compute the duals of some of the classical Banach spaces. For any $p \in \mathbb{R}$ with $1 \le p < \infty$, 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 9.6 *Let a, $b \ge 0$ and α, $\beta > 0$ with $\alpha + \beta = 1$. Then*

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

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 9.7 *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 $y = (y_n) \in \ell^q$,*

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

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

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

Proposition 9.8 *For any $x = (x_n) \in \ell^p$, with $p > 1$,*

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

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

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 $y_n = \overline{\operatorname{sgn} x_n} \, |x_n|^{p/q} / \|x\|^{p/q}$ 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 9.9 *For any $p \ge 1$ and $x, y \in \ell^p$, we have $x + y \in \ell^p$ and*

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

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

Theorem 9.10 *For any $1 \le p \le \infty$, ℓ^p is a Banach space. Moreover, if $1 \le p < \infty$, 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 $1 \le p < \infty$. Let $y_n = \lambda(e_n)$, where $e_n = (\delta_{nm})_{m \in \mathbb{N}} \in \ell^p$. Then for any $x = (x_n) \in \ell^p$,

$$\lambda(x) = \lambda\big(\sum_n x_n e_n\big) = \big(\sum_n x_n \lambda(e_n)\big) = \sum_n x_n y_n \,.$$

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

$$\sum_n |x_n\, y_n| \le \|\lambda\|\, \|x\|_p \,.$$

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

$$\sum_n |x_n\, y_n'| \le \|\lambda\|\, \|x\|_p$$

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

$$\|y'\|_q \le \|\lambda\| \,.$$

It follows that $y \in \ell^q$ (and that $\|y\|_q \le \|\lambda\|$). 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 $1 < p < \infty$. ∎

Now we consider c_0, the linear space of all complex sequences which converge to 0, equipped with the supremum norm

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

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 $\psi_z : c_0 \to \mathbb{C}$ 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

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

Thus we see that ψ_z is a bounded linear functional with norm $\|\psi_z\| \le \|z\|_1$. 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 $\|\psi_z\| \ge \sum_{n=1}^m |z_n|$. 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 $(\delta_{nm})_{m\in\mathbb{N}}$. For any given $N \in \mathbb{N}$, let

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

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

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

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

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

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

Remark 9.11 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 according to the above correspondence.

Theorem 9.12 *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 (real or complex, as appropriate) 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. (Clearly $\overline{S}$ is a closed subspace.) By Corollary 7.30, 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 $\lambda_{n_k} \to \Lambda$ as $k \to \infty$, that is,

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

as $n \to \infty$. However,

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

and so it follows that $\|\lambda_{n_k}\| \to 0$, as $k \to \infty$. But $\lambda_{n_k} \to \Lambda$ implies that $\|\lambda_{n_k}\| \to \|\Lambda\|$ 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 9.13 *For $1 \le p < \infty$ 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 $1 \le p < \infty$.

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 $\|x - \zeta\|_\infty \ge 1$ 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 9.14 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.

Suppose that X is a normed space with dual space X^*. Then, in particular, these are both topological spaces with respect to the topologies induced by their norms. We wish to consider topologies different from these norm topologies.

Definition 9.15 The $\sigma(X, X^*)$-topology on a normed space X is called the weak topology on X.

Thus, the weak topology is the weakest topology with respect to which all bounded linear functionals are continuous. This topology is a locally convex topology determined by the family of seminorms $\mathcal{P} = \{|\ell| : \ell \in X^*\}$. Since X^* separates points of X, a normed space X is a separated topological vector space when equipped with the weak topology. A net (x_ν) in X converges to x with respect to the weak topology if and only if $\ell(x_\nu) \to \ell(x)$ for each $\ell \in X^*$.

Since X^* is a linear space, it follows immediately from Corollary 7.9, that any linear functional on X is continuous with respect to the weak topology on X if and only if it is a member of X^*, i.e., X^* is precisely the set of all weakly continuous linear forms on X. Put another way, this says that every weakly continuous linear form on X is norm continuous. Nonetheless, the weak and norm topologies on X are not (always) equal, as we shall now show.

Proposition 9.16 *Every weak neighbourhood of* 0 *in an infinite dimensional normed space contains a one dimensional subspace.*

Proof Let U be any weak neighbourhood of 0 in the infinite dimensional normed space X. Then there is $r > 0$ and elements $\ell_1, \dots, \ell_n$ in X^*, the dual of X, such that

$$V(0, |\ell_1|, \dots, |\ell_n|; r) \subseteq U.$$

Since X is infinite dimensional, so is X^*. (If X^* were finite dimensional, then X^{**} would also be finite dimensional. However, this is impossible because we have seen that X is linearly isometrically isomorphic to a subspace of X^{**}.) By Proposition 5.16, there is $z \in X$ such that $z \neq 0$ and $\ell_i(z) = 0$ for all $1 \leq i \leq n$. Thus $tz \in U$ for all $t \in \mathbb{K}$. ■

Theorem 9.17 *The weak topology on an infinite dimensional normed space X is strictly weaker than the norm topology.*

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. Consider the 'open' unit ball

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

Then 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 some $z \in X$ with $z \neq 0$ such that $tz \in G$ for all $t \in \mathbb{K}$, by Proposition 9.16. Clearly this is not possible for values of t with $|t| \geq 1/\|z\|$. We conclude that G is not weakly open and the result follows. ■

Remark 9.18 We have shown that no weak neighbourhood of 0 can be norm bounded. Since boundedness is preserved under translations, we conclude that no non-empty weakly open set is 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 $: X^* \to \mathbb{C}$ given by $x : \ell \mapsto \ell(x)$, for $x \in X$ and $\ell \in X^*$—this is the map φ_x defined earlier.

Definition 9.19 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 $X^* \to \mathbb{C}$ as above.

The weak*-topology is also called the w^*-topology on X^*. Since X separates points of X^*, trivially, the w^*-topology is Hausdorff. An open neighbourhood base at 0 for the w^*-topology on X^* is given by

$$\{\lambda \in X^* : |\lambda(x_i) - \ell(x_i)| < r, \quad 1 \le i \le m\}$$

where $r > 0$, $m \in \mathbb{N}$ and $x_1, \ldots, x_m \in X$.

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 as we see next.

Theorem 9.20 *The normed space X is reflexive if and only if the weak and weak* topologies on X^* coincide.*

Proof Suppose that the weak and weak*-topologies on X^* coincide. By Corollary 7.9, X^{**} is precisely the set of all weakly continuous linear forms on X^*. Hence, if the two topologies are equal, every member of X^{**} is weak*-continuous, and so belongs to X. Thus, if φ_x denotes the map $\ell \mapsto \ell(x)$, for $x \in X$, we see that $x \mapsto \varphi_x$ is a map from X onto X^{**} and so X is reflexive.

The converse is clear from the definitions of the topologies. ■

Theorem 9.21 *The weak*-topology on the dual X^* of an infinite dimensional normed space X is strictly weaker than the norm topology on X^*.*

Proof For $x \in X$, the map $X^* \to \mathbb{K}$ given by $\ell \mapsto \ell(x)$ is certainly continuous when X^* is equipped with its norm topology, so the norm topology is stronger than the weak*-topology on X^*. To show that it is strictly stronger when X is infinite dimensional we shall mimic the corresponding result for the weak topology on X. Let $x_1, \ldots, x_n$ be given elements of X and let z be some member of X which is linearly independent of the x_i's. Let M be the linear subspace of X spanned by the elements $z, x_1, \ldots, x_n$ and let λ_z be the linear form on M given by $\lambda_z(z) = 1$ and $\lambda(x_i) = 0$ for $1 \le i \le n$. By Corollary 6.24, the form $\lambda : M \to \mathbb{K}$ is continuous and so, by the Hahn-Banach theorem, Theorem 7.27, has a continuous extension Λ, say, to the whole of X. Then $\Lambda \in X^*$ and $\Lambda(x_i) = 0$ for all $1 \le i \le n$. It follows that if $V = \{\lambda : |\lambda(x_i)| < r,\ 1 \le i \le n\}$ is a basic neighbourhood of 0 in X^*, there is $\Lambda \in X^*$ such $t\Lambda \in V$ for all $t \in \mathbb{K}$. Clearly V is not norm bounded and so we conclude that, for example, the ball $\{\ell \in X^* : \|\ell\| < 1\}$, whilst open in the norm topology is not open in the weak*-topology. (It cannot contain any weak*-open neighbourhood such as V.) The result follows. ■

Proposition 9.22 *The unit ball $\{\ell : \|\ell\| \le 1\}$ in X^* is closed in the w^*-topology.*

Proof Let $\ell_\nu \to \ell$ in the w^*-topology. Then $\ell_\nu(x) \to \ell(x)$, for each $x \in X$. But if $\|\ell_\nu\| \le 1$ for all ν, it follows that $|\ell_\nu(x)| \le \|x\|$ and so $|\ell(x)| \le \|x\|$ for all $x \in X$. That is, $\|\ell\| \le 1$ and we conclude that the unit ball is closed. ■

We will use Tychonov's theorem to show that the unit ball in the dual space of a normed space is actually compact in the w^*-topology. To do this, it is necessary to consider the unit ball of the dual space in terms of a suitable cartesian product. By way of a preamble, let us consider the dual space X^* of the normed space X in such terms. 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 $\mathbb{K}$, equipped with its usual topology. Let $Y = \prod_{x \in X} Y_x = \prod_{x \in X} \mathbb{K}$, 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-coordinate of γ_ℓ is $\ell(x) \in \mathbb{K} = Y_x$.

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 $X^* \to Y$ is one–one. Thus X^* can be thought of as a subset of $Y = \prod_{x \in X} \mathbb{K}$.

Suppose now that $\{\ell_\alpha\}$ is a net in X^* such that $\ell_\alpha \to \ell$ in X^* with respect to the w^*-topology. This is equivalent to the statement that $\ell_\alpha(x) \to \ell(x)$ for each $x \in X$. But then $\ell_\alpha(x) = p_x(\gamma_{\ell_\alpha}) \to \ell(x) = p_x(\gamma_\ell)$ for all $x \in X$, which, in turn, is equivalent to the statement that $\gamma_{\ell_\alpha} \to \gamma_\ell$ with respect to the product topology on $\prod_{x \in X} \mathbb{K}$.

We see, then, that the correspondence $\ell \leftrightsquigarrow \gamma_\ell$ 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 $|\ell(x)| \le \|x\|$. Let B_x denote the ball in $\mathbb{K}$ given by

$$B_x = \{t \in \mathbb{K} : |t| \le \|x\|\} .$$

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, Theorem 3.8, 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,

$$\widehat{Y} = \{\gamma \in Y : \gamma = \gamma_\ell,\ \text{some } \ell \in X_1^*\}.$$

According to the discussion above, we see that a net ℓ_α converges to ℓ in X_1^* with respect to the w^*-topology if and only if γ_{ℓ_α} converges to γ_ℓ in Y. In other words, the correspondence $\ell \leftrightsquigarrow \gamma_\ell$ is a homeomorphism between X_1^* and $\widehat{Y}$ when these are equipped with the induced topologies.

Proposition 9.23 *$\widehat{Y}$ is closed in Y.*

Proof Let (γ_λ) be a net in $\widehat{Y}$ such that $\gamma_\lambda \to \gamma$ in Y. Then $p_x(\gamma_\lambda) \to p_x(\gamma)$ 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

$$p_x(\gamma_{\ell_\lambda}) = \ell_\lambda(x) \to \gamma(x)$$

for each $x \in X_1^*$. It follows that for any $a \in \mathbb{K}$ 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., $|\gamma(x)| \le \|x\|$, 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. ∎

The compactness result we seek is now readily established.

Theorem 9.24 (Banach-Alaoglu) *Let X be a normed space and let X_1^* denote the unit ball in the dual X^*,*

$$X_1^* = \{\ell \in X^* : \|\ell\| \le 1\}.$$

Then X_1^* is a compact subset of X^* with respect to the w^*-topology.

Proof Using the notation established above, we know that Y is compact and since $\widehat{Y}$ is closed in Y, we conclude that $\widehat{Y}$ is also compact. Now, X_1^* and $\widehat{Y}$ are homeomorphic when given the induced topologies so it follows that X_1^* is compact in the induced w^*-topology. However, a subset of a topological space is compact if and only if it is compact with respect to its induced topology and so we conclude that X_1^* is w^*-compact. ∎

Example 9.25 Let $X = \ell^\infty$ and, for each $n \in \mathbb{N}$, let $\ell_m : X \to \mathbb{K}$ 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 $|\ell_m(x)| = |x_m| \le \|x\|_\infty$ 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 it 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 $\ell_{m_k} \to \ell$ in the w^*-topology if and only if $\ell_{m_k}(x) \to \ell(x)$ in $\mathbb{K}$ 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 $(\ell_{m_k}(z))$ cannot converge in $\mathbb{K}$.

10. Fréchet Spaces

In this chapter we will consider complete locally convex topological vector spaces—this class includes, in particular, Banach spaces.

Theorem 10.1 *Let $(X, \mathcal{T})$ be a topological vector space with topology determined by a countable separating family of seminorms. Then $(X, \mathcal{T})$ is metrizable. Moreover, the metric d can be chosen to be translation invariant in the sense that $d(x + a, y + a) = d(x, y)$ for any x, y and a in X.*

Proof Let $\{p_n : n \in \mathbb{N}\}$ be a countable separating family of seminorms which determine the topology $\mathcal{T}$ on X. For any x, y in X, put

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x - y)}{1 + p_n(x - y)} .$$

Then d is well-defined and clearly satisfies $d(x, x) = 0$, $d(x, y) = d(y, x) \geq 0$ and $d(x + a, y + a) = d(x, y)$, $x, y, a \in X$. Now, $p_n(x - z) \leq p_n(x - y) + p_n(y - z)$ and the map $t \mapsto t/(1 + t)$ is increasing on $[0, \infty)$ and so

$$\frac{p_n(x - z)}{1 + p_n(x - z)} \leq \frac{p_n(x - y)}{1 + p_n(x - y)} + \frac{p_n(y - z)}{1 + p_n(y - z)}$$

which implies that $d(x, z) \leq d(x, y) + d(y, z)$, $x, y, z \in X$. Finally, we see that $d(x, y) = 0$ implies that $p_n(x - y) = 0$ for all n, and so $x = y$ since $\{p_n\}$ is a separating family. Thus d is a translation invariant metric on X.

We must show that d induces the vector topology on X determined by the family $\{p_n\}$. To see this, we note that a net (x_ν) in X converges to x with respect to the vector topology $\mathcal{T}$ if and only if $(p_n(x_\nu - x))$ converges to 0 for each $n \in \mathbb{N}$. This is equivalent to $d(x_\nu, x) \to 0$, that is, $x_\nu \to x$ in the metric topology on X induced by d. It follows that the two topologies have the same closed sets and therefore the same open sets, that is, they coincide. ■

Definition 10.2 A sequence (x_n) in a topological vector space $(X, \mathcal{T})$ is said to be a Cauchy sequence if for any neighbourhood U of 0 there is some $N \in \mathbb{N}$ such that $x_n - x_m \in U$ whenever $n \geq m \geq N$.

Proposition 10.3 *Suppose that $(X, \mathcal{T})$ is a topological vector space such that the topology $\mathcal{T}$ is induced by a translation invariant metric d. Then a sequence (x_n) is a Cauchy sequence in $(X, \mathcal{T})$ if and only if it is a Cauchy sequence with respect to the metric d in the usual metric space sense. In particular, if d and d' are translation invariant metrics both inducing the same vector topology on a vector space X, then they have the same Cauchy sequences. Furthermore, (X, d) is complete if and only if (X, d') is complete.*

Proof Suppose that $(X, \mathcal{T})$ is a topological vector space such that $\mathcal{T}$ is given by the translation invariant metric d. Then for any $\varepsilon > 0$, the ball $B_\varepsilon = \{x : d(x, 0) < \varepsilon\}$ is an open neighbourhood of 0, and every neighbourhood of 0 contains such a ball. Given any sequence (x_n) in X, we have $d(x_n - x_m, 0) = d(x_n, x_m)$ so that $x_n - x_m \in B_\varepsilon$ if and only if $d(x_n, x_m) < \varepsilon$ and the first part of the proposition follows.

Now if d and d' are two invariant metrics both inducing the same topology, $\mathcal{T}$, then they clearly have the same Cauchy sequences, namely those sequences which are Cauchy sequences in $(X, \mathcal{T})$.

Suppose that (X, d) is complete, and suppose that (x_n) is a Cauchy sequence in (X, d'). Then (x_n) is a Cauchy sequence in $(X, \mathcal{T})$ and hence also in (X, d). Hence it converges in (X, d) and therefore also in $(X, \mathcal{T})$, to the same limit. But then it also converges in (X, d'), i.e., (X, d') is complete. ∎

This result means that completeness depends only on the topology determined by the family of seminorms and not on the particular family itself. For example, the topology would remain unaltered if every seminorm of the family were replaced by some non-zero multiple of itself, say, p_n were to be replaced by $s_n p_n$, where $s_n > 0$. The translation invariant metric d is changed by this, but the family of Cauchy sequences (and convergent sequences) remains unchanged.

Definition 10.4 A Fréchet space is a topological vector space whose topology is given by a countable separating family of seminorms and such that it is complete as a metric space with respect to the translation invariant metric as defined above via the family of seminorms.

As noted above, completeness does not depend on any particular translation invariant metric which may induce the vector topology.

Example 10.5 A Banach space is a prime example of a Fréchet space; the vector topology is determined by the collection of seminorms consisting of just the norm, and this is separating. The translation invariant metric d induced by the norm, as above, is given by $d(x,y) = \|x-y\|/(1+\|x-y\|)$. Clearly d and the usual metric (also translation invariant) given by the norm, namely, $(x,y) \mapsto \|x-y\|$ have the same Cauchy sequences.

Theorem 10.6 *Let $(X, \mathcal{T})$ be a first countable, Hausdorff locally convex topological vector space. Then there is a countable family $\mathcal{P}$ of seminorms on X such that $\mathcal{T} = \mathcal{T}_{\mathcal{P}}$, the topology determined by $\mathcal{P}$. In particular, $\mathcal{T}$ is induced by a translation invariant metric.*

Proof By hypothesis, there is a countable neighbourhood base at 0. However, any neighbourhood of 0 in a locally convex topological vector space contains a convex balanced neighbourhood so there is a countable neighbourhood base $\{U_n\}_{n\in\mathbb{N}}$ at 0 consisting of convex balanced sets. For each n, let p_n denote the Minkowski functional associated with U_n. The argument of Theorem 7.25 shows that $\mathcal{T} = \mathcal{T}_{\mathcal{P}}$, where $\mathcal{P}$ is the family $\{p_n : n \in \mathbb{N}\}$.

For any $x \in X$ with $x \neq 0$, there is some n such that $x \notin U_n$, since $\mathcal{T}$ is Hausdorff, by hypothesis. Hence $p_n(x) \geq 1$ and so $\mathcal{P}$ is separating.

The topology $\mathcal{T} = \mathcal{T}_{\mathcal{P}}$ is given by the translation invariant metric

$$d(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)},$$

for $x, y \in X$. ∎

Corollary 10.7 *If $(X, \mathcal{T})$ is a metrizable locally convex topological vector space, then $\mathcal{T}$ is determined by a countable family of seminorms on X.*

Proof If $(X, \mathcal{T})$ is metrizable, $\mathcal{T}$ is Hausdorff and first countable. The result now follows from the theorem. ∎

Theorem 10.8 *Let $(X, \mathcal{T})$ be a first countable separated locally convex topological vector space such that every Cauchy sequence (with respect to $\mathcal{T}$) converges. Then $(X, \mathcal{T})$ is a Fréchet space.*

Proof From the above, we see that $(X, \mathcal{T})$ is a topological vector space whose topology is determined by a countable family $\mathcal{P}$ of seminorms. The convergence of Cauchy sequences with respect to $\mathcal{T}$ is equivalent to their convergence with respect to the translation invariant metric d associated with the countable family of seminorms, $\mathcal{P}$. In other words, $(X, \mathcal{T})$ is equal to (X, d), as topological vector spaces, and (X, d) is complete, i.e., $(X, \mathcal{T})$ is a Fréchet space. ∎

Definition 10.9 A subset of a metric space is said to be nowhere dense if its closure has empty interior.

Example 10.10 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 $\overline{S} = \{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$ 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, $\overline{S}$ of S, contains no open ball $B(a;r)$ of positive radius.

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

Theorem 10.11 (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, $X \setminus \bigcup_{n\in\mathbb{N}} \overline{A_n} \subseteq X \setminus \bigcup_{n\in\mathbb{N}} A_n$, and a set is nowhere dense if and only 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 $X \setminus \overline{A_0} \neq \varnothing$. Now, $X \setminus \overline{A_0}$ is open, and non-empty, so there is $x_0 \in X \setminus \overline{A_0}$ and $r_0 > 0$ such that $B(x_0; r_0) \subseteq X \setminus \overline{A_0}$, that is, $B(x_0; r_0) \cap \overline{A_0} = \varnothing$. 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 lead to 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 $B(x_1; r_1) \cap A_1 = \varnothing$.

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 $B(x_2; r_2) \cap A_2 = \varnothing$.

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 $B(x_3; r_3) \cap A_3 = \varnothing$.

Continuing in this way, 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 $B(x_n; r_n) \cap A_n = \varnothing$.

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

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

Hence (x_n) is a Cauchy sequence in X and therefore there is some $x \in X$ such that $x_n \to x$. Since $x_n \in B(x_n; r_n) \subseteq \overline{B(x_N; r_N)}$, 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 $B(x_{N-1}; r_{N-1}) \cap A_{N-1} = \varnothing$. Hence $x \notin A_{N-1}$ for any N. This is our required contradiction and the result follows. ∎

Remark 10.12 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.

Corollary 10.13 *The intersection of any countable family of dense open sets in a complete metric space is dense.*

Proof Suppose that $\{G_n : n \in \mathbb{N}\}$ is a countable family of dense open subsets of a complete metric space. For each n, G_n is open, so its complement, G_n^c, is closed. Also, a set is dense if and only if its complement has no interior. Therefore each G_n^c is closed and nowhere dense. By the theorem, Theorem 10.11, the complement of the union $\bigcup_n G_n^c$ is dense, that is,

$$\bigcap_n G_n = \big(\bigcup_n G_n^c\big)^c$$

is dense, as required. ∎

Example 10.14 For each rational $q \in \mathbb{Q}$, let $F_q = \{q\}$. Then the complement of F_q in $\mathbb{R}$ is a dense open set, and the intersection of all these complements, as q runs over $\mathbb{Q}$, is a countable intersection equal to the set of irrationals and so is dense. On the other hand, if each F_q is considered as a subset of $\mathbb{Q}$, then the intersection of the complements (in $\mathbb{Q}$) is empty—which is evidently not dense in $\mathbb{Q}$.

Definition 10.15 Suppose that X and Y are topological vector spaces and that $\mathcal{E}$ is a collection of linear mappings from X into Y. The collection $\mathcal{E}$ is said to be equicontinuous if to every neighbourhood W of 0 in Y there is some neighbourhood V of 0 in X such that $T(V) \subseteq W$ for all $T \in \mathcal{E}$.

If a family $\mathcal{E}$ of linear maps is equicontinuous, then certainly each member of $\mathcal{E}$ is continuous at 0. However, by linearity of the mappings, this is equivalent to each mapping being continuous. In particular, if $\mathcal{E}$ consists of just a single member T, then equicontinuity of $\mathcal{E} = \{T\}$ is equivalent to the continuity of T. The point of the definition is that the neighbourhood V should not depend on any particular member of $\mathcal{E}$.

In normed spaces, equicontinuous families are the uniformly bounded families as the next result indicates.

Proposition 10.16 *A family $\mathcal{E}$ of linear maps from a normed space X into a normed space Y is equicontinuous if and only if there is $C > 0$ such that*

$$\|Tx\| \leq C\,\|x\|$$

for $x \in X$ and all $T \in \mathcal{E}$.

Proof Suppose that $\mathcal{E}$ is an equicontinuous family of linear maps from the normed space X into the normed space Y. Then, by definition, there is a neighbourhood of 0 in X such that $T(U) \subseteq \{y : \|y\| < 1\}$ for all $T \in \mathcal{E}$. But there is some $r > 0$ such that $\{x : \|x\| < r\} \subseteq U$, and so $\|Tx\| < 1$ whenever $\|x\| < r$, $T \in \mathcal{E}$. Putting $C = 2/r$, it follows that $\|Tx\| \leq C\,\|x\|$ for all $x \in X$ and for all $T \in \mathcal{E}$.

Conversely, suppose that there is $C > 0$ such that $\|Tx\| \leq C\|x\|$, for all $x \in X$ and $T \in \mathcal{E}$. Let W be any neighbourhood of 0 in Y. There is $\varepsilon > 0$ such that $\{y : \|y\|\varepsilon\} \subseteq W$ and so $\|Tx\| < \varepsilon$ for all $T \in \mathcal{E}$ whenever $\|x\| < \varepsilon/C$. That is, if we set $V = \{x : \|x\| < \varepsilon/C\}$, then $T(V) \subseteq W$ for all $T \in \mathcal{E}$. Thus, $\mathcal{E}$ is equicontinuous as claimed. ∎

Proposition 10.17 *Suppose that $\mathcal{E}$ is an equicontinuous family of linear maps from a topological vector space X into a topological vector space Y. For any bounded subset E in X there is a bounded set F in Y such that $T(E) \subseteq F$ for every $T \in \mathcal{E}$.*

Proof Suppose that E is a bounded set in X. Then there is some $s > 0$ such that $E \subset tV$ for all $t > s$. Let $F = \bigcup_{T \in \mathcal{E}} T(E)$. For $t > s$, we have

$$T(E) \subseteq T(tV) = t\,T(V) \subseteq t\,W$$

so that $F \subseteq tW$ and therefore F is bounded. ∎

Theorem 10.18 (Banach-Steinhaus theorem) *Suppose that $\mathcal{E}$ is a collection of continuous linear mappings from a Fréchet space X into a topological vector space Y such that for each $x \in X$ the set*

$$B(x) = \{Tx : T \in \mathcal{E}\}$$

is bounded in Y. Then $\mathcal{E}$ is equicontinuous.

Proof Let W be a neighbourhood of 0 in Y, which, without loss of generality, we may assume to be balanced, and let U be a balanced neighbourhood of 0 such that $U + U + U + U \subseteq W$. Since $\overline{U} \subseteq U + U$, we have that $\overline{U} + \overline{U} \subseteq W$. Let

$$A = \bigcup_{T \in \mathcal{E}} T^{-1}(\overline{U}) \,.$$

For any $x \in X$, $B(x)$ is bounded, by hypothesis, and so there is some $n \in \mathbb{N}$ such that $B(x) \subseteq nU$. It follows that $T((1/n)x) \in U$ for every $T \in \mathcal{E}$ which implies that $(1/n)x \in A$, i.e., $x \in nA$. Hence

$$X = \bigcup_{n \in \mathbb{N}} n\,A \,.$$

Now, every T in $\mathcal{E}$ is continuous and so A is closed and so, therefore, is each nA, $n \in \mathbb{N}$. By the Baire Category theorem, Theorem 10.11, some nA has an interior point; there is some $m \in \mathbb{N}$, $a \in A$ and some open set G with $ma \in G \subseteq mA$. We see that $a \in \frac{1}{m}G \subseteq A$ so that writing V for $\frac{1}{m}G - a$, V is a neighbourhood of 0 such that $V \subseteq A - a$. It follows that

$$T(V) \subseteq T(A) - Ta \subseteq \overline{U} \subseteq W$$

for every $T \in \mathcal{E}$, which shows that $\mathcal{E}$ is equicontinuous. ∎

Theorem 10.19 *Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of continuous linear mappings from a Fréchet space X into a topological vector space Y such that*

$$Tx = \lim_{n \to \infty} T_n x$$

exists for each $x \in X$. Then T is a continuous linear map from X into Y.

Proof It is clear that $T : x \mapsto Tx = \lim_n T_n x$ defines a linear map from X into Y. We must show that T is continuous. Let W be any neighbourhood of 0 in Y and let V be a neighbourhood of 0 such that $V + V \subseteq W$. For each $x \in X$, the sequence $(T_n x)$ converges, so it is bounded. By the Banach Steinhaus theorem, Theorem 10.18, there is a neighbourhood U of 0 in X such that $T_n(U) \subseteq V$ for all n. Now $T_n u \to Tu$ for each $u \in U$, so it follows that $Tu \in \overline{V}$ for $u \in U$. Hence

$$T(U) \subseteq \overline{V} \subseteq V + V \subseteq W$$

and we conclude that T is continuous. ∎

Example 10.20 Let X be the normed space of those sequences (x_k) of complex numbers with only a finite number of non-zero terms equipped with the usual component-wise addition and scalar multiplication and norm $\|(x_k)\| = \sup_k |x_k|$ $(= \max_k |x_k|)$. For each $n \in \mathbb{N}$, let T_n be the linear operator $T_n : X \to X$ with action given by $T_n(x_k) = (a_k x_k)$ where $a_k = \min\{k, n\}$. Clearly, each T_n is continuous $(\|T_n(x_k)\| \le n\|(x_k)\|)$. Moreover, $(T_n(x_k))$ converges, as $n \to \infty$, to $T(x_k)$, where $T : X \to X$ is the linear map $T(x_k) = (kx_k)$. However, the map T is not continuous—for example, if $(x^{(n)})$ is the sequence of elements of X with $x^{(n)}$ the complex sequence whose only non-zero term is the n^{th}, which is equal to $\frac{1}{n}$, then $x^{(n)} \to 0$ in X whereas $\|Tx^{(n)}\| = 1$ for all $n \in \mathbb{N}$.

Theorem 10.21 (Open Mapping theorem) *Suppose that X and Y are Fréchet spaces and that $T : X \to Y$ is a continuous linear mapping from X onto Y. Then T is an open mapping.*

Proof We must show that $T(G)$ is open in Y whenever G is open in X. If G is empty, so is $T(G)$, so suppose that G is a non-empty open set in X. Let $b \in T(G)$ with $b = T(a)$ for some $a \in G$. We need to show that b is an interior point of $T(G)$. By translating G by $-a$ and $T(G)$ by $T(-a) = -b$, this amounts to showing that for any neighbourhood V of 0 in X, 0 is an interior point of $T(V)$. Indeed, since $G - a$ is an open neighbourhood of 0, it would follow that 0 is an interior point of $T(G - a)$, that is, there is an open set W with $0 \in W \subseteq T(G - a)$. Hence $T(a) + W \subseteq T(G)$, which is to say that $b = T(a)$ is an interior point of $T(G)$.

Let d be a translation invariant metric compatible with the topology of X, and let V be any neighbourhood of 0. Then there is some $r > 0$ such that $\{x \in X : d(x, 0) < r\} \subseteq V$. Set

$$V_n = \{x \in X : d(x, 0) < 2^{-n} r\}, \quad n = 0, 1, 2, \ldots.$$

The first step is to show that there is an open neighbourhood W of 0 such that

$$W \subseteq \overline{T(V_1)}.$$

It is easy to see that $V_2 - V_2 \subseteq V_1$, and so

$$T(V_2) - T(V_2) \subseteq T(V_1)$$

and therefore

$$\overline{T(V_2)} - \overline{T(V_2)} \subseteq \overline{T(V_2) - T(V_2)} \subseteq \overline{T(V_1)}$$

using $-\overline{A} = \overline{(-A)}$ and $\overline{A} + \overline{B} \subseteq \overline{A + B}$, for any subsets A, B in a topological vector space. We claim that $\overline{T(V_2)}$ has non-empty interior. To see this, we note that

$$Y = T(X) = \bigcup_{k=1}^{\infty} k\, T(V_2)$$

because $X = \bigcup_{k=1}^{\infty} kV_2$, since V_2 is a neighbourhood of 0 and so is absorbing. By Baire's category theorem, $\overline{k_0 T(V_2)}$ has non-empty interior for some $k_0 \in \mathbb{N}$. But $y \mapsto k_0 y$ is a homeomorphism of Y onto Y so that $\overline{k_0 T(V_2)} = k_0 \overline{T(V_2)}$ and we deduce that $\overline{T(V_2)}$ has non-empty interior, as claimed.

In particular, this means that there is an open set U contained in $\overline{T(V_2)}$. Putting $W = U - U$, we see that W is an open neighbourhood of 0 with

$$W \subseteq \overline{T(V_2)} - \overline{T(V_2)} \subseteq \overline{T(V_1)}$$

which completes the first part of the argument.

Next we shall show that $\overline{T(V_1)} \subseteq T(V)$. To see this, we construct a sequence $(y_n)_{n\in\mathbb{N}}$, with $y_n \in \overline{T(V_n)}$, such that $y_{n+1} - y_n \in T(V_n)$. The y_n are defined recursively. First fix y_1 to be any point of $\overline{T(V_1)}$. Suppose that $n \geq 1$ and that y_n has been chosen in $\overline{T(V_n)}$. Arguing as above, but now with V_{n+1} instead of V_1, we see that $\overline{T(V_{n+1})}$ is a neighbourhood of 0. Hence $y_n - \overline{T(V_{n+1})}$ is a neighbourhood of y_n, and, since y_n belongs to the closure of $\overline{T(V_n)}$, we have

$$\big(y_n - \overline{T(V_{n+1})}\big) \cap T(V_n) \neq \varnothing.$$

Hence there is $x_n \in V_n$ such that $T(x_n) \in y_n - \overline{T(V_{n+1})}$. Set $y_{n+1} = y_n - T(x_n)$. Then $y_{n+1} \in \overline{T(V_{n+1})}$ which completes the construction of the sequence (y_n).

Let $s_n = x_1 + x_2 + \cdots + x_n$, for $n \in \mathbb{N}$. Then for $n > m$,

$$\begin{aligned}
d(s_n, s_m) &= d(s_n - s_m, 0) \\
&= d(x_n + \cdots + x_{m+1}, 0) \\
&\leq d(x_n + \cdots + x_{m+1}, x_{m+1}) + d(x_{m+1}, 0) \\
&< d(x_n + \cdots + x_{m+2}, 0) + \frac{1}{2^{m+1}} \qquad \text{since } x_{m+1} \in V_{m+1}, \\
&< \frac{1}{2^n} + \cdots + \frac{1}{2^{m+1}}.
\end{aligned}$$

Hence (s_n) is a Cauchy sequence in X and therefore converges (because X is complete) to some $x \in X$, with $d(x,0) < r$ (since $d(x,0) \leq d(x,x_1) + d(x_1,0) < d(x,x_1) + r/2$ and $d(s_n,x_1) = d(x_n + \cdots + x_2, 0) \leq r/2$ for all $n > 1$, so that $d(x,x_1) = \lim d(s_n,x_1) \leq r/2$). It follows that $x \in V_0 \subseteq V$. Furthermore,

$$T(s_n) = \sum_{j=1}^{n} T(x_j) = \sum_{j=1}^{n} (y_j - y_{j+1}) = y_1 - y_{n+1}.$$

Now, $y_n \in \overline{T(V_n)}$ and so there is some $v_n \in V_n$ such that $d(y_n, T(v_n)) < \frac{1}{n}$ and hence

$$d(y_n, 0) \leq d(y_n, T(v_n)) + d(T(v_n), 0) < \frac{1}{n} + d(T(v_n), 0).$$

Since $v_n \in V_n$, we have that $d(v_n, 0) < 2^{-n}r$ and so $v_n \to 0$. By the continuity of T (which has not been invoked so far), it follows that $T(v_n) \to 0$, and therefore we see that $y_n \to 0$. Furthermore, $s_n \to x$ implies that $T(s_n) \to T(x)$ so that $y_1 = T(x) \in T(V)$ and we conclude that $\overline{T(V_1)} \subseteq T(V)$.

Combining this with the first part, we see that there is some neighbourhood W of 0 in Y such that $W \subseteq \overline{T(V_1)} \subseteq T(V)$. Thus 0 is an interior point of $T(V)$, and the proof is complete. ■

Corollary 10.22 (Inverse mapping theorem) *Let X and Y be Fréchet spaces and suppose that $T : X \to Y$ is a one-one continuous linear mapping from X onto Y. Then $T^{-1} : Y \to X$ is continuous. In particular, for any given continuous seminorm p on X, there is a continuous seminorm q on Y such that*

$$p(x) \le q(T(x)) \quad \text{for all } x \in X.$$

Proof The inverse T^{-1} exists because T is a bijection, by hypothesis. Write S for T^{-1}, so that $S : Y \to X$. Let G be any set in X. Then

$$\begin{aligned} S^{-1}(G) &= \{y \in Y : S(y) \in G\} = \{y \in Y : T^{-1}(y) \in G\} \\ &= \{y \in Y : y = T(x) \text{ for some } x \in G\} = T(G). \end{aligned}$$

By the theorem, T is open, so that if G is open so is $T(G)$. Therefore T^{-1} is continuous.

Now suppose that p is a continuous seminorm on X. Then $y \mapsto p(S(y))$ is a continuous seminorm on Y. It follows that there is some $C > 0$ and seminorms $q_1, \dots, q_m$ on Y (belonging to any separating family of seminorms which determine the topology on Y) such that

$$p(S(y)) \le C\,(\,q_1(y) + \cdots + q_m(y)\,), \quad y \in Y.$$

Setting $q(y) = C(q_1(y) + \cdots + q_m(y))$, for $y \in Y$, we see that q is a continuous seminorm on Y and, replacing y by $T(x)$, we get

$$p(x) \le q(T(x))$$

for all $x \in X$, as required. ■

Corollary 10.23 *Suppose that X and Y are Banach spaces and that $T : X \to Y$ is a one-one continuous linear map from X onto Y. Then there are positive constants a and b such that*

$$a\|x\| \le \|T(x)\| \le b\|x\|$$

for all $x \in X$.

Proof By the corollary, $S = T^{-1}$ is continuous, and so there is some $C > 0$ such $\|S(y)\| \le C\|y\|$ for all $y \in Y$. Replacing y by $T(x)$ and setting $a = C^{-1}$, it follows that

$$a\|x\| = a\|S(y)\| \le \|y\| = \|T(x)\|, \qquad \text{for } x \in X.$$

The continuity of T implies that $\|T(x)\| \le b\|x\|$ for some positive constant b and all $x \in X$. ■

Corollary 10.24 *Suppose that $\mathcal{T}_1 \subseteq \mathcal{T}_2$ are vector topologies on a vector space X, such that X is a Fréchet space with respect to both. Then $\mathcal{T}_1 = \mathcal{T}_2$.*

Proof The identity map from $(X, \mathcal{T}_2) \to (X, \mathcal{T}_1)$ is a one-one continuous linear map. By the open mapping theorem, it is open, and therefore every $\mathcal{T}_2$ open set is also $\mathcal{T}_1$ open, i.e., $\mathcal{T}_2 \subseteq \mathcal{T}_1$, and so equality holds. ■

If X and Y are vector spaces over $\mathbb{K}$ (either both over $\mathbb{R}$ or both over $\mathbb{C}$) then the Cartesian product $X \times Y$ is a vector space when equipped with the obvious component-wise linear operations, namely $t(x, y) = (tx, ty)$ and $(x, y) + (x', y') = (x + x', y + y')$ for any $t \in \mathbb{K}$, $x, x' \in X$ and $y, y' \in Y$. If X and Y are topological vector spaces, then $X \times Y$ can be equipped with the product topology. It is not difficult to see that this is a vector topology thus making $X \times Y$ into a topological vector space. Indeed, if (x_ν, y_ν) and (x'_ν, y'_ν) are nets in $X \times Y$ converging to (x, y) and (x', y'), respectively, then $x_\nu \to x$ and $x'_\nu \to x'$ in X and $y_\nu \to y$ and $y'_\nu \to y'$ in Y. It follows that $x_\nu + x'_\nu \to x + x'$ and $y_\nu + y'_\nu \to y + y'$ and therefore $(x_\nu, y_\nu) + (x'_\nu, y'_\nu) \to (x + x', y + y')$ in $X \times Y$. Thus addition is continuous in $X \times Y$. In a similar way, one sees that scalar multiplication is continuous.

Now, if X and Y are Fréchet spaces, then so is $X \times Y$. In fact, if $\{p_n\}$ and $\{q_n\}$ are countable families of determining seminorms for X and Y, respectively, then $\{\rho_n\}$ is a determining family for the product topology on $X \times Y$, where ρ_n is given by

$$\rho_n((x, y)) = p_n(x) + q_n(y)$$

for $n \in \mathbb{N}$ and $(x, y) \in X \times Y$. The fact that this family does indeed determine the product topology on $X \times Y$ follows from the equivalence of the following statements; $(x_\nu, y_\nu) \to (x, y)$ in $X \times Y$, $x_\nu \to x$ in X and $y_\nu \to y$ in Y, $p_n(x_\nu - x) \to$

0 and $q_n(y_\nu - y) \to 0$ for every $n \in \mathbb{N}$, $\rho_n((x_\nu, y_\nu) - (x,y)) \to 0$ for every $n \in \mathbb{N}$, $(x_\nu, y_\nu) \to (x,y)$ in the vector topology determined by the family $\{\rho_n\}$.

Let d_X, d_Y and $d_{X\times Y}$ be the translation invariant metrics constructed from the appropriate families of seminorms, i.e.,

$$d_X(x,x') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x-x')}{1+p_n(x-x')}, \quad d_Y(y,y') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(y-y')}{1+p_n(y-y')},$$

and

$$d_{X\times Y}((x,y),(x',y')) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x-x')+q_n(y-y')}{1+p_n(x-x')+q_n(y-y')},$$

for $x, x' \in X$ and $y, y' \in Y$. We have $d_X(x,x') \le d_{X\times Y}((x,y),(x',y'))$ and also $d_Y(y,y') \le d_{X\times Y}((x,y),(x',y'))$, so that if $((x_k,y_k))_{k\in\mathbb{N}}$ is a Cauchy sequence in $X \times Y$, then its components (x_k) and (y_k) are Cauchy sequences in X and Y, respectively, and therefore converge to x and y, say. But then $p_n(x_k) \to p_n(x)$ and $q_n(y_k) \to q_n(y)$ as $k \to \infty$ for each $n \in \mathbb{N}$. This means that $((x_k, y_k))$ converges to (x,y) with respect to the metric $d_{X\times Y}$, i.e., $X \times Y$ is a Fréchet space.

Definition 10.25 Let $T : X \to Y$ be a linear map from the topological vector space X into the topological vector space Y. The graph of T is the subset $\Gamma(T)$ of $X \times Y$ given by

$$\Gamma(T) = \{(x, Tx) : x \in X\}.$$

Theorem 10.26 (Closed Graph theorem) *Suppose that X and Y are Fréchet spaces and $T : X \to Y$ is a linear map from X into Y. Then T is continuous if and only if it has a closed graph in $X \times Y$.*

Proof Suppose that T is continuous and suppose that $(x_n, Tx_n) \to (x,y)$ in $X \times Y$. Then $x_n \to x$ and so $Tx_n \to Tx$. But $Tx_n \to y$ and it follows that $y = Tx$ and therefore $(x,y) = (x,Tx) \in \Gamma(T)$, that is, $\Gamma(T)$ is closed.

Conversely, suppose that $\Gamma(T)$ is closed in $X \times Y$. Then $\Gamma(T)$ is a closed linear subspace of the Fréchet space $X \times Y$ and so is itself a Fréchet space—with respect to the restriction of the metric $d_{x\times Y}$. Let $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ be the projection maps. By definition of the topologies on these spaces, it is clear that both π_X and π_Y are continuous. Moreover, $\pi_X : X \times X$ is one-one and onto X and so, by the inverse mapping theorem, Corollary 10.22, its inverse, π_X^{-1} is continuous from X into $X \times Y$. But then T is given by

$$T : x \overset{\pi_X^{-1}}{\longmapsto} (x, Tx) \overset{\pi_Y}{\longmapsto} Tx,$$

i.e., $T = \pi_Y \circ \pi_X^{-1}$, the composition of two continuous maps, and so is continuous. ∎

Remark 10.27 It is sometimes easier to check that a map T has a closed graph than to check that it is continuous. For the latter, it is necessary to show that if $x_n \to x$, then Tx_n does, in fact, converge and has limit equal to Tx. To show that T has a closed graph, one starts with the hypotheses that both x_n and Tx_n converge, the first to x and the second to some y. All that remains is to show that $Tx = y$.

The following application to operators on a Hilbert space is of interest. It says that a symmetric operator defined on the whole Hilbert space is bounded. This is of interest in quantum mechanics where symmetric operators (or self-adjoint operators, to be precise) are used to represent physically observable quantities. It turns out that these are often unbounded operators. The following theorem says that it is no use trying to define such objects on the whole space. Instead one uses dense linear subspaces as domains of definition for unbounded operators.

Theorem 10.28 (Hellinger-Toeplitz) *Suppose that $T : \mathcal{H} \to \mathcal{H}$ is a linear operator on a Hilbert space $\mathcal{H}$ such that*

$$\langle Tx, y \rangle = \langle x, Ty \rangle$$

for every $x, y \in \mathcal{H}$. Then T is continuous.

Proof A Hilbert space is a Banach space, so is complete. We need only show that T has closed graph. So suppose that $(x_n, Tx_n) \to (x, y)$ in $\mathcal{H} \times \mathcal{H}$. For any $z \in \mathcal{H}$,

$$\langle Tx_n, z \rangle = \langle x_n, Tz \rangle \to \langle x, Tz \rangle = \langle Tx, z \rangle.$$

However, $Tx_n \to y$ and so

$$\langle y, z \rangle = \langle Tx, z \rangle$$

and we deduce that $y = Tx$. It follows that $\Gamma(T)$ is closed and so T is a continuous linear operator on $\mathcal{H}$. ■

We shall end this chapter with a brief discussion of projections. Let X be a linear space and suppose that V and W are subspaces of X such that $V \cap W = \{0\}$ and $X = \operatorname{span}\{V, W\}$. 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 $P : X \to V$ 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 $\operatorname{ran} P = V$ (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 $P : X \to X$ is a linear operator such that $P^2 = P$, that is, P is an idempotent. Set $V = \operatorname{ran} P$ 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

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

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 $Px \in V = \operatorname{ran} P$ 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 topological vector space and that $P : X \to X$ is a continuous linear operator such that $P^2 = P$. Then both $V = \operatorname{ran} P = \ker(\mathbb{1} - P)$ and $W = \ker P$ are closed subspaces of X and $X = V \oplus W$.

Conversely, suppose that X is a Fréchet space and $X = V \oplus W$, where V and W are closed linear subspaces of X. Define $P : X \to V$ as above so that $P^2 = P$ and $V = \operatorname{ran} P = \ker(\mathbb{1} - P)$ and $W = \ker P$. We wish to show that P is continuous. To see this we will show that P is closed and then appeal to the closed-graph theorem. Suppose, then, that $x_n \to x$ and $Px_n \to y$. 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, $(\mathbb{1} - P)x_n = x_n - Px_n \to x - y$ 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 Fréchet space X onto the Fréchet space V. By the closed-graph theorem, it follows that P is continuous. We have therefore proved the following theorem.

Theorem 10.29 *Suppose that V is a closed subspace of a Fréchet space X. Then there is a closed subspace W such that $X = V \oplus W$ if and only if there exists a continuous idempotent P with* $\operatorname{ran} P = V$.

Definition 10.30 We say that a closed subspace V in a topological vector space is complemented if there is a closed subspace W such that $X = V \oplus W$.

Theorem 10.31 *Suppose that V is a finite-dimensional subspace of a topological vector space X. Then V is closed and complemented.*

Proof Let $v_1, \dots, v_m$ be linearly independent elements of X which span V. Define $\ell_i : V \to \mathbb{K}$ by linear extension of the rule $\ell_i(v_j) = \delta_{ij}$ for $1 \le i, j \le m$. By Corollary 6.24, each ℓ_i is a continuous linear functional on V. By the Hahn-Banach theorem, we may extend each of these to continuous linear functionals on X, which we will also denote by ℓ_i. Then, if $v \in V$ is given by $v = t_1 v_1 + \cdots + t_m v_m$, we have $\ell_i(v) = t_i$ and so

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

Define $P : X \to X$ by

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

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