Class Notes for 189-354A.

by

S. W. Drury

Contents

1

Normed and Metric Spaces

We start by introducing the concept of a ***norm*** . This generalization of the absolute value on $\mathbb{R}$ (or $\mathbb{C}$) to the framework of vector spaces is central to modern analysis.

The zero element of a vector space V (over $\mathbb{R}$ or $\mathbb{C}$) will be denoted 0_V. For an element v of the vector space V the norm of v (denoted $\|v\|$) is to be thought of as the distance from 0_V to v, or as the "size" of v. In the case of the absolute value on the field of scalars, there is really only one possible candidate, but in vector spaces of more than one dimension a wealth of possibilities arises.

DEFINITION *A **norm** on a vector space V over $\mathbb{R}$ or $\mathbb{C}$ is a mapping*

$$v \longrightarrow \|v\|$$

from V to $\mathbb{R}^+$ with the following properties.

- $\|0_V\| = 0$.
- $v \in V, \|v\| = 0 \Rightarrow v = 0_V$.
- $\|tv\| = |t|\|v\| \qquad \forall t$ *a scalar*, $v \in V$.
- $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\| \qquad \forall v_1, v_2 \in V$.

The last of these conditions is called the ***subadditivity inequality*** . There are really two definitions here, that of a ***real norm*** applicable to real vector spaces and that of a ***complex norm*** applicable to complex vector spaces. However, every complex vector space can also be considered as a real vector space — one simply "forgets" how to multiply vectors by complex scalars that are not real scalars. This process is called ***realification*** . In such a situation, the two definitions are different. For instance,

$$\|x+iy\| = \max(|x|, 2|y|) \qquad (x, y \in \mathbb{R})$$

defines a perfectly good real norm on $\mathbb{C}$ considered as a real vector space. On the other hand, the only complex norms on $\mathbb{C}$ have the form

$$\|x+iy\| = t(x^2+y^2)^{\frac{1}{2}}$$

for some $t > 0$.

The inequality

$$\|t_1v_1 + t_2v_2 + \cdots + t_nv_n\| \leq |t_1|\|v_1\| + |t_2|\|v_2\| + \cdots + |t_n|\|v_n\|$$

holds for scalars $t_1, \ldots, t_n$ and elements $v_1, \ldots, v_n$ of V. It is an immediate consequence of the definition.

If $\| \;\; \|$ is a norm on V and $t > 0$ then

$$|||v||| = t\|v\|$$

defines a new norm $||| \;\; |||$ on V. We note that in the case of a norm there is often no natural way to normalize it. On the other hand, an absolute value is normalized so that $|1| = 1$, possible since the field of scalars contains a distinguished element 1.

1.1 Some Norms on Euclidean Space

Because of the central role of $\mathbb{R}^n$ as a vector space it is worth looking at some of the norms that are commonly defined on this space.

EXAMPLE On $\mathbb{R}^n$ we may define a norm by

$$\|(x_1, \ldots, x_n)\|_\infty = \max_{j=1}^{n} |x_j|. \tag{1.1}$$

□

EXAMPLE Another norm on $\mathbb{R}^n$ is given by

$$\|(x_1, \ldots, x_n)\|_1 = \sum_{j=1}^{n} |x_j|.$$

□

EXAMPLE The ***Euclidean norm*** on $\mathbb{R}^n$ is given by

$$\|(x_1, \ldots, x_n)\|_2 = \left(\sum_{j=1}^{n} |x_j|^2\right)^{\frac{1}{2}}.$$

This is the standard norm, representing the standard Euclidean distance to $\mathbf{0}$. The symbol $\mathbf{0}$ will be used to denote the zero vector of $\mathbb{R}^n$ or $\mathbb{C}^n$. □

Later we will generalize these examples by defining in case $1 \le p < \infty$

$$\|(x_1, \ldots, x_n)\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}.$$

In case that $p = \infty$ we use (1.1) to define $\| \ \|_\infty$. It will be shown (on page 49) that $\| \ \|_p$ is a norm.

1.2 Inner Product Spaces

Inner product spaces play a very central role in analysis. They have many applications. For example the physics of Quantum Mechanics is based on inner product spaces. In this section we only scratch the surface of the subject.

DEFINITION *A* ***real inner product space*** *is a real vector space V together with an inner product. An* ***inner product*** *is a mapping from $V \times V$ to $\mathbb{R}$ denoted by*

$$(v_1, v_2) \longrightarrow \langle v_1, v_2 \rangle$$

and satisfying the following properties

- $\langle w, t_1 v_1 + t_2 v_2 \rangle = t_1 \langle w, v_1 \rangle + t_2 \langle w, v_2 \rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{R}.$
- $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle \qquad \forall v_1, v_2 \in V.$
- $\langle v, v \rangle \ge 0 \qquad \forall v \in V.$
- *If $v \in V$ and $\langle v, v \rangle = 0$, then $v = 0_V$.*

The symmetry and the linearity in the second variable implies that the inner product is also linear in the first variable.

$$\langle t_1 v_1 + t_2 v_2, w \rangle = t_1 \langle v_1, w \rangle + t_2 \langle v_2, w \rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{R}.$$

EXAMPLE The ***standard inner product*** on $\mathbb{R}^n$ is given by

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j$$

□

The most general inner product on $\mathbb{R}^n$ is given by

$$\langle x, y \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} p_{j,k} x_j y_k$$

where the $n \times n$ real matrix $P = (p_{j,k})$ is a ***positive definite*** matrix. This means that

- P is a symmetric matrix.
- We have
$$\sum_{j=1}^{n} \sum_{k=1}^{n} p_{j,k} x_j x_k \geq 0$$
for every vector $(x_1, \ldots, x_n)$ of $\mathbb{R}^n$.
- The circumstance
$$\sum_{j=1}^{n} \sum_{k=1}^{n} p_{j,k} x_j x_k = 0$$
only occurs when $x_1 = 0, \ldots, x_n = 0$.

In the complex case, the definition is slightly more complicated.

DEFINITION *A* ***complex inner product space*** *is a complex vector space V together with a* ***complex inner product*** *, that is a mapping from $V \times V$ to $\mathbb{C}$ denoted*

$$(v_1, v_2) \longrightarrow \langle v_1, v_2 \rangle$$

and satisfying the following properties

- $\langle w, t_1 v_1 + t_2 v_2 \rangle = t_1 \langle w, v_1 \rangle + t_2 \langle w, v_2 \rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{C}.$
- $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle} \qquad \forall v_1, v_2 \in V.$
- $\langle v, v \rangle \geq 0 \qquad \forall v \in V.$

- *If $v \in V$ and $\langle v, v\rangle = 0$, then $v = 0_V$.*

It will be noted that a complex inner product is linear in its second variable and conjugate linear in its first variable.

$$\langle t_1 v_1 + t_2 v_2, w\rangle = \overline{t_1}\langle v_1, w\rangle + \overline{t_2}\langle v_2, w\rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{C}.$$

EXAMPLE The standard inner product on $\mathbb{C}^n$ is given by

$$\langle x, y\rangle = \sum_{j=1}^{n} \overline{x_j} y_j$$

□

The most general inner product on $\mathbb{C}^n$ is given by

$$\langle x, y\rangle = \sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}\overline{x_j} y_k$$

where the $n \times n$ complex matrix $P = (p_{j,k})$ is a ***positive definite*** matrix. This means that

- P is a hermitian matrix, in other words $p_{jk} = \overline{p_{kj}}$.
- We have
$$\sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}\overline{x_j} x_k \geq 0$$
for every vector $(x_1, \ldots, x_n)$ of $\mathbb{C}^n$.
- The circumstance
$$\sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}\overline{x_j} x_k = 0$$
only occurs when $x_1 = 0, \ldots, x_n = 0$.

DEFINITION *Let V be an inner product space. Then we define*

$$\|v\| = (\langle v, v\rangle)^{\frac{1}{2}} \tag{1.2}$$

the ***associated norm*** .

It is not immediately clear from the definition that the associated norm satisfies the subadditivity condition. Towards this, we establish the abstract Cauchy-Schwarz inequality.

PROPOSITION 1 (CAUCHY-SCHWARZ INEQUALITY) *Let V be an inner product space and $u, v \in V$. Then*

$$|\langle u, v\rangle| \le \|u\|\|v\| \tag{1.3}$$

holds.

Proof of the Cauchy-Schwarz Inequality. We give the proof in the complex case. The proof in the real case is slightly easier. If $v = 0_V$ then the inequality is evident. We therefore assume that $\|v\| > 0$. Similarly, we may assume that $\|u\| > 0$.

Let $t \in \mathbb{C}$. Then we have

$$\begin{aligned} 0 \le \|u + tv\|^2 &= \langle u + tv, u + tv\rangle \\ &= \langle u, u\rangle + \bar{t}\langle v, u\rangle + t\langle u, v\rangle + t\bar{t}\langle v, v\rangle \\ &= \|u\|^2 + 2\Re t\langle u, v\rangle + |t|^2\|v\|^2. \end{aligned} \tag{1.4}$$

Now choose t such that

$$t\langle u, v\rangle \text{ is real and } \le 0 \tag{1.5}$$

and

$$|t| = \frac{\|u\|}{\|v\|}. \tag{1.6}$$

Here, (1.6) designates the absolute value of t and (1.5) specifies its argument. Substituting back into (1.4) we obtain

$$2\frac{\|u\|}{\|v\|}|\langle u, v\rangle| \le \|u\|^2 + \left(\frac{\|u\|}{\|v\|}\right)^2 \|v\|^2$$

which simplifies to the desired inequality (1.3). ■

PROPOSITION 2 *In an inner product space (1.2) defines a norm.*

Proof. We verify the subadditivity of $v \longrightarrow \|v\|$. The other requirements of a norm are straightforward to establish. We have

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \|u\|^2 + \langle v,u\rangle + \langle u,v\rangle + \|v\|^2 \\ &= \|u\|^2 + 2\Re\langle u,v\rangle + \|v\|^2 \\ &\le \|u\|^2 + 2|\Re\langle u,v\rangle| + \|v\|^2 \\ &\le \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \\ &\le \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2 \end{aligned} \tag{1.7}$$

using the Cauchy-Schwarz Inequality (1.3). Taking square roots yields

$$\|u+v\| \le \|u\| + \|v\|$$

as required. ■

1.3 Geometry of Norms

It is possible to understand the concept of norm from the geometrical point of view. Towards this we associate with each norm a geometrical object — its unit ball.

DEFINITION *Let V be a normed vector space. Then the **unit ball** B of V is defined by*

$$B = \{v; v \in V, \|v\| \le 1\}.$$

DEFINITION *Let V be a vector space and let $B \subseteq V$. We say that B is **convex** iff*

$$t_1v_1 + t_2v_2 \in B \qquad \forall v_1, v_2 \in B,\ \forall t_1, t_2 \ge 0 \text{ such that } t_1 + t_2 = 1.$$

In other words, a set B is convex iff whenever we take two points of B, the line segment joining them lies entirely in B.

DEFINITION *Let V be a vector space and let $B \subseteq V$. We say that B satisfies the* ***line condition*** *iff for every $v \in V \setminus \{0_V\}$, there exists a constant $a \in]0, \infty[$ such that*

$$tv \in B \quad \Leftrightarrow \quad |t| \leq a.$$

The line condition says that the intersection of B with every one-dimensional subspace R of V is the unit ball for some norm on R. The line condition involves a multitude of considerations. It implies that the set B is ***symmetric*** about the zero element. The fact that $a > 0$ is sometimes expressed by saying that B is ***absorbing*** . This expresses the fact that every point v of V lies in some (large) multiple of B. Finally the fact that $a < \infty$ is a ***boundedness condition*** .

THEOREM 3 *Let V be a vector space and let $B \subseteq V$. Then the following two statements are equivalent.*

- *There is a norm on V for which B is the unit ball.*
- *B is convex and satisfies the line condition.*

Proof. We assume first that the first statement holds and establish the second. Let $v_1, v_2 \in B$ and let $t_1, t_2 > 0$ be such that $t_1 + t_2 = 1$. Then

$$\begin{aligned} \|t_1 v_1 + t_2 v_2\| &\leq \|t_1 v_1\| + \|t_2 v_2\| \\ &\leq |t_1|\|v_1\| + |t_2|\|v_2\| \\ &\leq t_1 + t_2 = 1. \end{aligned}$$

It follows that B is convex. Now let $v \in V$ and suppose that $v \neq 0_V$. Then it is straightforward to show that the line condition holds with $a = \|v\|^{-1}$.

The real meat of the Theorem is contained in the converse to which we now turn. Let B be a convex subset of V satisfying the line condition. We define for $v \in V \setminus \{0_V\}$

$$\|v\| = a^{-1}$$

where a is the constant of the line condition. We also define $\|0_V\| = 0$. We aim to show that $\| \ \|$ is a norm and that B is its unit ball. Let $v \neq 0_V$ and $s \neq 0$.

Then, applying the line condition to V and sv we have constants a and b with $\|v\| = a^{-1}$ and $\|sv\| = b^{-1}$ such that

$$tv \in B \quad \Leftrightarrow \quad |t| \leq a$$

and

$$r(sv) \in B \quad \Leftrightarrow \quad |r| \leq b.$$

Substituting $t = rs$ we find that

$$|rs| \leq a \quad \Leftrightarrow \quad |r| \leq b$$

so that $a = b|s|$. It now follows that

$$\|sv\| = b^{-1} = |s|a^{-1} = |s|\|v\|. \tag{1.8}$$

On the other hand if $s = 0$ or if $v = 0_V$, then (1.8) also holds.

We turn next to the subadditivity of the norm. Let v_1 and v_2 be non-zero vectors in V. Let $t_1 = \|v_1\|$ and $t_2 = \|v_2\|$. Then $t_1^{-1}v_1 \in B$ and $t_2^{-1}v_2 \in B$. Hence, we find that

$$\begin{aligned} v_1 + v_2 &= t_1 t_1^{-1} v_1 + t_2 t_2^{-1} v_2 \\ &= (t_1 + t_2)\left(\frac{t_1}{t_1 + t_2} t_1^{-1} v_1 + \frac{t_2}{t_1 + t_2} t_2^{-1} v_2\right) \\ &= (t_1 + t_2) v \end{aligned}$$

where $v \in B$ by the convexity of B. If $v_1 + v_2 \neq 0_V$ we have the desired conclusion

$$\|v_1 + v_2\| \leq t_1 + t_2 = \|v_1\| + \|v_2\| \tag{1.9}$$

by the definition of the $\| \ \|$. If $v_1 + v_2 = 0_V$, then (1.9) follows trivially. We also observe that (1.9) follows if either v_1 or v_2 vanishes. The remaining properties of the norm follow directly from the definition.

It is routine to check that for $v \in V \setminus \{0_V\}$

$$\|v\| \leq 1 \quad \Leftrightarrow \quad v \in B.$$

and both sides are true if $v = 0_V$. It follows that B is precisely the unit ball of $\| \ \|$. ∎

EXAMPLE Let us define the a subset B of $\mathbb{R}^2$ by

$$(x,y) \in B \ \text{ if } \begin{cases} x^2+y^2 \le 1 & \text{in case } x \ge 0 \text{ and } y \ge 0, \\ \max(-x,y) \le 1 & \text{in case } x \le 0 \text{ and } y \ge 0, \\ x^2+y^2 \le 1 & \text{in case } x \le 0 \text{ and } y \le 0, \\ \max(x,-y) \le 1 & \text{in case } x \ge 0 \text{ and } y \le 0. \end{cases}$$

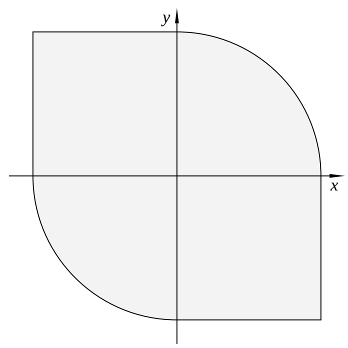

Figure 1.1: The unit ball for a norm on $\mathbb{R}^2$.

It is geometrically obvious that B is a convex subset of $\mathbb{R}^2$ and satisfies the line condition — see Figure 1.1. Therefore it defines a norm. Clearly this norm is given by

$$\|(x,y)\| = \begin{cases} (x^2+y^2)^{\frac{1}{2}} & \text{if } x \ge 0 \text{ and } y \ge 0, \\ \max(|x|,|y|) & \text{if } x \le 0 \text{ and } y \ge 0, \\ (x^2+y^2)^{\frac{1}{2}} & \text{if } x \le 0 \text{ and } y \le 0, \\ \max(|x|,|y|) & \text{if } x \ge 0 \text{ and } y \le 0. \end{cases}$$

□

1.4 Metric Spaces

In the previous section we discussed the concept of the norm of a vector. In a normed vector space, the expression $\|u-v\|$ represents the size of the difference $u-v$ of two vectors u and v. It can be thought of as the distance between u and v. Just as a vector space may have many possible norms, there can be many possible concepts of distance.

In this section we introduce the concept of a ***metric space*** . A metric space is simply a set together with a distance function which measures the distance between any two points of the space. While normed spaces give interesting examples of metric spaces, there are many interesting examples of metric spaces that do not come from norms.

DEFINITION *A **metric space** (X, d) is a set X together with a **distance function** or **metric** $d : X \times X \longrightarrow \mathbb{R}^+$ satisfying the following properties.*

- $d(x, x) = 0 \qquad \forall x \in X.$
- $x, y \in X, d(x, y) = 0 \quad \Rightarrow \quad x = y.$
- $d(x, y) = d(y, x) \qquad \forall x, y \in X.$
- $d(x, z) \le d(x, y) + d(y, z) \qquad \forall x, y, z \in X.$

The fourth axiom for a distance function is called the ***triangle inequality*** . It is easy to derive the ***extended triangle inequality***

$$d(x_1, x_n) \le d(x_1, x_2) + d(x_2, x_3,) + \cdots + d(x_{n-1}, x_n) \quad \forall x_1, \ldots, x_n \in X \tag{1.10}$$

directly from the axioms.

Sometimes we will abuse notation and say that X is a metric space when the intended distance function is understood.

Let X be a metric space and let $Y \subseteq X$. Then the restriction of the distance function of X to the subset $Y \times Y$ of $X \times X$ is a distance function on Y. Sometimes this is called the ***restriction metric*** or the ***relative metric*** . If the four axioms listed above hold for all points of X then *a fortiori* they hold for all points of Y. Thus every subset of a metric space is again a metric space in its own right. This idea will be used very frequently in the sequel.

EXAMPLE Let V be a normed vector space with norm $\|\ \|$. Then V is a metric space with the distance function

$$d(u, v) = \|u - v\|.$$

The reader should check that the triangle inequality is a consequence of the subadditivity of the norm. □

EXAMPLE As an example of an infinite dimensional normed vector space we consider the space ℓ^∞. Its elements are the bounded real sequences (x_n) and the norm is defined by

$$\|(x_n)\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

□

EXAMPLE Another example of an infinite dimensional normed vector space is the space ℓ^1. Its elements are the absolutely summable real sequences (x_n) and the norm is defined by

$$\|(x_n)\|_1 = \sum_{n=1}^{\infty} |x_n|.$$

□

EXAMPLE It follows that every subset X of a normed vector space is a metric space in the distance function induced from the norm. □

EXAMPLE Let $\langle\ ,\ \rangle$ and $\|\ \|$ denote the standard inner product and Euclidean norm on $\mathbb{R}^n$. Let S^{n-1} denote the unit sphere

$$S^{n-1} = \{x; x \in \mathbb{R}^n, \|x\| = 1\}$$

then we can define the ***geodesic distance*** between two points x and y of S^{n-1} by

$$d(x, y) = \arccos(\langle x, y\rangle). \tag{1.11}$$

We will show that d is a metric on S^{n-1}. This metric is of course different from the Euclidean distance $\|x - y\|$.

To verify that (1.11) is in fact a metric, at least the symmetry of the metric is evident. Suppose that $x, y \in S^{n-1}$ and that $d(x, y) = 0$. Then $\langle x, y\rangle = 1$ and

$$\|x - y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2 = 1 - 2 + 1 = 0.$$

It follows that $x = y$.

To establish the triangle inequality, let $x, y, z \in S^{n-1}$, $\theta = \arccos(\langle x, y\rangle)$ and $\varphi = \arccos(\langle y, z\rangle)$. Then we can write $x = \cos\theta\, y + \sin\theta\, u$ and $z = \cos\varphi\, y + \sin\varphi\, v$ where u and v are unit vectors orthogonal to y. An easy calculation now gives

$$\langle x, z\rangle = \cos\theta\ \cos\varphi + \langle u, v\rangle \sin\theta\ \sin\varphi.$$

Now, since $0 \leq \theta, \varphi \leq \pi$, we have $\sin\theta \, \sin\varphi \geq 0$. By the Cauchy-Schwarz Inequality (1.3), we find that $\langle u, v \rangle \geq -1$. Hence

$$\langle x, z \rangle \geq \cos\theta \, \cos\varphi - \sin\theta \, \sin\varphi = \cos(\theta + \varphi).$$

Since arccos is decreasing on $[-1, 1]$ this immediately yields

$$d(x, z) \leq \theta + \varphi = d(x, y) + d(y, z).$$

□

2

Topology of Metric Spaces

2.1 Neighbourhoods and Open Sets

It is customary to refer to the elements of a metric space as ***points*** . In this chapter we will develop the ***point-set topology*** of metric spaces. This is done through concepts such as ***neighbourhoods*** , ***open sets*** , ***closed sets*** and ***sequences*** . Any of these concepts can be used to define more advanced concepts such as the continuity of mappings from one metric space to another. They are, as it were, languages for the further development of the subject. We study them all and most particularly the relationships between them.

DEFINITION *Let (X, d) be a metric space. For $t > 0$ and $x \in X$, we define*

$$U(x, t) = \{y; y \in X, d(x, y) < t\}$$

and

$$B(x, t) = \{y; y \in X, d(x, y) \leq t\}.$$

*the **open ball** $U(x, t)$ centred at x of radius t and the corresponding **closed ball** $B(x, t)$.*

DEFINITION *Let V be a subset of a metric space X and let $x \in V$. Then we say that V is a **neighbourhood** of x or x is an **interior point** of V iff there exists $t > 0$ such that $U(x, t) \subseteq V$.*

Thus V is a neighbourhood of x iff all points sufficiently close to x lie in V.

PROPOSITION 4

- *If V is a neighbourbood of x and $V \subseteq W \subseteq X$. Then W is a neighbourhood of x.*
- *If $V_1, V_2, \ldots, V_n$ are finitely many neighbourhoods of x, then $\cap_{j=1}^n V_j$ is also a neighbourhood of x.*

Proof. The first statement is left as an exercise for the reader. For the second, applying the definition, we may find $t_1, t_2, \ldots, t_n > 0$ such that $U(x, t_j) \subseteq V_j$. It follows that

$$\bigcap_{j=1}^{n} U(x, t_j) \subseteq \bigcap_{j=1}^{n} V_j. \tag{2.1}$$

But the left-hand side of (2.1) is just $U(x, t)$ where $t = \min t_j > 0$. It now follows that $\cap_{j=1}^n V_j$ is a neighbourhood of x. ■

Neighbourhoods are a local concept. We now introduce the corresponding global concept.

DEFINITION *Let (X, d) be a metric space and let $V \subseteq X$. Then V is an **open subset** of X iff V is a neighbourhood of every point x that lies in V.*

EXAMPLE For all $t > 0$, the open ball $U(x, t)$ is an open set. To see this, let $y \in U(x, t)$, that is $d(x, y) < t$. We must show that $U(x, t)$ is a neighbourhood of y. Let $s = t - d(x, y) > 0$. We claim that $U(y, s) \subseteq U(x, t)$. To prove the claim, let $z \in U(y, s)$. Then $d(y, z) < s$. We now find that

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + s = t,$$

so that $z \in U(x, t)$ as required. □

EXAMPLE In $\mathbb{R}$ every interval of the form $]a, b[$ is an open set. Here, a and b are real and satisfy $a < b$. We also allow the possibilities $a = -\infty$ and $b = \infty$. □

THEOREM 5 *In a metric space (X, d) we have*

- *X is an open subset of X.*
- *$\emptyset$ is an open subset of X.*

- *If V_α is open for every α in some index set I, then $\cup_{\alpha\in I} V_\alpha$ is again open.*
- *If V_j is open for $j = 1, \ldots, n$, then the finite intersection $\cap_{j=1}^n V_j$ is again open.*

Proof. For every $x \in X$ and any $t > 0$, we have $U(x, t) \subseteq X$, so X is open. On the other hand, $\emptyset$ is open because it does not have any points. Thus the condition to be checked is vacuous.

To check the third statement, let $x \in \cup_{\alpha\in I} V_\alpha$. Then there exists $\alpha \in I$ such that $x \in V_\alpha$. Since V_α is open, V_α is a neighbourhood of x. The result now follows from the first part of Proposition 4.

Finally let $x \in \cap_{j=1}^n V_j$. Then since V_j is open, it is a neighbourhood of x for $j = 1, \ldots, n$. Now apply the second part of Proposition 4. ∎

DEFINITION *Let X be a set. Let $\mathcal{V}$ be a "family of open sets" satisfying the four conditions of Theorem 5. Then $\mathcal{V}$ is a* ***topology*** *on X and $(X, \mathcal{V})$ is a* ***topological space*** .

Not every topology arises from a metric. In these notes we are *not* concerned with topological spaces in their own right. For some applications topological spaces are needed to capture key ideas (like the weak$\star$ topology). On the other hand, some theorems true for general metric spaces are false for topological spaces (separation theorems for example). Finally some metric space concepts (such as uniform continuity) cannot be defined on topological spaces.

It is worth recording here that there is a complete description of the open subsets of $\mathbb{R}$. A subset V of $\mathbb{R}$ is open iff it is a disjoint union of open intervals (possibly of infinite length). Furthermore, such a union is necessarily countable.

2.2 Convergent Sequences

A sequence $x_1, x_2, x_3, \ldots$ of points of a set X is really a mapping from $\mathbb{N}$ to X. Normally, we denote such a sequence by (x_n). For $x \in X$ the sequence given by $x_n = x$ is called the ***constant sequence*** with value x.

DEFINITION *Let X be a metric space. Let (x_n) be a sequence in X. Then (x_n)* ***converges to*** *$x \in X$ iff for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$*

for all $n > N$. In this case, we write $x_n \longrightarrow x$ or

$$x_n \overset{n\to\infty}{\longrightarrow} x.$$

Sometimes, we say that x is the ***limit*** *of (x_n). Proposition 6 below justifies the use of the indefinite article. To say that (x_n) is a* ***convergent sequence*** *is to say that there exists some $x \in X$ such that (x_n) converges to x.*

EXAMPLE Perhaps the most familiar example of a convergent sequence is the sequence

$$x_n = \frac{1}{n}$$

in $\mathbb{R}$. This sequence converges to 0. To see this, let $\epsilon > 0$ be given. Then choose a natural number N so large that $N > \epsilon^{-1}$. It is easy to see that

$$n > N \qquad \Rightarrow \qquad \left|\frac{1}{n}\right| < \epsilon$$

Hence $x_n \longrightarrow 0$. □

PROPOSITION 6 *Let (x_n) be a convergent sequence in X. Then the limit is unique.*

Proof. Suppose that x and y are both limits of the sequence (x_n). We will show that $x = y$. If not, then $d(x, y) > 0$. Let us choose $\epsilon = \frac{1}{2}d(x, y)$. Then there exist natural numbers N_x and N_y such that

$$\begin{aligned} n > N_x &\qquad \Rightarrow \qquad d(x_n, x) < \epsilon, \\ n > N_y &\qquad \Rightarrow \qquad d(x_n, y) < \epsilon. \end{aligned}$$

Choose now $n = \max(N_x, N_y) + 1$ so that both $n > N_x$ and $n > N_y$. It now follows that

$$2\epsilon = d(x, y) \le d(x, x_n) + d(x_n, y) < \epsilon + \epsilon$$

a contradiction. ■

PROPOSITION 7 *Let X be a metric space and let (x_n) be a sequence in X. Let $x \in X$. The following conditions are equivalent to the convergence of (x_n) to x.*

- *For every neighbourhood V of x in X, there exists $N \in \mathbb{N}$ such that*

$$n > N \qquad \Rightarrow \qquad x_n \in V. \tag{2.2}$$

- *The sequence $(d(x_n, x))$ converges to 0 in $\mathbb{R}$.*

We leave the details of the proof to the reader. The first item here is significant because it leads to the concept of the ***tail*** of a sequence. The sequence (t_n) defined by $t_k = x_{N+k}$ is called the Nth tail sequence of (x_n). The set of points $T_N = \{x_n; n > N\}$ is the Nth tail set. The condition (2.2) can be rewritten as $T_N \subseteq V$.

Sequences provide one of the key tools for understanding metric spaces. They lead naturally to the concept of ***closed subsets*** of a metric space.

DEFINITION *Let X be a metric space. Then a subset $A \subseteq X$ is said to be **closed** iff whenever (x_n) is a sequence in A (that is $x_n \in A \quad \forall n \in \mathbb{N}$) converging to a limit x in X, then $x \in A$.*

The link between closed subsets and open subsets is contained in the following result.

THEOREM 8 *In a metric space X, a subset A is closed if and only if $X \setminus A$ is open.*

It follows from this Theorem that U is open in X iff $X \setminus U$ is closed.

Proof. First suppose that A is closed. We must show that $X \setminus A$ is open. Towards this, let $x \in X \setminus A$. We claim that there exists $\epsilon > 0$ such that $U(x, \epsilon) \subseteq X \setminus A$. Suppose not. Then taking for each $n \in \mathbb{N}$, $\epsilon_n = \frac{1}{n}$ we find that there exists $x_n \in A \cap U(x, \frac{1}{n})$. But now (x_n) is a sequence of elements of A converging to x. Since A is closed $x \in A$. But this is a contradiction.

For the converse assertion, suppose that $X \setminus A$ is open. We will show that A is closed. Let (x_n) be a sequence in A converging to some $x \in X$. If $x \in X \setminus A$ then since $X \setminus A$ is open, there exists $\epsilon > 0$ such that

$$U(x, \epsilon) \subseteq X \setminus A. \tag{2.3}$$

But since (x_n) converges to x, there exists $N \in \mathbb{N}$ such that $x_n \in U(x, \epsilon)$ for $n > N$. Choose $n = N + 1$. Then we find that $x_n \in A \cap U(x, \epsilon)$ which contradicts (2.3). ■

Combining now Theorems 5 and 8 we have the following corollary.

COROLLARY 9 *In a metric space* (X, d) *we have*

- X *is an closed subset of* X.
- $\emptyset$ *is an closed subset of* X.
- *If* A_α *is closed for every* α *in some index set* I, *then* $\cap_{\alpha \in I} A_\alpha$ *is again closed.*
- *If* A_j *is closed for* $j = 1, \ldots, n$, *then the finite union* $\cup_{j=1}^n A_j$ *is again closed.*

EXAMPLE In a metric space every singleton is closed. To see this we remark that a sequence in a singleton is necessarily a constant sequence and hence convergent to its constant value. □

EXAMPLE Combining the previous example with the last assertion of Corollary 9, we see that in a metric space, every finite subset is closed. □

EXAMPLE Let (x_n) be a sequence converging to x. Then the set

$$\{x_n; n \in \mathbb{N}\} \cup \{x\}$$

is a closed subset. □

EXAMPLE In $\mathbb{R}$, the intervals $[a, b]$, $[a, \infty[$ and $]-\infty, b]$ are closed subsets. □

EXAMPLE A more complicated example of a closed subset of $\mathbb{R}$ is the ***Cantor set*** . There are several ways of describing the Cantor set. Let $E_0 = [0, 1]$. To obtain E_1 from E_0 we remove the middle third of E_0. Thus $E_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. To obtain E_2 from E_1 we remove the middle thirds from both the constituent intervals of E_1. Thus

$$E_2 = [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1].$$

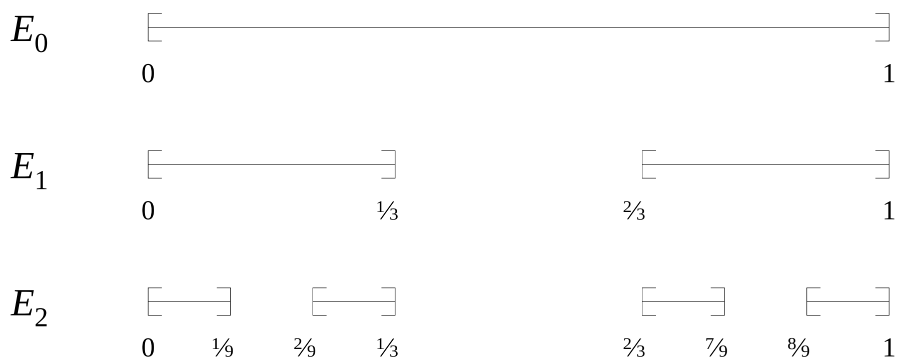

Figure 2.1: The sets E_0, E_1 and E_2.

Continuing in this way, we find that E_k is a union of 2^k closed intervals of length 3^{-k}. The Cantor set E is now defined as

$$E = \bigcap_{k=0}^{\infty} E_k.$$

By Corollary 9 it is clear that E is a closed subset of $\mathbb{R}$.

The sculptor Rodin once said that to make a sculpture one starts with a block of marble and removes everything that is unimportant. This is the approach that we have just taken in building the Cantor set. The second way of constructing the Cantor set works by building the set from the inside out.

Let us define

$$K = \{\sum_{k=1}^{\infty} \omega_k 3^{-k}; \omega_k \in \{0, 2\}, k = 1, 2, \ldots\}.$$

A moment's thought shows us that the points $\sum_{k=1}^{n} \omega_k 3^{-k}$ given by the 2^n choices of ω_k for $k = 1, 2, \ldots, n$ are precisely the left hand endpoints of the 2^n constituent subintervals of E_n. Also a straightforward estimate on the tail sum

$$0 \le \sum_{k=n+1}^{\infty} \omega_k 3^{-k} \le \sum_{k=n+1}^{\infty} 2 \cdot 3^{-k} \le 3^{-n},$$

shows that each sum $\sum_{k=1}^{\infty} \omega_k 3^{-k}$ lies in E_n for each $n \in \mathbb{N}$. It follows that $K \subseteq E$.

For the reverse inclusion, suppose that $x \in E$. Then for every $n \in \mathbb{N}$, let x_n be the left hand endpoint of the subinterval of E_n to which x belongs. Then

$$|x - x_n| \leq 3^{-n}. \tag{2.4}$$

We write

$$x_n = \sum_{k=1}^{n} \omega_k 3^{-k} \tag{2.5}$$

where ω_k takes one or other of the values 0 and 2. It is not difficult to see that the values of ω_k do not depend on the value n under consideration. Indeed, suppose that (2.5) holds for a specific value of n. Then $x \in [x_n, x_n + 3^{-n}]$. At the next step, we look to see whether x lies in the left hand third or the right hand third of this interval. This determines x_{n+1} by

$$x_{n+1} = x_n + \omega_{n+1} 3^{-(n+1)}$$

where $\omega_{n+1} = 0$ if it is the left hand interval and $\omega_{n+1} = 2$ if it is the right hand interval. The values of ω_k for $k = 1, 2, \ldots, n$ are not affected by this choice. It now follows from (2.5) and (2.4) that

$$x = \sum_{k=1}^{\infty} \omega_k 3^{-k}$$

so that $x \in K$ as required. □

2.3 Continuity

The primary purpose of the preceding sections is to define the concept of ***continuity*** of mappings. This concept is the mainspring of mathematical analysis.

DEFINITION *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Let $x \in X$. Then f is **continuous at** x iff for all $\epsilon > 0$, there exists $\delta > 0$ such that*

$$z \in U(x, \delta) \qquad \Rightarrow \qquad f(z) \in U(f(x), \epsilon). \tag{2.6}$$

The $\forall \ldots \exists \ldots$ combination suggests the role of the "devil's advocate" type of argument. Let us illustrate this with an example.

EXAMPLE The mapping $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = x^2$ is continuous at $x = 1$. To prove this, we suppose that the devil's advocate provides us with a number $\epsilon > 0$ chosen cunningly small. We have to "reply" with a number $\delta > 0$ (depending on ϵ) such that (2.6) holds. In the present context, we choose

$$\delta = \min(\tfrac{1}{4}\epsilon, 1)$$

so that for $|x - 1| < \delta$ we have

$$|x^2 - 1| \le |x - 1||x + 1| < (\tfrac{1}{4}\epsilon)(3) < \epsilon$$

since $|x - 1| < \delta$ and $|x + 1| = |(x - 1) + 2| \le |x - 1| + 2 < 3$. □

EXAMPLE Continuity at a point — a single point that is, does not have much strength. Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}, \\ x & \text{if } x \in \mathbb{Q}. \end{cases}$$

This function is continuous at 0 but at no other point of $\mathbb{R}$. □

EXAMPLE An interesting contrast is provided by the function $g : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$g(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \text{ or if } x = 0, \\ \frac{1}{q} & \text{if } x = \frac{p}{q} \text{ where } p \in \mathbb{Z} \setminus \{0\},\ q \in \mathbb{N} \text{ are coprime.} \end{cases}$$

The function g is continuous at x iff x is zero or irrational. To see this, we first observe that if $x \in \mathbb{Q} \setminus \{0\}$, then $g(x) \neq 0$ but there are irrational numbers z as close as we like to x which satisfy $g(z) = 0$. Thus g is not continuous at the points of $\mathbb{Q} \setminus \{0\}$. On the other hand, if $x \in \mathbb{R} \setminus \mathbb{Q}$ or $x = 0$, we can establish continuity of g at x by an epsilon delta argument. We agree that whatever $\epsilon > 0$ we will always choose $\delta < 1$. Then the number of points z in the interval $]x - \delta, x + \delta[$ where $|g(z)| \ge \epsilon$ is finite because such a z is necessarily a rational number that can be expressed in the form $\frac{p}{q}$ where $1 \le q < \epsilon^{-1}$. With only finitely many points to avoid, it is now easy to find $\delta > 0$ such that

$$|z - x| < \delta \qquad \Longrightarrow \qquad |g(z) - g(x)| = |g(z)| < \epsilon.$$

□

There are various other ways of formulating continuity at a point.

THEOREM 10 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Let $x \in X$. Then the following statements are equivalent.*

- *f is continuous at x.*
- *For every neighbourhood V of $f(x)$ in Y, $f^{-1}(V)$ is a neighbourhood of x in X.*
- *For every sequence (x_n) in X converging to x, the sequence $(f(x_n))$ converges to $f(x)$ in Y.*

Proof. We show that the first statement implies the second. Let f be continuous at x and suppose that V is a neighbourhood of $f(x)$ in Y. Then there exists $\epsilon > 0$ such that $U(f(x), \epsilon) \subseteq V$ in Y. By definition of continuity at a point, there exists $\delta > 0$ such that

$$\begin{aligned} z \in U(x,\delta) \quad &\Rightarrow \quad f(z) \in U(f(x),\epsilon) \\ &\Rightarrow \quad f(z) \in V \\ &\Rightarrow \quad z \in f^{-1}(V). \end{aligned}$$

Hence $f^{-1}(V)$ is a neighbourhood of x in X.

Next, we assume the second statement and establish the third. Let (x_n) be a sequence in X converging to x. Let $\epsilon > 0$. Then $U(f(x), \epsilon)$ is a neighbourhood of $f(x)$ in Y. By hypothesis, $f^{-1}(U(f(x), \epsilon))$ is a neighbourhood of x in X. By the first part of Proposition 7 there exists $N \in \mathbb{N}$ such that

$$n > N \quad \Rightarrow \quad x_n \in f^{-1}(U(f(x),\epsilon)).$$

But this is equivalent to

$$n > N \quad \Rightarrow \quad f(x_n) \in U(f(x),\epsilon).$$

Thus $(f(x_n))$ converges to $f(x)$ in Y.

Finally we show that the third statement implies the first. We argue by contradiction. Suppose that f is not continuous at x. Then there exists $\epsilon > 0$ such that for all $\delta > 0$, there exists $z \in X$ with $d(x, z) < \delta$, but $d(f(x), f(z)) \geq \epsilon$. We take choice $\delta = \frac{1}{n}$ for $n = 1, 2, \ldots$ in sequence. We find that there exist x_n in X with $d(x, x_n) < \frac{1}{n}$, but $d(f(x), f(x_n)) \geq \epsilon$. But now, the sequence (x_n) converges to x in X while the sequence $(f(x_n))$ does not converge to $f(x)$ in Y. ■

We next build the global version of continuity from the concept of continuity at a point.

DEFINITION *Let X and Y be metric spaces and let $f : X \longrightarrow Y$. Then the mapping f is* ***continuous*** *iff f is continuous at every point x of X.*

There are also many possible reformulations of global continuity.

THEOREM 11 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Then the following statements are equivalent to the continuity of f.*

- *For every open set U in Y, $f^{-1}(U)$ is open in X.*
- *For every closed set A in Y, $f^{-1}(A)$ is closed in X.*
- *For every convergent sequence (x_n) in X with limit x, the sequence $(f(x_n))$ converges to $f(x)$ in Y.*

Proof. Let f be continuous. We check that the first statement holds. Let $x \in f^{-1}(U)$. Then $f(x) \in U$. Since U is open in Y, U is a neighbourhood of $f(x)$. Hence, by Theorem 10 $f^{-1}(U)$ is a neighbourhood of x. We have just shown that $f^{-1}(U)$ is a neighbourhood of each of its points. Hence $f^{-1}(U)$ is open in X. For the converse, we assume that the first statement holds. Let x be an arbitrary point of X. We must show that f is continuous at x. Again we plan to use Theorem 10. Let V be a neighbourhood of $f(x)$ in Y. Then, there exists $t > 0$ such that $U(f(x), t) \subseteq V$. It is shown on page 15 that $U(f(x), t)$ is an open subset of Y. Hence using the hypothesis, $f^{-1}(U(f(x), t))$ is open in X. Since $x \in f^{-1}(U(f(x), t))$, this set is a neighbourhood of x, and it follows that so is the larger subset $f^{-1}(V)$.

The second statement is clearly equivalent to the first. For instance if A is closed in Y, then $Y \setminus A$ is an open subset. Then

$$X \setminus f^{-1}(A) = f^{-1}(Y \setminus A)$$

is open in X and it follows that $f^{-1}(A)$ is closed in X. The converse entirely similar.

The third statement is equivalent directly from the definition. ■

One very useful condition that implies continuity is the Lipschitz condition.

DEFINITION *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Then f is a* ***Lipschitz map*** *iff there is a constant C with $0 < C < \infty$ such that*

$$d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2) \qquad \forall x_1, x_2 \in X.$$

In the special case that $C = 1$ we say that f is a ***nonexpansive mapping*** *. In the even more restricted case that*

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2) \qquad \forall x_1, x_2 \in X,$$

we say that f is an ***isometry*** *.*

PROPOSITION 12 *Every Lipschitz map is continuous.*

Proof. We work directly. Let $\epsilon > 0$. The set $\delta = C^{-1}\epsilon$. Then $d_X(z, x) < \delta$ implies that

$$d_Y(f(z), f(x)) \leq C d_X(z, x) \leq C\delta = \epsilon.$$

as required. ■

2.4 Compositions of Functions

DEFINITION *Let X, Y and Z be sets. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be mappings. Then we can make a new mapping $h : X \longrightarrow Z$ by $h(x) = g(f(x))$. In other words, to map by h we first map by f from X to Y and then by g from Y to Z. The mapping h is called the* ***composition*** *or* ***composed mapping*** *of f and g. It is usually denoted by $h = g \circ f$.*

Composition occurs in very many situations in mathematics. It is the primary tool for building new mappings out of old.

THEOREM 13 *Let X, Y and Z be metric spaces. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be continuous mappings. Then the composition $g \circ f$ is a continuous mapping from X to Z.*

THEOREM 14 *Let X, Y and Z be metric spaces. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be mappings. Suppose that $x \in X$, that f is continuous at x and that g is continuous at $f(x)$. Then the composition $g \circ f$ is a continuous at x.*

Proof of Theorems 13 and 14. There are many possible ways of proving these results using the tools from Theorem 11 and 10. It is even relatively easy to work directly from the definition.

Let us use sequences. In the local case, we take x as a fixed point of X whereas in the global case we take x to be a generic point of X.

Let (x_n) be a sequence in X convergent to x. Then since f is continuous at x, $(f(x_n))$ converges to $f(x)$. But, then using the fact that g is continuous at $f(x)$, we find that $(g(f(x_n)))$ converges to $g(f(x))$. This says that $(g \circ f(x_n))$ converges to $g \circ f(x)$. Since this holds for every sequence (x_n) convergent to x, it follows that $g \circ f$ is continuous (respectively continuous at x). ■

2.5 Product Spaces and Mappings

In order to discuss combinations of functions we need some additional machinery.

DEFINITION *Let (X, d_X) and (Y, d_Y) be metric spaces. Then we define a **product metric** d on the product set $X \times Y$ which allows us to consider $X \times Y$ as a **product metric space** . We do this as follows*

$$d((x_1, y_1), (x_2, y_2)) = \max(d_X(x_1, x_2), d_Y(y_1, y_2)) \tag{2.7}$$

PROPOSITION 15 *Equation (2.7) defines a bona fide metric on $X \times Y$.*

Proof. The first three conditions in the definition of a metric (on page 11) are obvious. It remains to check the triangle inequality. Let $(x_1, y_1), (x_2, y_2)$ and (x_3, y_3) be three generic points of $X \times Y$. Then $d((x_1, y_1), (x_3, y_3))$ is the maximum of $d_X(x_1, x_3)$ and $d_Y(y_1, y_3)$. Let us suppose without loss of generality that $d_X(x_1, x_3)$ is the larger of the two quantities. Then, by the triangle inequality on X, we have

$$d_X(x_1, x_3) \le d_X(x_1, x_2) + d_X(x_2, x_3). \tag{2.8}$$

But the right hand side of (2.8) is in turn less than

$$d((x_1, y_1), (x_2, y_2)) + d((x_2, y_2), (x_3, y_3))$$

providing the required result. ∎

With the definition out of the way, the next step is to see how it relates to other topological constructs.

LEMMA 16 *Let X and Y be metric spaces. Let $x \in X$ and let (x_n) be a sequence in X. Let $y \in Y$ and let (y_n) be a sequence in Y. Then the sequence $((x_n, y_n))$ converges to (x, y) in $X \times Y$ if and only if the sequence (x_n) converges to x in X and the sequence (y_n) converges to y in Y.*

Proof. First, suppose that $((x_n, y_n))$ converges to (x, y) in $X \times Y$. We must show that (x_n) converges to x in X. (It will follow similarly that (y_n) converges to y in Y.) This amounts then to showing that the projection $\pi : X \times Y \longrightarrow X$ onto the first coordinate, given by

$$\pi((x, y)) = x$$

is continuous. But the definition of the product metric ensures that π is nonexpansive (see page 25) and hence is continuous. The key inequality is

$$d_X(x_1, x_2) \leq d_{X \times Y}((x_1, y_1), (x_2, y_2)).$$

For the converse, we have to get our hands dirtier. Let $\epsilon > 0$. Then there exists N such that $d_X(x_n, x) < \epsilon$ for $n > N$. Also, there exists M such that $d_Y(y_n, y) < \epsilon$ for $n > M$. It follows that for $n > \max(N, M)$ both of the above inequalities hold, so that

$$\max(d_X(x_n, x), d_Y(y_n, y)) < \epsilon.$$

But this is exactly equivalent to

$$d_{X \times Y}((x_n, y_n), (x, y)) < \epsilon$$

as required for the convergence of $((x_n, y_n))$ to (x, y). ∎

There is a simple way to understand neighbourhoods and hence open sets in product spaces.

PROPOSITION 17 *Let X and Y be metric spaces. Let $x \in X$ and $y \in Y$. Let $U \subseteq X \times Y$. Then the following two statements are equivalent*

- *U is a neighbourhood of (x, y).*
- *There exist V a neighbourhood of x and W a neighbourhood of y such that $V \times W \subseteq U$.*

Proof. Suppose that the first statement holds. Then there exists $t > 0$ such that $U_{X \times Y}((x, y), t) \subseteq U$. But it is easy to check that

$$U_{X \times Y}((x, y), t) = U_X(x, t) \times U_Y(y, t).$$

Of course, $U_X(x, t)$ is a neighbourhood of x in X and $U_Y(y, t)$ is a neighbourhood of y in Y.

Conversely, let V and W be neighbourhoods of x and y in X and Y respectively. Then there exist $t, s > 0$ such that $U_X(x, t) \subseteq V$ and $U_Y(y, s) \subseteq W$. It is then easy to verify that

$$U_{X \times Y}((x, y), \min(t, s)) \subseteq U_X(x, t) \times U_Y(y, s) \subseteq V \times W \subseteq U,$$

so that U is a neighbourhood of (x, y) as required. ■

Next, we introduce ***product mappings*** .

DEFINITION *Let X, Y, P and Q be sets. Let $f : X \longrightarrow P$ and $g : Y \longrightarrow Q$. Then we define the* ***product mapping*** *$f \times g : X \times Y \longrightarrow P \times Q$ by*

$$(f \times g)(x, y) = (f(x), g(y)).$$

PROPOSITION 18 *Let X, Y, P and Q be metric spaces. Let $f : X \longrightarrow P$ and $g : Y \longrightarrow Q$ be continuous mappings. Then the product mapping $f \times g$ is also continuous.*

Proof. We argue using sequences. We could equally well use neighbourhoods or epsilons and deltas. Let $((x_n, y_n))$ be an arbitrary sequence in $X \times Y$ converging to (x, y). Then (x_n) converges to x in X by Lemma 16. By Theorem 11 we find that $(f(x_n))$ converges to $f(x)$. Similar reasoning shows that $(g(y_n))$ converges to $g(y)$. Now we use Lemma 16 again to show that $((f(x_n), g(y_n)))$ converges to $(f(x), g(y))$. Finally since $((x_n, y_n))$ is an arbitrary sequence in $X \times Y$ converging to (x, y), it follows again by Theorem 16 that $f \times g$ is continuous. ■

There is also a local version of Proposition 18. We leave both the statement and the proof to the reader.

2.6 The Diagonal Mapping and Pointwise Combinations

DEFINITION *Let X be a metric space. Then the **diagonal mapping** on X is the mapping $\Delta_X : X \longrightarrow X \times X$ given by*

$$\Delta_X(x) = (x, x) \qquad \forall x \in X.$$

If X is a metric space it is easy to check that Δ_X is an isometry (for the definition, see page 25). In particular, Δ_X is a continuous mapping. This gives us the missing link to discuss the continuity of pointwise combinations.

THEOREM 19 *Let P, Q and R be metric spaces. Let $\mu : P \times Q \longrightarrow R$ be a continuous mapping. Let $f : X \longrightarrow P$ and $g : X \longrightarrow Q$ also be continuous mappings. Then the combination $h : X \longrightarrow R$ given by*

$$h(x) = \mu(f(x), g(x)) \qquad \forall x \in X$$

is also continuous.

Proof. It suffices to write $h = \mu \circ (f \times g) \circ \Delta_X$ and to apply Theorem 13 and Proposition 18 together with the continuity of Δ_X. ■

There are numerous examples of Theorem 19. In effect, the examples that follow are examples of continuous binary operations.

EXAMPLE Let $P = Q = R = \mathbb{R}$. Let $\mu(x, y) = x + y$, addition on $\mathbb{R}$. Then if $f, g : X \longrightarrow \mathbb{R}$ are continuous so is the sum function $f + g$ defined by

$$(f + g)(x) = f(x) + g(x) \qquad \forall x \in X.$$

It remains to check the continuity of μ. We have

$$\begin{aligned}|\mu(x_1,y_1)-\mu(x_2,y_2)| &= |(x_1-x_2)+(y_1-y_2)| \\ &\le |x_1-x_2|+|y_1-y_2| \\ &\le d_{\mathbb{R}\times\mathbb{R}}((x_1,y_1),(x_2,y_2))+d_{\mathbb{R}\times\mathbb{R}}((x_1,y_1),(x_2,y_2)) \\ &= 2d_{\mathbb{R}\times\mathbb{R}}((x_1,y_1),(x_2,y_2)),\end{aligned}$$

so that μ is Lipschitz with constant $C=2$ and hence continuous. □

EXAMPLE Let $P=Q=R=\mathbb{R}$. Let $\mu(x,y)=xy$, multiplication on $\mathbb{R}$. Then if $f,g:X\longrightarrow\mathbb{R}$ are continuous so is the pointwise product function fg defined by

$$(fg)(x)=f(x)g(x) \qquad \forall x\in X.$$

We check that μ is continuous at (x_1,y_1). Observe that

$$xy-x_1y_1=x_1(y-y_1)+(x-x_1)y_1+(x-x_1)(y-y_1)$$

so that

$$|xy-x_1y_1|\le|x_1||y-y_1|+|x-x_1||y_1|+|x-x_1||y-y_1|$$

Now let $\epsilon>0$ be given. We choose $\delta=\min(1,(|x_1|+|y_1|+1)^{-1}\epsilon)$. Then

$$d_{\mathbb{R}\times\mathbb{R}}((x,y),(x_1,y_1))<\delta$$

implies that

$$\begin{aligned}|xy-x_1y_1| &< |x_1|\delta+\delta|y_1|+\delta^2 \\ &\le (|x_1|+|y_1|+1)\delta \\ &\le \epsilon.\end{aligned}$$

This estimate establishes that μ is continuous at (x_1,y_1). □

We leave the reader to check that addition and multiplication are continuous operations in $\mathbb{C}$. Two other operations on $\mathbb{R}$ that are continuous are max and min. We leave the reader to show that these are distance decreasing.

EXAMPLE One very important binary operation on a metric space is the distance function itself. Let X be a metric space, $P = Q = X$ and $R = \mathbb{R}^+$. Let $\mu(x, y) = d(x, y)$. We check that μ is continuous. By the extended triangle inequality (page 11) we have

$$\begin{aligned} d(x_2, y_2) &\le d(x_2, x_1) + d(x_1, y_1) + d(y_1, y_2) \\ &\le d(x_1, y_1) + 2d_{X\times X}((x_1, y_1), (x_2, y_2)), \end{aligned}$$

and similarly

$$d(x_1, y_1) \le d(x_2, y_2) + 2d_{X\times X}((x_1, y_1), (x_2, y_2)).$$

We may combine these two inequalities into one as

$$|d(x_1, y_1) - d(x_2, y_2)| \le 2d_{X\times X}((x_1, y_1), (x_2, y_2)).$$

This shows that the distance function is Lipschitz with constant $C = 2$, and hence is continuous. □

Other examples of continuous binary operations are found in the context of normed spaces. Let us recall that in a normed space $(V, \| \ \|)$, the metric d_V is given by

$$d_V(v_1, v_2) = \|v_1 - v_2\| \qquad \forall v_1, v_2 \in V.$$

We will treat only the case of real normed spaces. The complex case is similar.

EXAMPLE In a normed space $(V, \| \ \|)$, the vector addition operator is continuous. Let $\mu(v, w) = v + w$. We have

$$\begin{aligned} \|\mu(v_1, w_1) - \mu(v_2, w_2)\| &= \|(v_1 - v_2) + (w_1 - w_2)\| \\ &\le \|v_1 - v_2\| + \|w_1 - w_2\| \\ &\le d_{V\times V}((v_1, w_1), (v_2, w_2)) + d_{V\times V}((v_1, w_1), (v_2, w_2)) \\ &= 2d_{V\times V}((v_1, w_1), (v_2, w_2)), \end{aligned}$$

so that μ is Lipschitz with constant $C = 2$. □

While the previous example parallelled addition in $\mathbb{R}$, the next is similar to multiplication in $\mathbb{R}$.

EXAMPLE In a normed space $(V, \| \ \|)$, the scalar multiplication operator is continuous. Thus $P = \mathbb{R}$, $Q = R = V$, and $\mu : \mathbb{R}\times V \longrightarrow V$ is the map $\mu(t, v) = tv$. We leave the details to the reader. □

EXAMPLE Now let V be a real inner product space. Then the inner product is continuous. Thus $P = Q = V$, $R = \mathbb{R}$ and $\mu(v, w) = \langle v, w \rangle$.

We check that μ is continuous at (v_1, w_1). Observe that

$$\langle v, w \rangle - \langle v_1, w_1 \rangle = \langle v_1, w - w_1 \rangle + \langle v - v_1, w_1 \rangle + \langle v - v_1, w - w_1 \rangle$$

so that by the Cauchy-Schwarz inequality (page 6) we have

$$|\langle v, w \rangle - \langle v_1, w_1 \rangle| \leq \|v_1\| \|w - w_1\| + \|v - v_1\| \|w_1\| + \|v - v_1\| \|w - w_1\|.$$

Now let $\epsilon > 0$ be given. We choose $\delta = \min(1, (\|v_1\| + \|w_1\| + 1)^{-1}\epsilon)$. Then

$$d_{V \times V}((v, w), (v_1, w_1)) < \delta$$

implies that

$$\begin{aligned} |\langle v, w \rangle - \langle v_1, w_1 \rangle| &< \|v_1\|\delta + \delta\|w_1\| + \delta^2 \\ &\leq (\|v_1\| + \|w_1\| + 1)\delta \\ &\leq \epsilon. \end{aligned}$$

This estimate establishes that μ is continuous at (v_1, w_1). □

2.7 Interior and Closure

We return to discuss subsets and sequences in metric spaces in greater detail. Let X be a metric space and let A be an arbitrary subset of X. Then $\emptyset$ is an open subset of X contained in A, so we can define the ***interior*** $\text{int}(A)$ of A by

$$\text{int}(A) = \bigcup_{U \text{ open } \subseteq A} U. \tag{2.9}$$

By Theorem 5 (page 16), we see that $\text{int}(A)$ is itself an open subset of X contained in A. Thus $\text{int}(A)$ is the unique open subset of X contained in A which in turn contains all open subsets of X contained in A. There is a simple characterization of $\text{int}(A)$ in terms of interior points (page 14).

PROPOSITION 20 *Let X be a metric space and let $A \subseteq X$. Then*

$$\text{int}(A) = \{x; x \textit{ is an interior point of } A\}.$$

Proof. Let $x \in \text{int}(A)$. Then since $\text{int}(A)$ is open, it is a neighbourhood of x. But then the (possibly) larger set A is also a neighbourhood of x. This just says that x is an interior point of A.

For the converse, let x be an interior point of A. Then by definition, there exists $t > 0$ such that $U(x,t) \subseteq A$. But it is shown on page 15, that $U(x,t)$ is open. Thus $U = U(x,t)$ figures in the union in (2.9), and since $x \in U(x,t)$ it follows that $x \in \text{int}(A)$. ■

EXAMPLE The interior of the closed interval $[a,b]$ of $\mathbb{R}$ is just $]a,b[$. □

EXAMPLE The Cantor set E has empty interior in $\mathbb{R}$. Suppose not. Let x be an interior point of E. Then there exist $\epsilon > 0$ such that $U(x,\epsilon) \subseteq E$. Choose now n so large that $3^{-n} < \epsilon$. Then we also have $U(x,\epsilon) \subseteq E_n$. For the notation see page 20. This says that E_n contains an open interval of length $2(3^{-n})$ which is clearly not the case. □

By passing to the complement and using Theorem 8 (page 18) we see that there is a unique closed subset of X containing A which is contained in every closed subset of X which contains A. The formal definition is

$$\text{cl}(A) = \bigcap_{E \text{ closed } \supseteq A} E. \tag{2.10}$$

The set $\text{cl}(A)$ is called the ***closure*** of A. We would like to have a simple characterization of the closure.

PROPOSITION 21 *Let X be a metric space and let $A \subseteq X$. Let $x \in X$. Then $x \in \text{cl}(A)$ is equivalent to the existence of a sequence of points (x_n) in A converging to x.*

Proof. Let $x \in \text{cl}(A)$. Then x is not in $\text{int}(X \setminus A)$. Then by Proposition 20, x is not an interior point of $X \setminus A$. Then, for each $n \in \mathbb{N}$, there must be a point $x_n \in A \cap U(x, \frac{1}{n})$. But now, $x_n \in A$ and (x_n) converges to x.

For the converse, let (x_n) be a sequence of points of A converging to x. Then $x_n \in \text{cl}(A)$ and since $\text{cl}(A)$ is closed, it follows from the definition of a closed set that $x \in \text{cl}(A)$. ■

While Proposition 21 is perfectly satisfactory for many purposes, there is a subtle variant that is sometimes necessary.

DEFINITION *Let X be a metric space and let $A \subseteq X$. Let $x \in X$. Then x is an* ***accumulation point*** *or a* ***limit point*** *of A iff $x \in \mathrm{cl}(A \setminus \{x\})$.*

PROPOSITION 22 *Let X be a metric space and let $A \subseteq X$. Let $x \in X$. Then the following statements are equivalent.*

- $x \in \mathrm{cl}(A)$.
- *$x \in A$ or x is an accumulation point of A.*

Proof. That the second statement implies the first follows easily from Proposition 21. We establish the converse. Let $x \in \mathrm{cl}(A)$. We may suppose that $x \notin A$, for else we are done. Now apply the argument of Proposition 21 again. For each $n \in \mathbb{N}$, there is a point $x_n \in A \cap U(x, \frac{1}{n})$. Since $x \notin A$, we have $A = A \setminus \{x\}$. Thus we have found $x_n \in A \setminus \{x\}$ with (x_n) converging to x. ■

DEFINITION *Let X be a metric space and let $A \subseteq X$. Let $x \in A$. Then x is an* ***isolated point*** *of A iff there exists $t > 0$ such that $A \cap U(x, t) = \{x\}$.*

We leave the reader to check that a point of A is an isolated point of A if and only if it is not an accumulation point of A.

A very important concept related to closure is the concept of density.

DEFINITION *Let X be a metric space and let $A \subseteq X$. Then A is said to be* ***dense*** *in X if $\mathrm{cl}(A) = X$.*

If A is dense in X, then by definition, for every $x \in X$ there exists a sequence (x_n) in A converging to x.

PROPOSITION 23 *Let f and g be continuous mappings from X to Y. Suppose that A is a dense subset of X and that $f(x) = g(x)$ for all $x \in A$. Then $f(x) = g(x)$ for all $x \in X$.*

Proof. Let $x \in X$ and let (x_n) be a sequence in A converging to x. Then $f(x_n) = g(x_n)$ for all $n \in \mathbb{N}$. So the sequences $(f(x_n))$ and $(g(x_n))$ which converge to $f(x)$ and $g(x)$ respectively, are in fact identical. By the uniqueness of the limit, Proposition 6 (page 17), it follows that $f(x) = g(x)$. This holds for all $x \in X$ so that $f = g$. ■

We leave the proof of the following Proposition to the reader.

PROPOSITION 24 *Let A be a dense subset of a metric space X and let B be a dense subset of a metric space Y. Then $A \times B$ is dense in $X \times Y$.*

2.8 Limits in Metric Spaces

DEFINITION *Let X be a metric space and let $t > 0$. Then for $x \in X$ the **deleted open ball** $U'(x, t)$ is defined by*

$$U'(x,t) = \{z; z \in X, 0 < d(x,z) < t\} = U(x,t) \setminus \{x\}.$$

Let A be a subset of X then it is routine to check that x is an accumulation point of A if and only if for all $t > 0$, $U'(x,t) \cap A \neq \emptyset$. Deleted open balls are also used to define the concept of a ***limit*** .

DEFINITION *Let X and Y be metric spaces. Let x be an accumulation point of X. Let $f : X \setminus \{x\} \longrightarrow Y$. Then $f(z)$ has **limit** y **as** z **tends to** x **in** X, in symbols*

$$\lim_{z \to x} f(z) = y \tag{2.11}$$

if and only if for all $\epsilon > 0$ there exists $\delta > 0$ such that

$$z \in U'(x,\delta) \Longrightarrow f(z) \in U(y,\epsilon).$$

*In the same way one also defines $f(z)$ has **a limit as** z **tends to** x **in** X, which simply means that (2.11) holds for some $y \in Y$.*

Note that in the above definition, the quantity $f(x)$ is undefined. The purpose of taking the limit is to "attach a value" to $f(x)$. The following Lemma connects this idea with the concept of continuity at a point. We leave the proof to the reader.

LEMMA 25 *Let X and Y be metric spaces. Let x be an accumulation point of X. Let $f : X \setminus \{x\} \longrightarrow Y$. Suppose that (2.11) holds for some $y \in Y$. Now define $\tilde{f} : X \longrightarrow Y$ by*

$$\tilde{f}(z) = \begin{cases} f(z) & \text{if } z \in X \setminus \{x\}, \\ y & \text{if } z = x. \end{cases}$$

Then $\tilde{f}$ is continuous at x.

One of the most standard uses of limits is in the definition of the derivative.

DEFINITION *Let $g :]a, b[\longrightarrow V$ where V may as well be a general normed vector space. Let $t \in]a, b[$. Then the quotient*

$$f(s) = (s - t)^{-1}(g(s) - g(t)) \in V$$

is defined for s in $]a, b[\setminus \{t\}$. It is not defined at $s = t$. If

$$\lim_{s \to t} f(s)$$

exists, then we say that g is ***differentiable*** *at t and the value of the limit is denoted $g'(t)$ and called the* ***derivative*** *of g at t. It is an element of V.*

2.9 Distance to a Subset

DEFINITION *Let X be a metric space and let A be a non-empty subset of X. Then we may define for every element $x \in X$, the real number $\mathrm{dist}_A(x) \geq 0$ by*

$$\mathrm{dist}_A(x) = \inf_{a \in A} d(x, a).$$

This is the distance from x to the subset A. We view dist_A as a mapping $\mathrm{dist}_A : X \longrightarrow \mathbb{R}^+$.

PROPOSITION 26 *Let X be a metric space and let $A \subseteq X$. Then*

- $\mathrm{dist}_A : X \longrightarrow \mathbb{R}^+$ *is continuous.*
- $\mathrm{dist}_A(x) = 0 \Leftrightarrow x \in \mathrm{cl}(A)$.
- $\mathrm{dist}_A(x) = \mathrm{dist}_{\mathrm{cl}(A)}(x) \qquad \forall x \in X$.

Proof. Let $x_1, x_2 \in X$ and $a \in A$. Then by the triangle inequality

$$d(x_1, a) \leq d(x_1, x_2) + d(x_2, a).$$

Take infimums of both sides as a runs over the elements of A to obtain

$$\mathrm{dist}_A(x_1) \leq d(x_1, x_2) + \mathrm{dist}_A(x_2). \tag{2.12}$$

An exactly similar argument yields

$$\operatorname{dist}_A(x_2) \le d(x_1, x_2) + \operatorname{dist}_A(x_1). \tag{2.13}$$

Now we combine (2.12) and (2.13) to find that

$$|\operatorname{dist}_A(x_1) - \operatorname{dist}_A(x_2)| \le d(x_1, x_2), \tag{2.14}$$

which asserts that dist_A is nonexpansive. The first assertion follows.

The second assertion follows directly from the definition of $\operatorname{cl}(A)$.

For the third assertion, it is clear that $\operatorname{dist}_{\operatorname{cl}(A)}(x) \le \operatorname{dist}_A(x)$ since $\operatorname{cl}(A)$ is a (possibly) larger set than A. It therefore remains to show that $\operatorname{dist}_{\operatorname{cl}(A)}(x) \ge \operatorname{dist}_A(x)$. By the definition of $\operatorname{dist}_{\operatorname{cl}(A)}(x)$, it suffices to take a an arbitrary point of $\operatorname{cl}(A)$ and show that

$$\operatorname{dist}_A(x) \le d(a, x). \tag{2.15}$$

Using the fact that $a \in \operatorname{cl}(A)$, we see that there is a sequence (a_n) of points of A converging to a. By definition of $\operatorname{dist}_A(x)$ we have

$$\operatorname{dist}_A(x) \le d(a_n, x) \tag{2.16}$$

But since d is a continuous function on $X \times X$, it now follows that $d(a_n, x) \longrightarrow d(a, x)$ as $n \longrightarrow \infty$. Combining this with (2.16) yields (2.15) as required. ■

2.10 Separability

In this text, we use the term ***countable*** to mean finite or countably infinite. Thus a set A is countable iff it can be put in one to one correspondence with some subset of $\mathbb{N}$.

DEFINITION *A metric space X is said to be **separable** iff it has a countable dense subset.*

EXAMPLE The real line $\mathbb{R}$ is a separable metric space with the standard metric because the set $\mathbb{Q}$ of rational numbers is dense in $\mathbb{R}$. □

The nomenclature is somewhat misleading. Separability has nothing to do with separation. In fact separability is a measure of the smallness of a metric space. Unfortunately this fact is not obvious. The following Theorem clarifies the situation.

THEOREM 27 *Let X be a separable metric space. Let Y be a subset of X. Then Y is separable when considered as a metric space with the restriction metric.*

Proof. Let A be a countable dense subset of X. Then it is certainly possible that $A \cap Y = \emptyset$. We need therefore to build a subset of Y in a more complicated way. Let (a_n) be an enumeration of A. By the definition of $\mathrm{dist}_Y(a_n)$ we can deduce the existence of an element $b_{n,k}$ of Y such that

$$d(a_n, b_{n,k}) < \mathrm{dist}_Y(a_n) + \tfrac{1}{k}. \tag{2.17}$$

We will show that the set $\{b_{n,k}; n, k \in \mathbb{N}\}$ is dense in Y. Towards this, let $y \in Y$. We will show that for every $\epsilon > 0$ there exist n and k such that $b_{n,k} \in U(y, \epsilon)$. We choose n such that $d(a_n, y) < \frac{1}{3}\epsilon$, possible because A is dense in X. We choose k so large that $\frac{1}{k} < \frac{1}{3}\epsilon$. It follows that

$$\begin{aligned} d(b_{n,k}, y) &\le d(b_{n,k}, a_n) + d(a_n, y) \\ &\le \mathrm{dist}_Y(a_n) + \tfrac{1}{k} + d(a_n, y) \\ &\le d(a_n, y) + \tfrac{1}{k} + d(a_n, y) \\ &< \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon = \epsilon. \end{aligned}$$

as required. ■

Much easier is the following Theorem the proof of which we leave as an exercise.

THEOREM 28 *Let X and Y be separable metric spaces. Then $X \times Y$ is again a separable metric space (with the product metric).*

The following result is needed in applications to measure theory.

THEOREM 29 *In a separable metric space X, every open subset U can be written as a countable union of open balls.*

Proof. We leave the reader to prove the Theorem in case that $U = X$ and assume henceforth that $U \ne X$. By Theorem 27, the set U itself possesses a countable dense subset. Let us enumerate this subset as (x_n). We claim that

$$U = \bigcup_n U(x_n, \tfrac{1}{2}\,\mathrm{dist}_{X \setminus U}(x_n)). \tag{2.18}$$

Obviously, the right hand side of (2.18) is contained in the left hand side. To establish the claim, we let $x \in U$ and show that x is in the right hand member of (2.18). Let $t = \operatorname{dist}_{X \setminus U}(x) > 0$ because of Proposition 26 and since $X \setminus U$ is a closed set. Using the density of (x_n), we may find $n \in \mathbb{N}$ such that

$$d(x, x_n) < \tfrac{1}{3}t. \tag{2.19}$$

By (2.14), we have that

$$|\operatorname{dist}_{X\setminus U}(x) - \operatorname{dist}_{X\setminus U}(x_n)| \le d(x, x_n), \tag{2.20}$$

and it follows from (2.19), (2.20) and the definition of t that

$$\operatorname{dist}_{X\setminus U}(x_n) > \tfrac{2}{3}t.$$

Then we have

$$d(x, x_n) < \tfrac{1}{3}t = \tfrac{1}{2}(\tfrac{2}{3}t) < \tfrac{1}{2}\operatorname{dist}_{X\setminus U}(x_n).$$

It follows that $x \in U(x_n, \frac{1}{2}\operatorname{dist}_{X\setminus U}(x_n))$ as required. ■

EXAMPLE Every subset of $\mathbb{R}^d$ is separable. □

EXAMPLE The normed vector space ℓ^∞ (page 12) is not separable. To see this, suppose that S is a dense subset of ℓ^∞. Let $\omega = (\omega_n)$ be a sequence taking the values ± 1. There are uncountably many such sequences ω. For each such sequence, there is a sequence $s = (s_n)$ in S such that $\|s - \omega\| < \frac{1}{3}$. It is easy to see that two distinct values of ω necessarily lead to distinct elements of S. It follows that S is also uncountable. □

EXAMPLE On the other hand, the space ℓ^1 (page 12) is separable. Let S be the set of sequences with rational entries eventually zero. Then S is a countable set. Given a sequence $x = (x_n)$ in ℓ^1 and a strictly positive real number ϵ, we first choose N so large that

$$\sum_{n>N} |x_n| < \frac{\epsilon}{2}.$$

Let $y = (y_n)$ be the truncated sequence given by

$$y_n = \begin{cases} x_n & \text{if } n \le N, \\ 0 & \text{if } n > N. \end{cases}$$

Then $\|x - y\|_1 < \frac{1}{2}\epsilon$. It now remains to find a slightly perturbed sequence $s = (s_n) \in S$ such that $\|y - s\|_1 < \frac{1}{2}\epsilon$. We leave this as an exercise. For more on this example see Proposition 32 □

2.11 Relative Topologies

We remarked on page 11 that if X is a metric space and Y is a subset of X then Y can be considered as a metric space in its own right. From the point of view of convergent sequences, this causes no problems. The sequences in Y that converge in Y to an element of Y are simply the sequences in Y that converge in X to an element of Y. Of course, it is possible to have a sequence of elements of Y which converges in X to an element of $X \setminus Y$. Such a sequence will not converge in Y.

The situation with regard to open and closed sets is more complicated, and certainly more difficult to understand. A subset A of Y can be said to be open in Y or said to be open in X. These concepts are different in general. To distinguish the difference, we sometimes say that A is ***relatively open*** when it is an open subset of the subset Y. In general the adverb ***relatively*** is reserved for properties considered with respect to the subset (in this case Y) rather than the whole space (in this case X). Thus when we say that A is ***relatively closed*** , we mean that it is closed in Y. If A is ***relatively dense*** , then it is dense in Y.

Let us consider an example to illustrate the difference.

EXAMPLE Let $X = \mathbb{R}$ with the usual metric and $Y = [0, 1]$ with the relative metric. Then the subset $A = [0, \frac{1}{2}[$ of Y is not open in X because $0 \in A$ and every neighbourhood of 0 in X contains small negative numbers that are not in A. However 0 *is* an interior point of A with respect to Y. This is because $U_Y(0, \epsilon) = [0, \epsilon[\subseteq A$ provided $0 < \epsilon < \frac{1}{2}$. Those small negative numbers are not in Y and do not cause a problem when we are considering openness in Y. The reader should ponder this point until he understands it, because it is fundamental to so much that follows. In fact the subset A is open relative to Y. □

EXAMPLE Let $X = \mathbb{R}$ with the usual metric and $Y = [0, 1[$ with the relative metric. Then the subset $A = [\frac{1}{2}, 1[$ is not closed in X, but it *is* closed in Y. The skeptic will immediately consider the sequence $(x_n = \frac{n}{n+1})$ which lies in A and "converges to 1". This is certainly true in X, but it is not true that $x_n \longrightarrow 1$ in Y for the simple reason that $1 \notin Y$. □

What is required is a way of understanding the open subsets of Y in terms of those of X. The following result fills that role.

THEOREM 30 *Let X be a metric space and let $Y \subseteq X$.*

- *A subset U of Y is open in Y iff there exists an open subset V of X such that $U = V \cap Y$.*

- *A subset F of Y is closed in Y iff there exists a closed subset E of X such that $F = E \cap Y$.*

Proof. We work on the first statement. Let U be a subset of Y open in Y. By definition, for every $y \in U$ there exists $t_y > 0$ such that $U_Y(y, t_y) \subseteq U$. Now define

$$V = \bigcup_{y \in U} U_X(y, t_y).$$

Then V is an open subset of X by Theorem 5 (page 16). We have

$$\begin{aligned} V \cap Y &= \bigcup_{y \in U} (U_X(y, t_y) \cap Y) \\ &= \bigcup_{y \in U} U_Y(y, t_y) \\ &= U, \end{aligned}$$

since, for every $y \in U$, we have $y \in U_X(y, t_y)$.

Conversely, if V is open in X and $y \in V \cap Y$, then there exists $t > 0$ such that $U_X(y, t) \subseteq V$. Then obviously

$$U_Y(y, t) = U_X(y, t) \cap Y \subseteq V \cap Y.$$

Thus $V \cap Y$ is a neighbourhood of each of its points in Y. In other words $V \cap Y$ is open in Y. This completes the proof of the first assertion. The second assertion follows immediately from the first and Theorem 8 (page 18). ■

EXAMPLE Consider $Y = \mathbb{R}$ embedded as the real axis in $X = \mathbb{R}^2$. The interval $]-1, 1[$ is a relatively open subset of the real axis Y. It is clearly not an open subset of $\mathbb{R}^2$. However, the disc

$$\{(x, y); x^2 + y^2 < 1\}$$

is open in the plane X and meets the real axis Y in precisely $]-1, 1[$. □

COROLLARY 31 *We maintain the notations of the Theorem. Thus X is a metric space, Y is a subset of X which we are considering as a metric space in its own right. Further U and F are subsets of Y*

- *If U is open in X, then it is open in Y.*
- *If Y is open in X and U is open in Y, then U is open in X.*
- *If F is closed in X, then it is closed in Y.*
- *If Y is closed in X and F is closed in Y, then F is closed in X.*

We can use relative topologies to elucidate the proof of the fact that the sequence space ℓ^1 is separable on page 39. Here there are three spaces $X = \ell^1$, Y the set of all real sequences that are eventually zero, and S the set of all rational sequences that are eventually zero. We have $S \subset Y \subset X$. We show that Y is dense in X and that S is relatively dense in Y. The density of S in X then follows from the following general principle which might be called the ***transitivity of density*** .

PROPOSITION 32 *Let X be a metric space and let $S \subseteq Y \subseteq X$. Suppose that Y is dense in X and that S is relatively dense in Y. Then S is dense in X.*

Proof. Let $\epsilon > 0$ and suppose that $x \in X$. Then, since Y is dense in X there exists $y \in Y$ such that $d(x, y) < \frac{1}{2}\epsilon$. Now, since S is dense in Y, there exists $s \in S$ such that $d(y, s) < \frac{1}{2}\epsilon$. The triangle inequality now yields $d(x, s) < \epsilon$ as required. ■

THEOREM 33 (GLUEING THEOREM) *Let X and Y be metric spaces. Let X_1 and X_2 be subsets of X such that $X = X_1 \cup X_2$. Let $f_j : X_j \longrightarrow Y$ be continuous maps for $j = 1, 2$. Suppose that f_1 and f_2 agree on their overlap — explicitly*

$$f_1(x) = f_2(x) \qquad \forall x \in X_1 \cap X_2,$$

so that the ***glued mapping*** $f : X \longrightarrow Y$ *given by*

$$f(x) = \begin{cases} f_1(x) & \text{if } x \in X_1, \\ f_2(x) & \text{if } x \in X_2, \end{cases}$$

is well defined. Suppose that one or other of the two following conditions holds.

- *Both X_1 and X_2 are open in X.*

- *Both X_1 and X_2 are closed in X.*

Then f is continuous.

Proof. Suppose that both X_1 and X_2 are open in X. We work with sequences. Let $x \in X$ and suppose that (x_n) is a sequence in X converging to x. Without loss of generality we may suppose that $x \in X_1$. Then since X_1 is open in X, the sequence (x_n) is eventually in X_1. Explicitly, there exists $N \in \mathbb{N}$ such that $x_n \in X_1$ for $n > N$. Since this tail of the sequence converges to x in X_1 and since f_1 is continuous as a mapping from X_1 to Y, the image sequence of the tail converges to $f_1(x)$. But this just says that $(f(x_n))$ converges to $f(x)$.

Let us go back and fill in the details in glorious technicolour. We define a new sequence (the tail) by $z_k = x_{N+k}$. We claim that z_k converges to x. Towards this, let $\epsilon > 0$. Then since (x_n) converges to x, there exists $M \in \mathbb{N}$ such that $d(x, x_n) < \epsilon$ for $n > M$. Then, certainly $d(x, z_k) < \epsilon$ for $k > M$. This proves the claim. Since for all k, $z_k \in X_1$ and since f_1 is continuous on X_1 we now find that $(f(z_k))$ converges to $f(x)$ in X. Now we claim that $(f(x_n))$ converges to $f(x)$. Let $\epsilon > 0$. Then there exists $K \in \mathbb{N}$, such that $k > K$ implies $d(f(z_k), f(x)) < \epsilon$. Then, for $n > N + K$ we have $d(f(x_n), f(x)) = d(f(z_k), f(x)) < \epsilon$ where $k = n - N > K$ as needed.

In case that X_1 and X_2 are both closed in X we use a completely different strategy, namely the characterization of continuity by closed subsets in Theorem 11 (page 24). Let A be a closed subset of Y. We must show that $f^{-1}(A)$ is closed in X. We write $f^{-1}(A) = (f^{-1}(A) \cap X_1) \cup (f^{-1}(A) \cap X_2)$ possible since $X = X_1 \cup X_2$. It is enough to show that the two sets $f^{-1}(A) \cap X_1$ and $f^{-1}(A) \cap X_2$ are closed in X. Without loss of generality we need only handle the first of these. Now $f^{-1}(A) \cap X_1 = f_1^{-1}(A)$, so that, by the continuity of f_1 this set is closed in X_1. Therefore, according to the last assertion of Corollary 31, it is also closed in X since X_1 is itself closed in X. ■

EXAMPLE The Glueing Theorem is used in homotopy theory. Let f and g be continuous maps from a metric space X to a metric space Y. Then we say that f and g are ***homotopic*** iff there exist a continuous map

$$F : [0,1] \times X \longrightarrow Y$$

such that

$$F(0, x) = f(x) \qquad \forall x \in X$$

and

$$F(1,x) = g(x) \qquad \forall x \in X.$$

It turns out that being homotopic is an equivalence relation. We leave the reflexivity and symmetry conditions to be verified by the reader. We now sketch the transitivity.

Let g and h also be homotopic. Then there is (with slight change in notation) a continuous mapping

$$G : [1,2] \times X \longrightarrow Y$$

such that

$$G(1,x) = g(x) \qquad \forall x \in X$$

and

$$G(2,x) = h(x) \qquad \forall x \in X.$$

Since the subsets $[0,1] \times X$ and $[1,2] \times X$ are closed in $[0,2] \times X$, the mappings F and G can be glued together to make a continuous mapping

$$H : [0,2] \times X \longrightarrow Y$$

such that

$$H(0,x) = f(x) \qquad \forall x \in X$$

and

$$H(2,x) = h(x) \qquad \forall x \in X.$$

It follows that f and h are homotopic. □

2.12 Uniform Continuity

For many purposes, continuity of mappings is not enough. The following strong form of continuity is often needed.

DEFINITION *Let X and Y be metric spaces and let $f : X \longrightarrow Y$. Then we say that f is* ***uniformly continuous*** *iff for all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$x_1, x_2 \in X, d_X(x_1, x_2) < \delta \qquad \Rightarrow \qquad d_Y(f(x_1), f(x_2)) < \epsilon. \quad (2.21)$$

In the definition of continuity, the number δ is allowed to depend on the point x_1 as well as ϵ.

EXAMPLE The function $f(x) = x^2$ is continuous, but not uniformly continuous as a mapping $f : \mathbb{R} \longrightarrow \mathbb{R}$. Certainly the identity mapping $x \longrightarrow x$ is continuous because it is an isometry. So f, which is the pointwise product of the identity mapping with itself is also continuous. We now show that f is not uniformly continuous. Let us take $\epsilon = 1$. Then, we must show that for all $\delta > 0$ there exist points x_1 and x_2 with $|x_1 - x_2| < \delta$, but $|x_1^2 - x_2^2| \geq 1$. Let us take $x_2 = x - \frac{1}{4}\delta$ and $x_1 = x + \frac{1}{4}\delta$. Then

$$x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2) = x\delta.$$

It remains to choose $x = \delta^{-1}$ to complete the argument. □

EXAMPLE Any function satisfying a Lipschitz condition (page 25) is uniformly continuous. Let X and Y be metric spaces. Let $f : X \longrightarrow Y$ with constant C. Then

$$d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2) \qquad \forall x_1, x_2 \in X.$$

Given $\epsilon > 0$ it suffices to choose $\delta = C^{-1}\epsilon > 0$ in order for $d_X(x_1, x_2) < \delta$ to imply $d_Y(f(x_1), f(x_2)) < \epsilon$. □

It should be noted that one cannot determine (in general) if a mapping is uniformly continuous from a knowledge only of the open subsets of X and Y. Thus, uniform continuity is not a topological property. It depends upon other aspects of the metrics involved.

In order to clarify the concept of uniform continuity and for other purposes, one introduces the ***modulus of continuity*** ω_f of a function f. Suppose that $f : X \longrightarrow Y$. Then $\omega_f(t)$ is defined for $t \geq 0$ by

$$\omega_f(t) = \sup\{d_Y(f(x_1), f(x_2)); x_1, x_2 \in X, d_X(x_1, x_2) \leq t\}. \qquad (2.22)$$

It is easy to see that the uniform continuity of f is equivalent to

$$\forall \epsilon > 0,\ \exists \delta > 0 \text{ such that } 0 < t < \delta \quad \Rightarrow \quad \omega_f(t) < \epsilon.$$

We observe that $\omega_f(0) = 0$ and regard $\omega_f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$. Then the uniform continuity of f is also equivalent to the continuity of ω_f at 0.

2.13 Subsequences

Subsequences are used extensively in analysis. Some advanced metric space concepts such as compactness can be handled quite nicely using subsequences. We start by defining a subsequence of the sequence of natural numbers.

DEFINITION *A sequence (n_k) of natural numbers is called a **natural subsequence** if $n_k < n_{k+1}$ for all $k \in \mathbb{N}$.*

Since $n_1 \geq 1$, a straightforward induction argument yields that $n_k \geq k$ for all $k \in \mathbb{N}$.

DEFINITION *Let (x_n) be a sequence of elements of a set X. A **subsequence** of (x_n) is a sequence (y_k) of elements of X given by*

$$y_k = x_{n_k}$$

where (n_k) is a natural subsequence.

The key result about subsequences is very easy and is left as an exercise for the reader.

LEMMA 34 *Let (x_n) be a sequence in a metric space X converging to an element $x \in X$. Then any subsequence (x_{n_k}) also converges to x.*

One way of showing that a sequence fails to converge is to find two convergent subsequences with different limits. Indeed, this idea can also be turned around. One way of showing that two sequences converge to the same limit is to build a new sequence that possesses both of the given sequences as subsequences. It is then enough to establish the convergence of the new sequence. This idea will be used in our discussion of completeness.

3

A Metric Space Miscellany

In this chapter we introduce some topics from metric spaces that are slightly out of the mainstream and which can be tackled with the rather meagre knowledge of the subject that we have amassed up to this point. This chapter is primarily intended to enrich the material presented thus far.

3.1 The p-norms on $\mathbb{R}^n$

Let $1 \le p < \infty$. We define

$$\|(x_1, \ldots, x_n)\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{\frac{1}{p}}.$$

Our aim is to show that $\| \ \|_p$ is a norm. It is easy to verify all the conditions defining a norm except the last one — the subadditivity condition.

In case that $p = \infty$ we use (1.1) to define $\| \ \|_\infty$. This fits into the scheme in that

$$\max_{k=1}^{n} |x_k| = \lim_{p \longrightarrow \infty} \|(x_1, \ldots, x_n)\|_p.$$

Let $1 \le p \le \infty$. Then we define $p' = \frac{p}{p-1}$ the ***conjugate index*** of p. In case $p = 1$ we take $p' = \infty$, and in case $p = \infty$ we take $p' = 1$. We have

$$\frac{1}{p} + \frac{1}{p'} = 1,$$

so that the relationship between index and conjugate index is symmetric.

PROPOSITION 35 (HÖLDER'S INEQUALITY) *For $x, y \in \mathbb{C}^n$ we have*

$$|\sum_{j=1}^{n} x_j y_j| \le \|x\|_p \, \|y\|_{p'} \tag{3.1}$$

If $p = 1$ or $p = \infty$, Hölder's Inequality is easy to verify. In the general case we use the following lemma.

LEMMA 36 *Let $x \ge 0$ and $y \ge 0$. Let $1 < p < \infty$ and let p' be the conjugate index of p, so that $1 < p' < \infty$. Then*

$$xy \le \frac{1}{p}x^p + \frac{1}{p'}y^{p'}. \tag{3.2}$$

Proof. First of all, if $x = 0$ or $y = 0$ the inequality is obvious. We therefore assume that $x > 0$ and $y > 0$.

Next, observe that if $t > 0$ and we replace x by $t^{\frac{1}{p}}x$ and y by $t^{\frac{1}{p'}}y$ in (3.2) then since $\frac{1}{p} + \frac{1}{p'} = 1$, (3.2) is multiplied by t and its content is unchanged. Choosing t appropriately (in fact with $t = y^{-p'}$), we can assume without loss of generality that $y = 1$. The problem is now reduced to one-variable calculus.

Let us define a function f on $]0, \infty[$ by

$$f(x) = \frac{1}{p}x^p - x + \frac{1}{p'}.$$

Taking the derivative of f we obtain

$$f'(x) = x^{p-1} - 1.$$

Since $p > 1$ this leads to

$$f'(x) \ge 0 \qquad \text{if } x \ge 1, \tag{3.3}$$

$$f'(x) \le 0 \qquad \text{if } x \le 1. \tag{3.4}$$

It follows from (3.3), (3.4) and the Mean-Value Theorem that

$$f(x) \ge f(1) \qquad \text{if } x \ge 1, \tag{3.5}$$

$$f(x) \ge f(1) \qquad \text{if } x \le 1. \tag{3.6}$$

Since $f(1) = 0$, (3.5) and (3.6) lead to

$$f(x) \geq 0 \qquad \forall x > 0. \tag{3.7}$$

But (3.7) is equivalent to (3.2) in case $y = 1$, completing the proof of Lemma 36. ■

Proof of Hölder's Inequality. We first suppose that $\|x\|_p = 1$ and $\|y\|_{p'} = 1$. Then, by multiple applications of Lemma 36 we have

$$\begin{aligned} |\sum_{j=1}^{n} x_j y_j| &\leq \sum_{j=1}^{n} |x_j||y_j| \\ &\leq \sum_{j=1}^{n} \frac{1}{p}|x_j|^p + \frac{1}{p'}|y_j|^{p'} \\ &= \frac{1}{p}\|x\|_p^p + \frac{1}{p'}\|y\|_{p'}^{p'} \\ &= \frac{1}{p} + \frac{1}{p'} = 1. \end{aligned} \tag{3.8}$$

For the general case, we first observe that if $\|x\|_p = 0$, then $x = 0$ and the result is straightforward. We may assume that $\|x\|_p > 0$ and similarly that $\|y\|_{p'} > 0$. Then, applying (3.8) with x replaced by $\|x\|_p^{-1}x$ and y replaced by $\|y\|_{p'}^{-1}y$, we obtain

$$|\sum_{j=1}^{n} \frac{x_j}{\|x\|_p} \frac{y_j}{\|y\|_{p'}}| \leq 1.$$

Finally, multiplying by $\|x\|_p\|y\|_{p'}$ yields Hölder's inequality. ■

THEOREM 37 (MINKOWSKI'S INEQUALITY) *Let $1 \leq p \leq \infty$ and $x, y \in \mathbb{R}^n$. Then*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \tag{3.9}$$

holds.

Proof. The result is easy if $p = 1$ or if $p = \infty$. We therefore suppose that $1 < p < \infty$. We have

$$\begin{aligned}
\|x+y\|_p^p &= \sum_{j=1}^n |x_j + y_j|^p \\
&= \sum_{j=1}^n |x_j + y_j||x_j + y_j|^{p-1} \\
&\leq \sum_{j=1}^n (|x_j| + |y_j|)|x_j + y_j|^{p-1} \\
&= \sum_{j=1}^n |x_j||x_j + y_j|^{p-1} + \sum_{j=1}^n |y_j||x_j + y_j|^{p-1} \qquad (3.10)
\end{aligned}$$

The key is to apply Hölder's inequality to each of the two sums in (3.10). We have

$$\begin{aligned}
\sum_{j=1}^n |x_j||x_j + y_j|^{p-1} &\leq \left(\sum_{j=1}^n |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{j=1}^n |x_j + y_j|^{p'(p-1)}\right)^{\frac{1}{p'}} \\
&= \|x\|_p \|x+y\|_p^{p-1}. \qquad (3.11)
\end{aligned}$$

since $p'(p-1) = p$ and $\frac{1}{p'} = (p-1)\frac{1}{p}$. Similarly

$$\sum_{j=1}^n |y_j||x_j + y_j|^{p-1} \leq \|y\|_p \|x+y\|_p^{p-1}. \qquad (3.12)$$

Combining now (3.10), (3.11) and (3.12), we obtain

$$\|x+y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x+y\|_p^{p-1}. \qquad (3.13)$$

Now if $\|x+y\|_p = 0$ we have the conclusion (3.9). If not, then it is legitimate to divide (3.13) by $\|x+y\|_p^{p-1}$ and again the conclusion follows. ■

The p-norms are used most frequently in the cases $p = 1$, $p = 2$ and $p = \infty$. The case $p = 2$ is special in that the 2-norm is the Euclidean norm which arises from an inner product. In particular the standard Cauchy-Schwarz inequality

$$|\sum_{k=1}^n x_k y_k| \leq \left(\sum_{j=1}^n |x_j|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^n |y_k|^2\right)^{\frac{1}{2}}$$

is just the case $p = 2$ of Hölder's Inequality (3.1).

3.2 Minkowski's Inequality and convexity

The proof of Theorem 3.9 is really very slick. However, it is not easy to understand the motivating forces behind the proof. We have seen in Theorem 3 (which relates to the line condition) that the subadditivity of a norm is related to convexity. If there is justice, it should be possible to understand Minkowski's Inequality as a convexity inequality. This is the purpose of this section.

DEFINITION *Let $a < b$ and suppose that $f :]a, b[\longrightarrow \mathbb{R}$. Then we say that f is* ***convex*** *, or more precisely a* ***convex function*** *iff it satisfies the inequality*

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2)$$

for all $x_1, x_2 \in]a, b[$ and for all t satisfying $0 \leq t \leq 1$.

The rationale for this definition and one way in which it relates to convex sets is that f is a convex function iff the region

$$\{(x, y); a < x < b, y > f(x)\}$$

lying above the graph of f is a convex subset of the plane $\mathbb{R}^2$.

The connection with norms is also clear. If $\| \ \|$ is a norm on a real vector space V then the function

$$f(t) = \|v_1 + tv_2\|$$

is convex on $\mathbb{R}$ for every fixed v_1 and v_2 in V. This result has a converse.

LEMMA 38 *Let $\| \ \|$ be a quantity defined on a real vector space V and such that the first three conditions of the definition of a norm hold. Suppose that*

$$f(t) = \|v_1 + tv_2\| \tag{3.14}$$

is convex on $\mathbb{R}$ for every fixed v_1 and v_2 in V. Then $\| \ \|$ satisfies the fourth condition and in consequence is a norm on V.

Proof. Let u_1 and u_2 be elements of V. We must show that

$$\|u_1 + u_2\| \leq \|u_1\| + \|u_2\|.$$

Towards this we write $v_1 = \frac{1}{2}(u_1 + u_2)$ and $v_2 = \frac{1}{2}(u_1 - u_2)$. Then using the fact that the function f in (3.14) is convex, we have

$$\|\tfrac{1}{2}(u_1 + u_2)\| = f(0) \leq \tfrac{1}{2}(f(-1) + f(1)) = \tfrac{1}{2}(\|u_1\| + \|u_2\|).$$

The desired result now follows from the homogeneity of $\| \ \|$. ■

The next step is to understand convex functions using differential calculus. The precise statement of the result depends on the degree of smoothness of the function. In fact, convex functions necessarily have a certain degree of regularity, but this issue is beyond the scope of this discussion.

THEOREM 39 *Let $a < b$ and suppose that $f :]a, b[\longrightarrow \mathbb{R}$. Then we have*

- *If f is differentiable, then f is convex on $]a, b[$ iff f' is increasing (in the wide sense) on $]a, b[$.*
- *If f is twice differentiable, then f is convex on $]a, b[$ iff f'' is nonnegative on $]a, b[$.*

Proof. It is enough to prove the first statement. First, we assume that f' is increasing in the wide sense on $]a, b[$. Then, using the Mean Value Theorem we have

$$\begin{aligned}
&(1-t)f(x_1) + tf(x_2) - f((1-t)x_1 + tx_2) \\
&= (1-t)(f(x_1) - f((1-t)x_1 + tx_2)) + t(f(x_2) - f((1-t)x_1 + tx_2)) \\
&= (1-t)t(x_1 - x_2)f'(\xi_1) - t(1-t)(x_1 - x_2)f'(\xi_2) \qquad (3.15) \\
&= (1-t)t(x_1 - x_2)(f'(\xi_1) - f'(\xi_2)) \\
&\geq 0 \qquad (3.16)
\end{aligned}$$

where in (3.15) ξ_1 is between $(1-t)x_1 + tx_2$ and x_1, and ξ_2 is between x_2 and $(1-t)x_1 + tx_2$. The crucial point is that the two quantities $x_1 - x_2$ and $\xi_1 - \xi_2$ have the same sign. Together with the fact that f' is increasing in the wide sense, this justifies (3.16). This result is often called Jensen's Inequality.

Conversely, suppose that f is convex on $]a, b[$. Let

$$a < x_1 < x_2 < b \qquad (3.17)$$

and suppose that $0 < t < (x_2 - x_1)$. Now let

$$s = \frac{x_2 - x_1 - t}{x_2 - x_1}$$

a number that satisfies $0 \le s \le 1$ and is defined so that

$$x_1 + t = (1-s)x_1 + sx_2 \tag{3.18}$$

$$x_2 - t = sx_1 + (1-s)x_2 \tag{3.19}$$

Applying the convexity of f to (3.18) and (3.19) yields

$$f(x_1+t) \le (1-s)f(x_1) + sf(x_2)$$
$$f(x_2-t) \le sf(x_1) + (1-s)f(x_2)$$

which combine to give

$$f(x_1+t) + f(x_2-t) \le f(x_1) + f(x_2). \tag{3.20}$$

But (3.20) can be rewritten in the form

$$\frac{f(x_1+t) - f(x_1)}{t} \le \frac{f(x_2) - f(x_2-t)}{t}.$$

Passing to the limit as $t \to 0$ we find that $f'(x_1) \le f'(x_2)$. Since x_1 and x_2 are arbitrary points satisfying (3.17) we see that f' is increasing in the wide sense. ■

In theory, our plan should now be to use Lemma 38 and Theorem 39 to establish the Minkowski Inequality. However, in order to succeed, we will need to get at the second derivative and unfortunately for $1 < p < 2$ the function

$$t \longrightarrow \{\sum_{j=1}^{n} |x_j + ty_j|^p\}^{\frac{1}{p}}$$

is not twice differentiable. To avoid this problem, we realise that it is enough to establish

$$\left\{\sum_{j=1}^{n} (x_j + y_j)^p\right\}^{\frac{1}{p}} \le \left\{\sum_{j=1}^{n} x_j^p\right\}^{\frac{1}{p}} + \left\{\sum_{j=1}^{n} y_j^p\right\}^{\frac{1}{p}} \tag{3.21}$$

in the case that $x_j, y_j \ge 0$. Indeed, by continuity, it will be enough to prove (3.21) in the case $x_j, y_j > 0$. For this, it suffices to establish that $f''(0) \ge 0$ where

$$f(t) = \left\{\sum_{j=1}^{n} (x_j + ty_j)^p\right\}^{\frac{1}{p}}$$

supposing that $x_j > 0$ and $y_j \in \mathbb{R}$. The proof of this fact follows that of Lemma 38. We leave the details to the reader.

This is much better because f is twice differentiable in a neighbourhood of 0. To aid calculations, let us set

$$f(t) = \{g(t)\}^{\frac{1}{p}}, \quad g(t) = \sum_{j=1}^{n} (x_j + ty_j)^p.$$

Then

$$f'(t) = \frac{1}{p}\{g(t)\}^{\frac{1}{p}-1} g'(t),$$
$$f''(t) = \frac{1}{p}\left(\frac{1}{p} - 1\right)\{g(t)\}^{\frac{1}{p}-2}(g'(t))^2 + \frac{1}{p}\{g(t)\}^{\frac{1}{p}-1} g''(t).$$

When the derivatives of g are calculated, the condition $f''(0) \geq 0$ finally boils down to

$$\left\{\sum_{j=1}^{n} x_j^p\right\}\left\{\sum_{j=1}^{n} x_j^{p-2} y_j^2\right\} \geq \left\{\sum_{j=1}^{n} x_j^{p-1} y_j\right\}^2$$

which is true in light of

$$\sum_{j=1}^{n}\sum_{k=1}^{n}(x_j y_k - x_k y_j)^2 x_j^{p-2} x_k^{p-2} \geq 0.$$

Again, we leave the details to the reader.

3.3 The sequence spaces ℓ^p

DEFINITION *Let $1 \leq p < \infty$. The the space ℓ^p is the vector space of all real sequences (x_k) for which the expression*

$$\|(x_k)\|_p = \left(\sum_{k \in \mathbb{N}} |x_k|^p\right)^{\frac{1}{p}} \tag{3.22}$$

is finite. The vector space operations on ℓ^p are defined coordinatewise. Thus

$$(tx + sy)_k = tx_k + sy_k \qquad k \in \mathbb{N}$$

where $x = (x_k)$ and $y = (y_k)$ are given elements of ℓ^p and t and s are reals.

Unfortunately, it is not immediately obvious that ℓ^p is a vector space, nor is it clear that (3.22) defines a genuine norm.

To see that ℓ^p is a vector space, we first invoke Minkowski's Inequality (3.9) in the form

$$\left(\sum_{k=1}^{n} |tx_k + sy_k|^p\right)^{\frac{1}{p}} \leq |t| \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} + |s| \left(\sum_{k=1}^{n} |y_k|^p\right)^{\frac{1}{p}} \tag{3.23}$$

Assuming that x, y in ℓ^p and bounding the right hand side of (3.23) we obtain

$$\left(\sum_{k=1}^{n} |tx_k + sy_k|^p\right)^{\frac{1}{p}} \leq |t|\|x\|_p + |s|\|y\|_p \tag{3.24}$$

for all $n \in \mathbb{N}$. Letting now n tend to ∞ on the left in (3.24) we find

$$\|tx + sy\|_p \leq |t|\|x\|_p + |s|\|y\|_p.$$

since the left hand side of (3.24) increases with n. This shows simultaneously that ℓ^p is a vector space and that (3.22) defines a norm on ℓ^p.

We would next like to establish the sequence space version of Hölder's Inequality. First, use (3.1) in the form

$$\sum_{j=1}^{n} |x_j||y_j| \leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^{p'}\right)^{\frac{1}{p'}}.$$

Again, let n tend to ∞ on the right to obtain

$$\sum_{j=1}^{n} |x_j y_j| \leq \|x\|_p\|y\|_{p'}$$

Letting n tend to infinity on the left, we again obtain an increasing limit

$$\sum_{j=1}^{\infty} |x_j y_j| \leq \|x\|_p\|y\|_{p'}.$$

Finally this gives

$$|\sum_{j=1}^{\infty} x_j y_j| \leq \|x\|_p \|y\|_{p'}. \tag{3.25}$$

the sequence space version of Hölder's Inequality.

We leave the reader to check two points. Firstly, ℓ^2 is an inner product space under

$$\langle (x_j), (y_j) \rangle = \sum_{j=1}^{\infty} x_j y_j.$$

The norm associated to this inner product is just the ℓ^2 norm. Secondly, the space ℓ^p is separable for $1 \leq p < \infty$.

3.4 Premetrics

Some examples of Metric Spaces stem naturally from the concept of a ***premetric*** .

DEFINITION *A* ***premetric function*** *on a set* X *is a function* $\rho : X \times X \longrightarrow [0, \infty]$ *such that*

- $\rho(x, x) = 0 \qquad \forall x \in X.$
- $\rho(x, y) > 0 \qquad \forall x, y \in X$ *such that* $x \neq y.$
- $\rho(x, y) = \rho(y, x) \qquad \forall x, y \in X.$

We think of $\rho(x, y)$ as the cost of moving from x to y in a single step. If it is impossible to move from x to y in a single step, this cost is infinite. A path from x to y is a finite chain $x = x_1, x_2, \ldots, x_{n-1}, x_n = y$ such that each link in the chain can be achieved in a single step, that is

$$\rho(x_j, x_{j+1}) < \infty \qquad \forall j = 1, 2, \ldots, n-1.$$

The metric is then defined by

$$d(x, y) = \inf \sum_{j=1}^{n-1} \rho(x_j, x_{j+1}) \tag{3.26}$$

where the infimum is taken over all paths from x to y. The metric function d then automatically satisfies the triangle inequality. However two remaining conditions have to be checked.

- For all x and y in X there must be some path from x to y in which each link has finite cost.
- It remains to be checked that $d(x, y) = 0$ implies that $x = y$. This may not be easy. It may be possible for a path from x to y with many links of very small cost to have arbitrarily small total cost.

EXAMPLE Let X be the set of finite character strings on a finite alphabet, say the lower case letters "a" through "z". We say that two strings t and s are adjacent iff one can be transformed into the other by one of the following operations (which simulate typing errors).

- Deletion of a single character anywhere in the string.
- Insertion of a single character anywhere in the string.
- Replacement of one character in the string by some other character.
- Transposition of two adjacent characters in the string.

We define $\rho(s, t) = 0$ if $s = t$, $\rho(s, t) = 1$ if s and t are adjacent and $\rho(s, t) = \infty$ in all other cases. Then (3.26) defines an integer valued metric on X. □

EXAMPLE A more general example relates to an undirected graph with (possibly infinite) vertex set V and edge set E. We assume that each edge $e \in E$ has a "weight" $w_e > 0$ attached to it. The we can define a premetric function

$$\rho_V(u, v) = \begin{cases} w_{\{u,v\}} & \text{if } \{u, v\} \in E, \\ \infty & \text{otherwise.} \end{cases}$$

Let us assume that the graph is connected in the sense that any two vertices can be linked by a *finite* path (chain of edges). Then it is clear that the function d_V defined by (3.26) is everywhere finite.

The function d_V may fail to be a metric on V however. Consider a graph with vertices x, y and z_{kj} for $j = 1, \ldots, k - 1$ and $k \in \mathbb{N}$. The edges are $\{x, y\}$ corresponding to $k = 1$, $\{x, z_{21}\}$ and $\{z_{21}, y\}$ corresponding to $k = 2$, $\{x, z_{31}\}$, $\{z_{31}, z_{32}\}$ and $\{z_{32}, y\}$ corresponding to $k = 3$ and so forth. Let the edges corresponding to a given value of k have weight k^{-2}. The it is clear that $d_V(x, y) = 0$ in spite of the fact that $x \neq y$ since for each $k \in \mathbb{N}$ there is a path from x to y having k links each with weight k^{-2} for a total cost of k^{-1}.

Nevertheless, there will also be many cases in which d_V is a genuine metric. □

EXAMPLE A more interesting example relates to this last example in case that d_V does define a metric. We construct a set X by "joining with a line segment" any two vertices that are linked by an edge. Thus each edge of the graph is "replaced" by a line segment and these line segments are "glued together" at the vertices. The descriptive notation

$$t\langle u\rangle + (1-t)\langle v\rangle \qquad (\{u,v\} \in E, t \in [0,1]) \tag{3.27}$$

specifies a typical point of X. The "scalar multiplications" and $+$ in this expression are purely symbolic and in no way represent algebraic operations. It is also understood that the expression $(1-t)\langle v\rangle + t\langle u\rangle$ represents exactly the same point as in (3.27). The point $1\langle v\rangle + 0\langle u\rangle$ represents the vertex v and is independent of u.

A premetric function will now be defined using the same weights $w_e > 0$ of the previous example. Two points can be joined in a single step if they lie on a common segment. Supposing that this segment corresponds to the edge $e = \{u, v\}$, we set

$$\rho(t\langle u\rangle + (1-t)\langle v\rangle, s\langle u\rangle + (1-s)\langle v\rangle) = w_e|t-s|$$

for such points. In all other cases we set $\rho(x,y) = \infty$. We assume as before that the graph is connected in the sense that any two vertices can be linked by a *finite* path. Then it is clear that the metric d_X defined by (3.26) is everywhere finite. It remains to show that

$$d_X(x,y) = 0 \qquad \Longrightarrow \qquad x = y.$$

Consider a path from x to y

$$x = x_1, x_2, \ldots, x_{n-1}, x_n = y \tag{3.28}$$

such that each link in the chain can be achieved in a single step. We claim that we can find a chain $x = \xi_1, \xi_2, \ldots, \xi_k = y$ at least as efficient as (3.28) and such that $\xi_2, \ldots, \xi_{k-1}$ are vertex points and all the points in the path are distinct. Typically the path (3.28) will pass through several vertex points. Let us suppose that x_ℓ and x_m are vertex points and that none of the intervening points $x_{\ell+1}, \ldots, x_{m-1}$ are vertex points. Then these intervening points necessarily lie on the same segment and it follows from the extended triangle inequality for $[0,1]$ that it is at least as efficient to remove them. In this way we obtain a path $x = \xi_1, \xi_2, \ldots, \xi_k = y$ in which $\xi_2, \ldots, \xi_{k-1}$ are vertex points. If a vertex appears twice in this path, for instance if $\xi_p = \xi_q$ with $p < q$, then we can obtain a more efficient path by omitting the points $\xi_{p+1}, \ldots, \xi_q$. We repeat this procedure until all the vertex points in the path are distinct.

It is now clear that if x and y are vertex points then we have $d_X(x, y) = d_V(x, y)$. Furthermore we can calculate d_X completely from d_V. Let $x = t\langle u\rangle + (1-t)\langle v\rangle$ and $y = s\langle z\rangle + (1-s)\langle w\rangle$ corresponding to distinct edges e and f respectively. Then $d_X(x, y)$ is the minimum of the four quantities

$$\begin{aligned}
&(1-t)w_e + (1-s)w_f + d_V(u, z),\\
&tw_e + (1-s)w_f + d_V(v, z),\\
&(1-t)w_e + sw_f + d_V(u, w),\\
&tw_e + sw_f + d_V(v, w).
\end{aligned}$$

If $x = t\langle u\rangle + (1-t)\langle v\rangle$ and $y = s\langle u\rangle + (1-s)\langle v\rangle$ lie on the same segment corresponding to the edge e and we assume without loss of generality that $t < s$, then $d_X(x, y)$ is the minimum of the two quantities

$$\begin{aligned}
&(1 + t - s)w_e + d_V(u, v)\\
&(s - t)w_e
\end{aligned}$$

In particular one can verify that $d_X(x, y) = 0$ implies $x = y$. □

EXAMPLE One particular example based on the previous one will be needed for a counterexample later in these notes. Let S be any set. Form a graph with vertex set $V = S \cup \{c\}$ where c is a special vertex called the centre. The edges of the graph all have the form $\{c, s\}$ where $s \in S$ and they all have unit weight. The corresponding space X is the ***star space*** based on S. In this case d_X is a metric.

The metric d_X can be described colloquially as follows. If two points lie on the same segment, the distance between them is the standard linear distance along the segment. If two points lie on different segments, then the distance between them is the linear distance of the first to the centre plus the linear distance of the second to the centre. □

3.5 Operator Norms

In this section we study continuous linear mappings between two normed vector spaces.

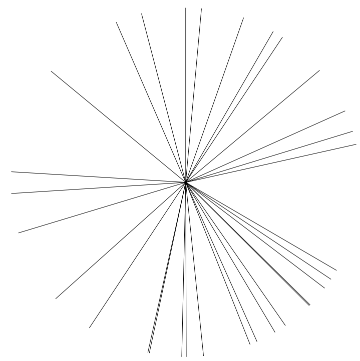

Figure 3.1: A typical star space.

THEOREM 40 *Let U and V be normed vector spaces. Let T be a linear mapping from U to V. Then the following are equivalent.*

- *T is continuous from U to V.*
- *T is continuous at 0_U.*
- *T is uniformly continuous from U to V.*
- *There exists a constant C such that $\|T(u)\|_V \le C\|u\|_U$ for all $u \in U$.*

Proof. We show that the fourth condition implies the third. In the fourth condition we replace u by $u_1 - u_2$ where u_1 and u_2 are arbitrary elements of U. Then, using the linearity of T in the form $T(u_1 - u_2) = T(u_1) - T(u_2)$ we see that the Lipschitz condition (page 25)

$$\|T(u_1) - T(u_2)\|_V \le C\|u_1 - u_2\|_U$$

holds. Thus T is uniformly continuous.

It is easy to see that the third condition implies the first, and the first condition implies the second.

It remains only to show that the second condition implies the fourth. For this we take $\epsilon = 1$ in the definition of the continuity of T at 0_U. There exists $\delta > 0$ such that

$$\|w\|_U < \delta \Rightarrow \|T(w)\|_V < 1. \tag{3.29}$$

We take $C = 2\delta^{-1}$. Then for $u \in U$ let $w = tu$ where $t = \frac{2}{3}\delta\|u\|^{-1}$. Then $\|w\| = \frac{2}{3}\delta < \delta$ and it follows from (3.29) that $\|T(w)\|_V < 1$ or equivalently $\|T(u)\|_V < \frac{3}{2}\delta^{-1}\|u\| \leq C\|u\|$. ■

Sometimes a continuous linear mapping is described as ***bounded linear*** . This is a different use of the word "bounded" from the one we have already met — bounded linear maps are not bounded in the metric space sense. Care is needed to make the correct interpretation of the word. We shall make use of the term ***continuous linear*** instead.

The space of all continuous linear maps from U to V is denoted $\mathcal{CL}(U,V)$. For $T \in \mathcal{CL}(U,V)$ we define

$$\|T\|_{\mathcal{CL}(U,V)} = \sup_{\|u\|\leq 1} \|T(u)\|. \tag{3.30}$$

It is an exercise to show that (3.30) defines a norm on $\mathcal{CL}(U,V)$ called the ***operator norm*** . A most particular case arises when U is finite dimensional. It then turns out that *all* linear mappings from U to V are continuous linear and we can view (3.30) as defining a norm on the space $\mathcal{L}(U,V)$ of all such linear maps. This fact is by no means obvious and can be obtained as a consequence of Corollary 84 on page 106.

If U, V and W are all normed spaces, $T \in \mathcal{CL}(U,V)$ and $S \in \mathcal{CL}(V,W)$ then it is easy to see that

$$\|S \circ T\| \leq \|S\|\|T\|$$

where all the norms are the appropriate operator norms. A particular case that is often used arises when all the space are equal. If $T \in \mathcal{CL}(U,U)$ then

$$\|T^n\| \leq \|T\|^n \quad \forall n \in \mathbb{Z}^+. \tag{3.31}$$

In particular the operator norm of the identity map I is unity. This is the case $n = 0$ in (3.31).

Operator norms are in general difficult to compute explicitly. When the underlying norms are Euclidean, the following result from linear algebra helps.

THEOREM 41 (SINGULAR VALUE DECOMPOSITION THEOREM) *Let A be a real $m \times n$ matrix. Then there exist orthogonal matrices U and V of shapes $m \times m$ and $n \times n$ respectively such that $A = UBV$ and B satisfies $b_{jk} = 0$ if $j \neq k$. Further the diagonal values b_{jj} for $j = 1, 2, \ldots, \min(m,n)$ may be taken nonnegative. They are called the* ***singular values*** *of A. Any two such decompositions yield the same singular values up to rearrangement.*

If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is given by

$$(Tx)_i = \sum_{j=1}^{n} a_{ij} x_j$$

and the norms taken on $\mathbb{R}^n$ and $\mathbb{R}^m$ are the Euclidean norms, then the operator norm of T is seen to be the largest singular value of the $m \times n$ matrix A. We leave the proof to the reader.

3.6 Continuous Linear Forms

If V is a normed real vector space, then it has a ***dual space*** , $V^\star$, the linear space of all linear mappings from V to $\mathbb{R}$. If V is finite dimensional, then *all* linear forms on V are continuous (this fact is not obvious and can be obtained from Corollary 84). If V is infinite dimensional, then one may have linear forms that are not continuous.

EXAMPLE Let F be the linear subspace of ℓ^1 of finitely supported sequences (x_n). A sequence is finitely supported if it satisfies the criterion

$$\exists N \in \mathbb{N} \text{ such that } (x_n = 0 \qquad \forall n \geq N).$$

Then F is a normed vector space with the norm restricted from ℓ^1. The form φ given by

$$(x_n) \quad \xrightarrow{\varphi} \quad \sum_{n=1}^{\infty} n x_n \tag{3.32}$$

is a perfectly good linear form on F. Note that the sum in (3.32) is in fact a finite sum, even though it is written as an infinite one. This ensures that φ is everywhere defined in F. Clearly $\varphi(e_n) = n$ while $\|e_n\|_{\ell^1} = 1$ so that φ is not continuous on F. □

EXAMPLE Finding a discontinuous linear form on ℓ^1 itself is considerably more difficult. Let e_n denote the sequence in ℓ^1 that has a 1 in the n-th place and 0 everywhere else. Consider the set $\{e_1, e_2, \ldots\}$. This set is linearly independent in ℓ^1. It is not however a spanning set. In fact its linear span is just the set F of the previous example. According to the basis extension theorem for linear spaces, we may extend it to a basis of ℓ^1 with $\{f_\alpha; \alpha \in I\}$, where I is some index set. It needs to be said that linear bases in infinite dimensional spaces are strange things — one needs to be able to write *every* vector in the space as a *finite* linear combination of basis vectors. So strange in fact that the ***Axiom of Choice*** is normally used in order to show the existence of $\{f_\alpha; \alpha \in I\}$. Indeed, nobody has ever written down an *explicit* example of such a set! Let f be one of these f_α's. Then we define the form θ as the mapping that takes an element of ℓ^1 to the coefficient of f in its corresponding basis representation. We will show that θ fails to be continuous. Indeed, θ vanishes on F, because the basis representation of an element (x_n) of F is just the *finite* sum

$$(x_n) = \sum_{n=1}^{\infty} x_n e_n.$$

The term in f is not required and the corresponding coefficient is zero. Since F is a dense subset of ℓ^1, it follows from Proposition 23, that if θ were continuous then θ would be identically zero. But $\theta(f) = 1$ so this is not possible. □

The space V' of continuous linear forms on a normed vector space V has a natural norm, namely the operator norm (using $|\ \ |$ as a norm on $\mathbb{R}$).

$$\|\varphi\|_{V'} = \sup_{\|v\|_V \le 1} |\varphi(v)|. \tag{3.33}$$

A most interesting situation develops in the case that V is finite dimensional and $V' = V^\star$. Here, since $V^\star$ is itself again finite dimensional, the second dual $V'' = V^{\star\star}$ and it is well known that $V^{\star\star}$ and V are naturally isomorphic. This allows us to construct a second dual norm on V. The following result asserts that this second dual norm is identical to the original norm on V.

PROPOSITION 42 *Let V be a finite dimensional real normed vector space. Let $v \in V$. Then*

$$\|v\| = \sup_{\|\varphi\|_{V'} \le 1} |\varphi(v)|. \tag{3.34}$$

This Proposition will be obtained as a consequence of the following Theorem, the proof of which will be given later (on page 142).

THEOREM 43 (SEPARATION THEOREM FOR CONVEX SETS) *Let C be an open convex subset of V a finite dimensional real normed vector space. Let $v \in V \setminus C$. Then there exists a linear form φ on V such that $\varphi(c) < \varphi(v)$ for all $c \in C$.*

Geometrically, the conclusion of Theorem 43 asserts that the convex set C lies entirely in the open halfspace

$$\{u; u \in V, \varphi(u) < \varphi(v)\}$$

whose boundary is the affine hyperplane

$$M = \{u; u \in V, \varphi(u) = \varphi(v)\}.$$

Proof of Proposition 42. Clearly we have

$$\sup_{\|\varphi\|_{V'} \leq 1} |\varphi(v)| \leq \|v\|. \tag{3.35}$$

If equality does not hold in (3.35), then after renormalizing we can suppose that there exists $v \in V$ such that $\|v\| = 1$ while $|\varphi(v)| < \|\varphi\|_{V'}$ for all $\varphi \in V'$. Now let $C = \{c; c \in V, \|c\| < 1\}$ the open unit ball of V. Clearly C is an open convex subset of V. Also $v \notin C$. Hence, by the Separation Theorem, there exists $\theta \in V'$ such that $\theta(c) < \theta(v)$ for all $c \in C$. But since $c \in C$ implies that $-c \in C$ we can also write

$$|\theta(c)| < \theta(v) \qquad \forall c \in C.$$

It now follows that $\|\theta\|_{V'} \leq \theta(v)$. Finally this yields

$$\|\theta\|_{V'} \leq \theta(v) \leq |\theta(v)| < \|\theta\|_{V'},$$

the required contradiction. ∎

3.7 Equivalent Metrics

There are various notions of ***equivalence*** for metrics on a set X. The standard definition follows.

DEFINITION *Let X be a set and suppose that d_1 and d_2 are two metrics on X. Then d_1 and d_2 are said to be **metrically equivalent** if there exist constants C_1 and C_2 with $0 < C_j < \infty$ for $j = 1, 2$ such that*

$$C_1 d_1(x_1, x_2) \leq d_2(x_1, x_2) \leq C_2 d_1(x_1, x_2)$$

for all $x_1, x_2 \in X$.

We leave the reader to show that metric equivalence is an equivalence relation. This is the strongest form of equivalence. Some authors call this ***uniform equivalence*** , but in these notes we reserve this terminology for a concept to be defined shortly. Metric equivalence is not very useful, except in the case of normed spaces where there is really only one form of equivalence and we drop the adverb *metrically*.

DEFINITION *Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. Then two norms $\|\ \|_1$ and $\|\ \|_2$ are* ***equivalent*** *iff there exist strictly positive constants C_1 and C_2 such that*

$$C_1\|v\|_1 \le \|v\|_2 \le C_2\|v\|_1$$

for all $v \in V$.

In the metric space context there are much more interesting forms of equivalence that preserve underlying properties.

DEFINITION *Let X be a set and suppose that d_1 and d_2 are two metrics on X. Then d_1 and d_2 are said to be* ***topologically equivalent*** *iff the metric spaces (X, d_1) and (X, d_2) have the same open sets.*

It is clear from the definition that topological equivalence is an equivalence relation. There is a more subtle way of rephrasing the definition. Two metrics d_1 and d_2 are topologically equivalent iff the identity mapping I_X on X is continuous as a mapping from (X, d_1) to (X, d_2) and also from (X, d_2) to (X, d_1). This makes it clear that one could also say that two metrics are topologically equivalent iff they have the same convergent sequences. There are many other equivalent formulations.

This idea also suggests the final form of equivalence.

DEFINITION *Let X be a set and suppose that d_1 and d_2 are two metrics on X. Then d_1 and d_2 are said to be* ***uniformly equivalent*** *iff the identity mapping I_X on X is uniformly continuous as a mapping from (X, d_1) to (X, d_2) and also as a map from (X, d_2) to (X, d_1).*

Metric equivalence implies uniform equivalence and uniform equivalence implies topological equivalence.

For normed spaces, all three forms of equivalence are the same. This follows immediately from Theorem 40.

EXAMPLE On $\mathbb{R}$ consider

- $d_1(x, y) = |x - y|$.
- $d_2(x, y) = 2|x - y|$.
- $d_3(x, y) = \arctan(|x - y|)$.
- $d_4(x, y) = |\arctan(x) - \arctan(y)|$.

It is not immediately obvious that d_3 is a metric. To see this, one needs to establish

$$\arctan(x + y) \leq \arctan x + \arctan y \tag{3.36}$$

for $x, y \geq 0$. Let $t = \arctan x$ and $s = \arctan y$. Then in case that $t + s \geq \frac{\pi}{2}$, (3.36) is obvious. Thus, we may assume that $t, s \geq 0$ and that $t + s < \frac{\pi}{2}$. We need to show that

$$\tan t + \tan s \leq \tan(t + s). \tag{3.37}$$

But (3.37) follows from the trig identity

$$\tan(t + s) = \frac{\tan t + \tan s}{1 - \tan t \tan s},$$

and the observation that $1 \geq 1 - \tan t \tan s > 0$ since $t + s < \frac{\pi}{2}$.

It is immediately obvious that d_1, d_2 and d_4 are metrics on $\mathbb{R}$. It is straightforward to show that d_1 and d_2 are metrically equivalent, that d_1 and d_3 are uniformly equivalent, but not metrically equivalent and finally that d_1 and d_4 are topologically equivalent but not uniformly equivalent. □

3.8 The Abstract Cantor Set

Consider X_j to be a copy of the two point space $\{0, 1\}$ for $j = 1, 2, 3, \ldots$. To define the abstract Cantor set X we simply consider

$$X = \prod_{j=1}^{\infty} X_j,$$

the infinite product of the X_j. In effect, a point x of X is a sequence $x = (x_j)$ with $x_j \in \{0, 1\}$ for $j = 1, 2, 3, \ldots$.

Next, we define a metric on X. We want entries far up the sequence to have less weight than entries near the beginning of the sequence, so we define

$$d((x_j),(y_j)) = \sum_{j=1}^{\infty} 2^{-j}|x_j - y_j|.$$

Observe that since $|x_j - y_j| \leq 2$, the series always converges. The use of the weights 2^{-j} is somewhat arbitrary here. It is routine to verify that d defines a metric on X. It will be observed that convergence in (X,d) is ***coordinatewise*** or ***pointwise convergence***. The case is somewhat special here because two coordinates either agree or differ by 1.

We see that if (x_j) and (y_j) agree in the first k coordinates, then $d((x_j),(y_j)) \leq 2^{-k}$. Conversely, if $d((x_j),(y_j)) \leq 2^{-k}$ then (x_j) and (y_j) agree in the first $k-1$ coordinates.

The mapping $\alpha : X \longrightarrow \mathbb{R}$ given by

$$\alpha((x_j)) = 2\sum_{j=1}^{\infty} 3^{-j}x_j$$

maps X onto the standard Cantor set in $\mathbb{R}$. The metric

$$d_1((x_j),(y_j)) = |\alpha(x) - \alpha(y)| = 2\sum_{j=1}^{\infty} |3^{-j}(x_j - y_j)|.$$

on X reflects the standard metric on $\mathbb{R}$ through the mapping α. For k an integer with say $k \geq 3$ it is easy to show that

$$d(x,y) \leq 2^{-(k+1)} \Longrightarrow d_1(x,y) \leq 3^{-k}$$

and

$$d_1(x,y) \leq 3^{-(k+1)} \Longrightarrow d(x,y) \leq 2^{-k}.$$

It follows that d and d_1 are uniformly equivalent metrics on X.

3.9 The Quotient Norm

Let V be a normed vector space and let N be a closed linear subspace. Then we can consider the quotient space $Q = V/N$. This is a new linear space with a complicated definition highly unpopular with students.

One starts by defining a relation $\sim$ on V by

$$v_1 \sim v_2 \qquad \Leftrightarrow \qquad v_1 - v_2 \in N.$$

Here v_1, v_2 denote elements of V. One verifies that $\sim$ is an equivalence relation on V. There is then a quotient space Q and a canonical projection π,

$$\pi : V \longrightarrow Q.$$

It is now possible to show that Q can be given the structure of a linear space in such a way that π is a linear mapping. In addition one has that $\ker(\pi) = N$.

We now define a norm on Q known as the quotient norm by

$$\|q\|_Q = \inf_{\pi(v)=q} \|v\|_V, \tag{3.38}$$

for $q \in Q$. The infimum is taken over all elements $v \in V$ such that $\pi(v) = q$. It is more or less obvious that $\| \ \|_Q$ is homogenous.

To show the subadditivity of the norm, we argue by contradiction. Suppose that there exists $\epsilon > 0$, $q_1, q_2 \in Q$ such that

$$\|q_1 + q_2\| \geq \|q_1\| + \|q_2\| + 3\epsilon. \tag{3.39}$$

Then using the definition (3.38), we can find $v_1, v_2 \in V$ such that $\pi(v_j) = q_j$ and

$$\|v_j\|_V \leq \|q_j\| + \epsilon,$$

for $j = 1, 2$. Obviously, $\pi(v_1 + v_2) = q_1 + q_2$ so that

$$\|q_1 + q_2\| \leq \|v_1\| + \|v_2\| \leq \|q_1\| + \|q_2\| + 2\epsilon.$$

This contradiction with (3.39) establishes the subadditivity.

There is one final detail that requires a little proof. Suppose that $q \in Q$ and that $\|q\|_Q = 0$. Then, using (3.38) we can find a sequence (v_j) of elements of V with $\pi(v_j) = q$ for $j = 1, 2, \ldots$ and $\|v_j\|$ tending to zero. Clearly $v_j \longrightarrow 0_V$ and hence $(v_1 - v_j) \longrightarrow v_1$. Since $(v_1 - v_j) \in N$ and since N is supposed to be closed in V, we conclude that $v_1 \in N$ and consequently that $q = 0_Q$.

Some nice points lie outside our present reach since they depend on compactness. If V is finite dimensional then any linear subspace N of V is necessarily closed. Furthermore, in this case, the infimum of (3.38) is necessarily attained. A consequence is that the unit ball of Q is just the direct image by π of the unit ball of V. In the finite dimensional case, this gives a geometric way of understanding the quotient norm.

It should be pointed out that one can try to define general quotient metrics in much that same way, but the issues are much more problematic. If

$$\pi : X \longrightarrow Q.$$

is a quotienting of a metric space X we can define

$$d_Q(q_1, q_2) = \inf_{\substack{\pi(x_1)=q_1 \\ \pi(x_2)=q_2}} d_X(x_1, x_2)$$

but only under very stringent additional assumptions will this define a metric on Q.

4

Completeness

In this chapter we will assume that the reader is familiar with the completeness of $\mathbb{R}$. Usually $\mathbb{R}$ is defined as the unique order-complete totally ordered field. The order completeness postulate is that every subset B of $\mathbb{R}$ which is bounded above possesses a least upper bound (or supremum). From this the metric completeness of $\mathbb{R}$ is deduced. Metric completeness is formulated in terms of the convergence of Cauchy sequences. It is true that in making the link between the two for $\mathbb{R}$, one uses the Bolzano–Weierstrass Theorem which is a form of compactness. Nevertheless, we believe that for metric spaces, completeness is a more fundamental concept than compactness and should be treated first.

DEFINITION *Let X be a metric space. Let (x_n) be a sequence in X. Then (x_n) is a* ***Cauchy sequence*** *iff for every number $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that*

$$p, q > N \qquad \Rightarrow \qquad d(x_p, x_q) < \epsilon.$$

LEMMA 44 *Every convergent sequence is Cauchy.*

Proof. Let X be a metric space. Let (x_n) be a sequence in X converging to $x \in X$. Then given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \frac{1}{2}\epsilon$ for $n > N$. Thus for $p, q > N$ the triangle inequality gives

$$d(x_p, x_q) \le d(x_p, x) + d(x, x_q) < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon.$$

Hence (x_n) is Cauchy. ■

The Cauchy condition on a sequence says that the diameters of the successive tails of the sequence converge to zero. One feels that this is almost equivalent to convergence except that no limit is explicitly mentioned. Sometimes, Cauchy sequences fail to converge because the "would be limit" is not in the space. It is the existence of such "gaps" in the space that prevent it from being complete.

DEFINITION *Let X be a metric space. Then X is* ***complete*** *iff every Cauchy sequence in X converges in X.*

EXAMPLE The real line $\mathbb{R}$ is complete. □

EXAMPLE The set $\mathbb{Q}$ of rational numbers is not complete. Consider the sequence defined inductively by

$$x_1 = 2 \quad \text{and} \quad x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right), \quad n = 1, 2, \ldots. \tag{4.1}$$

Then one can show that (x_n) converges to $\sqrt{2}$ in $\mathbb{R}$. It follows that (x_n) is a Cauchy sequence in $\mathbb{Q}$ which does not converge in $\mathbb{Q}$. Hence $\mathbb{Q}$ is not complete.

To fill in the details, observe first that (4.1) can also be written in both of the alternative forms

$$2x_n(x_{n+1} - \sqrt{2}) = (x_n - \sqrt{2})^2,$$
$$x_{n+1} - x_n = -\left(\frac{x_n^2 - 2}{2x_n}\right).$$

We now observe the following in succession.

- $x_n > 0$ for all $n \in \mathbb{N}$.
- $x_n > \sqrt{2}$ for all $n \in \mathbb{N}$.
- x_n is decreasing with n.
- $x_n \leq 2$ for all $n \in \mathbb{N}$.
- $|x_{n+1} - \sqrt{2}| \leq \dfrac{|x_n - \sqrt{2}|^2}{2\sqrt{2}}$ for all $n \in \mathbb{N}$.
- $|x_{n+1} - \sqrt{2}| \leq \dfrac{2 - \sqrt{2}}{2\sqrt{2}}|x_n - \sqrt{2}|$ for all $n \in \mathbb{N}$.

The convergence of (x_n) to $\sqrt{2}$ follows easily. □

4.1 Boundedness and Uniform Convergence

DEFINITION *Let A be a subset of a metric space X. Then the* ***diameter*** *of A is defined by*

$$\operatorname{diam}(X) = \sup_{x_1, x_2 \in A} d(x_1, x_2). \tag{4.2}$$

We say that A is a ***bounded subset*** *iff $\operatorname{diam}(A) < \infty$. We note that we regard the subset $A = \emptyset$ as being bounded, even though formally the supremum in (4.2) is illegal. We say that the metric space X is bounded iff it is a bounded subset of itself. We say that a sequence (x_n) is bounded iff its underlying set is bounded. We say that a function $f : Y \longrightarrow X$ is bounded iff $f(Y)$ is a bounded subset of X.*

These definitions are coloured by the fact that they relate to metric spaces. In the case of a normed vector space V, there is another alternative, provided by the distinguished element 0_V.

DEFINITION *A subset A of a normed vector space V is bounded iff*

$$\sup_{v \in A} \|v\| < \infty.$$

A moment's thought shows that the two concepts of boundedness are equivalent.

The following lemma follows directly from the definition of a Cauchy sequence and Lemma 44.

LEMMA 45

- *Every Cauchy sequence is bounded.*
- *Every convergent sequence is bounded.*

Our main immediate motivation for introducing boundedness at this point is the construction of additional examples of complete metric spaces. Let X be a non-empty set and Y a metric space. Then we denote by $B(X, Y)$ the set of all bounded mappings from X to Y. We can turn this into a metric space by defining the metric

$$d(f, g) = \sup_{x \in X} d_Y(f(x), g(x)). \tag{4.3}$$

This is the ***supremum metric*** or ***uniform metric*** . In case that Y is a normed vector space, it is easy to check that

$$\|f\|_{B(X,Y)} = \sup_{x\in X} \|f(x)\|_Y \tag{4.4}$$

is a norm on $B(X,Y)$ and that the metric that it induces is precisely the one given by (4.3). The norm defined by (4.4) is called the ***supremum norm*** or ***uniform norm*** .

The convergence of $B(X,Y)$ is called ***uniform convergence*** . On the other hand, we say that a sequence of functions (f_n) converges pointwise iff $(f_n(x))$ converges for every x in X. If a sequence converges uniformly, then it converges pointwise. But the converse is not true.

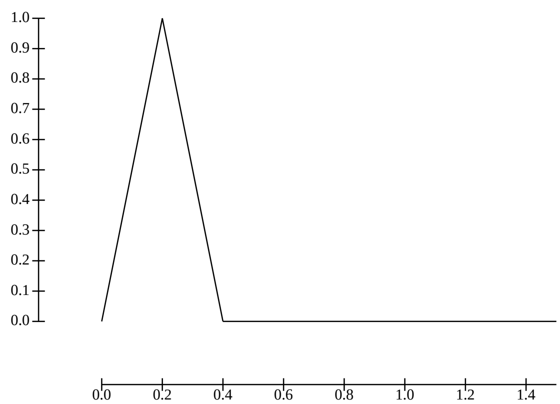

Figure 4.1: The function f_5.

EXAMPLE Consider the case $X = \mathbb{R}^+$, $Y = \mathbb{R}$ and suppose that the sequence (f_n) is given by

$$f_n(x) = \begin{cases} nx & \text{if } 0 \le x \le n^{-1}, \\ 2 - nx & \text{if } n^{-1} \le x \le 2n^{-1}, \\ 0 & \text{if } x \ge 2n^{-1}. \end{cases} \tag{4.5}$$

Now, if $x = 0$ then $f_n(0) = 0$ for all n so that $f_n(0) \longrightarrow 0$ as $n \longrightarrow \infty$. On the other hand, if $x > 0$, then for large values of n, it is the third line of (4.5) that applies. Once again, we obtain $f_n(x) \longrightarrow 0$. Thus the sequence (f_n) converges pointwise to 0.

This convergence is not uniform. To see this, we simply put $x = n^{-1}$. Then $f_n(\frac{1}{n}) = 1$ and it follows that $d_{B(\mathbb{R}^+,\mathbb{R})}(f_n, 0) = 1$ for all n. $\square$

While uniform convergence is just convergence in the metric of $B(X, Y)$, there is *no metric* which gives rise to pointwise convergence. From the point of view of Metric Spaces, pointwise convergence of sequences of functions is some kind of *rogue* convergence that does not fit the theory. In these notes we just have to live with this unfortunate circumstance. However there is a topology on $B(X, Y)$ (and on other spaces of functions) with the property that convergence in this topology is exactly pointwise convergence. The need to have a unified theory of convergence therefore forces one into the realm of topological spaces with all of its associated pathology.

PROPOSITION 46 *If Y is complete, so is $B(X, Y)$.*

Proof. The pattern of most completeness proofs is the same. Take a Cauchy sequence. Use some existing completeness information to deduce that the sequence converges in some weak sense. Use the Cauchy condition again to establish that the sequence converges in the metric sense.

Let (f_n) be a Cauchy sequence in $B(X, Y)$. Then, for each $x \in X$, it is straightforward to check that $(f_n(x))$ is a Cauchy sequence in Y and hence converges to some element of Y. This can be viewed as a rule for assigning an element of Y to every element of X — in other words, a function f from X to Y. We have just shown that (f_n) converges to f pointwise.

Now let $\epsilon > 0$. Then for each $x \in X$ there exists $N_x \in \mathbb{N}$ such that

$$q > N_x \qquad \Rightarrow \qquad d(f_q(x), f(x)) < \tfrac{1}{3}\epsilon. \tag{4.6}$$

Now we reuse the Cauchy condition — there exists $N \in \mathbb{N}$ such that

$$p, q > N \qquad \Rightarrow \qquad \sup_{x \in X} d(f_p(x), f_q(x)) < \tfrac{1}{3}\epsilon. \tag{4.7}$$

Now, combining (4.6) and (4.7) with the triangle inequality and choosing q explicitly as $q = \max(N, N_x) + 1$, we find that

$$p > N \qquad \Rightarrow \qquad d(f_p(x), f(x)) < \tfrac{2}{3}\epsilon \quad \forall x \in X. \tag{4.8}$$

We emphasize the crucial point that N depends only on ϵ. It does not depend on x. Thus we may deduce

$$p > N \qquad \Rightarrow \qquad \sup_{x \in X} d(f_p(x), f(x)) < \epsilon. \tag{4.9}$$

from (4.8).

This would be the end of the proof, if it were not for the fact that we still do not know that $f \in B(X,Y)$. For this, choose an explicit value of ϵ, say $\epsilon = 1$. Then, using the corresponding specialization of (4.9), we see that there exists $r \in \mathbb{N}$ such that

$$\sup_{x \in X} d(f_r(x), f(x)) < 1. \tag{4.10}$$

Now, substitute (4.10) into the extended triangle inequality

$$d(f(x_1), f(x_2)) \leq d(f(x_1), f_r(x_1)) + d(f_r(x_1), f_r(x_2)) + d(f_r(x_2), f(x_2))$$

to obtain

$$d(f(x_1), f(x_2)) \leq 1 + d(f_r(x_1), f_r(x_2)) + 1.$$

It now follows that since f_r is bounded, so is f. Finally, with the knowledge that $f \in B(X,Y)$ we see that (f_n) converges to f in $B(X,Y)$ by (4.9). ■

There is an alternative way of deducing (4.9) from (4.7) which worth mentioning. Conceptually it is simpler than the argument presented above, but perhaps less rigorous. We write (4.7) in the form

$$p, q > N \qquad \Rightarrow \qquad d(f_p(x), f_q(x)) < \tfrac{1}{3}\epsilon. \tag{4.11}$$

where x is a general point of X. The vital key is that N depends only on ϵ and not on x. Now, letting $q \longrightarrow \infty$ in (4.11) we find

$$p > N \qquad \Rightarrow \qquad d(f_p(x), f(x)) \leq \tfrac{1}{3}\epsilon. \tag{4.12}$$

because $f_q(x)$ converges pointwise to $f(x)$. Here we are using the fact that $[0, \frac{1}{3}\epsilon]$ is a closed subset of $\mathbb{R}$. Since N depends only on ϵ we can then deduce (4.9) from (4.12).

EXAMPLE An immediate Corollary of the above is that the space ℓ^∞ is complete. The same is true of ℓ^p for $1 \le p < \infty$. We sketch the details. Let (x_n) be a Cauchy sequence of elements of ℓ^p. Then each such element x_n is actually a sequence x_{nk} of real numbers. It is easy to see that for each fixed k, $(x_{nk})_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$. Using the completeness of $\mathbb{R}$ we infer the existence of $\xi_k \in \mathbb{R}$ such that

$$x_{nk} \longrightarrow \xi_k$$

as $n \to \infty$. We now use again the fact that (x_n) is a Cauchy sequence. Let $\epsilon > 0$. Then there exist $N \in \mathbb{N}$ such that

$$m, n > N \Longrightarrow \|x_m - x_n\|_p < \epsilon.$$

Then, for all $m, n > N$ and for all $K \in \mathbb{N}$ we have

$$\sum_{k=1}^{K} |x_{mk} - x_{nk}|^p \le \epsilon^p.$$

Letting $m \to \infty$, this leads to

$$\sum_{k=1}^{K} |\xi_k - x_{nk}|^p \le \epsilon^p,$$

because only finitely many values of k are involved. Finally letting $K \to \infty$ we get $\|\xi - x_n\|_p \le \epsilon$ for all $n > N$. This gives the desired convergence to an element ξ of ℓ^p. As above, a little extra work is necessary to show that $\xi \in \ell^p$. □

4.2 Subsets and Products of Complete Spaces

We seek other ways of building new complete spaces from old.

PROPOSITION 47

- *Let X be a complete metric space. Let Y be a closed subset of X. Then Y is complete (as a metric space in its own right).*
- *Let X be a metric space. Let Y be a subset of X which is complete (as a metric space in its own right). Then Y is closed in X.*

Proof. We establish the first statement. Let (x_n) be a Cauchy sequence in Y. Then (x_n) is a Cauchy sequence in X. Since X is complete, there exists $x \in X$ such that (x_n) converges to x. But since the sequence (x_n) lies in Y and Y is closed, $x \in Y$.

For the second statement, let (x_n) be a sequence in Y converging to some element x of X. We aim to show that $x \in Y$. By Lemma 44, (x_n) is a Cauchy sequence in X. Hence (x_n) is also a Cauchy sequence in Y. But Y is complete. It follows that there exists an element $y \in Y$ such that (x_n) converges to y. Then by Proposition 6, $x = y$. Hence $x \in Y$. ■

This Proposition gives the correct impression that completeness is a kind of global closedness property. We have the following useful corollary.

COROLLARY 48 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$ be an isometry. Suppose that X is complete and that $f(X)$ is dense in Y. Then f is onto.*

Proof. Since X is complete and f is an isometry, $f(X)$ is complete and hence closed in Y. But since $f(X)$ is also dense in Y, $f(X) = Y$. ■

PROPOSITION 49 *Let X and Y be complete spaces. Then so is $X \times Y$.*

Proof. Let (x_n, y_n) be a Cauchy sequence in $X \times Y$. Then it is easy to see that the component sequences (x_n) and (y_n) are Cauchy in X and Y respectively. Since X and Y are both complete, it follows that there exist limits x and y respectively. The result now follows directly from Lemma 16. ■

PROPOSITION 50 *Let X and Y be metric spaces. Let (f_n) be a sequence of continuous functions $f_n : X \longrightarrow Y$, converging uniformly to a function f. Then f is also continuous.*

Proof. The proof we give is by epsilons and deltas. Let $x_0 \in X$ — we will show that f is continuous at x_0. Suppose that $\epsilon > 0$. Then by the uniform convergence, there exists $N \in \mathbb{N}$ such that

$$n > N \qquad \Rightarrow \qquad \sup_{x \in X} d_Y(f(x), f_n(x)) < \tfrac{1}{3}\epsilon. \tag{4.13}$$

Let us fix $n = N + 1$. Now we use the fact that this one function f_n of the sequence is continuous at x_0. There exists $\delta > 0$ such that

$$d_X(x, x_0) < \delta \qquad \Rightarrow \qquad d_Y(f_n(x), f_n(x_0)) < \tfrac{1}{3}\epsilon. \tag{4.14}$$

Combining (4.13) and (4.14) we now obtain for $x \in X$ satisfying $d_X(x, x_0) < \delta$

$$\begin{aligned} d_Y(f(x), f(x_0)) &\le d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(x_0)) + d_Y(f_n(x_0), f(x_0)) \\ &< \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon + \tfrac{1}{3}\epsilon = \epsilon. \end{aligned}$$

This shows that f is continuous. ■

EXAMPLE We give next an important example of Proposition 47. Let X and Y be metric spaces. Then we denote by $C(X, Y)$ the subset of $B(X, Y)$ consisting of bounded continuous functions from X to Y. We claim that $C(X, Y)$ is a closed subset of $B(X, Y)$. This is an immediate consequence of Proposition 50. Applying the first assertion of Proposition 47 shows that $C(X, Y)$ is a complete space if Y is. □

EXAMPLE The star space X based on a set S is always complete. The definition is found on page 59. Let (x_n) be any sequence in X. Then we can denote $x_n = (1 - t_n)\langle c\rangle + t_n\langle s_n\rangle$ where c denotes the centre of X, (t_n) is a sequence in $[0, 1]$ and (s_n) is a sequence in S.

Suppose now that (x_n) is a Cauchy sequence in X. If $t_n \longrightarrow 0$ as $n \longrightarrow \infty$ then $x_n \longrightarrow c$ in X and we have convergence. Thus, to establish convergence we may assume that there exists a strictly positive number ϵ such that for all $N \in \mathbb{N}$ there exists $p > N$ such that $t_p > \epsilon$. Apply now the Cauchy condition with this ϵ. There exists $N \in \mathbb{N}$ such that

$$p, q > N \qquad \Rightarrow \qquad d_X(x_p, x_q) < \epsilon.$$

Choose p as described above. If $s_p \ne s_q$ so that x_p and x_q lie on different rays then $d_X(x_p, x_q) = t_p + t_q > \epsilon$ a contradiction. Hence $s_q = s_p$ for all $q > N$. Thus a tail of the sequence lies on a single ray where the metric is essentially just that of $[0, 1]$. In other words, (t_n) is a Cauchy sequence in $[0, 1]$. Since $[0, 1]$ is complete (it is a closed subset of $\mathbb{R}$), it follows that there exists $t \in [0, 1]$ such that $t_n \longrightarrow t$. The sequence (s_n) is eventually constant, so that (x_n) converges to $(1 - t)\langle c\rangle + t\langle s_p\rangle$. □

A major application of completeness in normed spaces is the existence of ***absolutely convergent sums*** .

PROPOSITION 51 *Let V be a complete normed space and let v_j be elements of V for $j \in \mathbb{N}$. Suppose that*

$$\sum_{j=1}^{\infty} \|v_j\|_V < \infty.$$

Then the sequence of partial sums (s_n) given by

$$s_n = \sum_{j=1}^{n} v_j$$

converges to an element $s \in V$. Furthermore we have the norm estimate

$$\|s\|_V \leq \sum_{j=1}^{\infty} \|v_j\|_V. \tag{4.15}$$

In this situation, it is natural to denote

$$s = \sum_{j=1}^{\infty} v_j.$$

Proof. We show that (s_j) is a Cauchy sequence. Let $1 \leq q \leq p$. Then applying the extended triangle inequality (1.10) to

$$s_p - s_q = \sum_{j=q+1}^{p} v_j,$$

we obtain

$$\|s_p - s_q\|_V \leq \sum_{j=q+1}^{p} \|v_j\|_V \leq \sum_{j=q+1}^{\infty} \|v_j\|_V. \tag{4.16}$$

But since the right hand term of (4.16) tends to 0 as $q \longrightarrow \infty$, it follows that (s_j) is a Cauchy sequence. Since V is complete we deduce that (s_j) converges to some element s of V. Putting $q = 0$ in (4.16) shows that

$$\|s_p\|_V \leq \sum_{j=1}^{\infty} \|v_j\|_V. \tag{4.17}$$

Since the norm is continuous on V, we see that (4.15) follows from (4.17) as $p \longrightarrow \infty$. ■

4.3 Contraction Mappings

DEFINITION *Let X be a metric space. Let $f : X \longrightarrow X$. Then f is a* ***contraction mapping*** *iff there exists a constant α with $0 \leq \alpha < 1$ such that*

$$d_X(f(x_1), f(x_2)) \leq \alpha d_X(x_1, x_2).$$

The following Theorem will be used extensively in the calculus section of this book.

THEOREM 52 (CONTRACTION MAPPING THEOREM) *Let f be a contraction mapping on a complete non-empty metric space X. Then there is a unique point $x \in X$ such that $f(x) = x$.*

Proof. Let $x_1 \in X$. Define $x_n \in X$ inductively by $x_{n+1} = f(x_n)$ ($n \in \mathbb{N}$). An easy induction argument establishes that

$$d(x_n, x_{n+1}) \leq \alpha^{n-1} d(x_1, x_2) \qquad \forall n \in \mathbb{N}.$$

We then apply the extended triangle inequality to obtain for $p \leq q$

$$\begin{aligned} d(x_p, x_q) &\leq \sum_{n=p}^{q-1} d(x_n, x_{n+1}), \\ &\leq \sum_{n=p}^{q-1} \alpha^{n-1} d(x_1, x_2), \\ &\leq \alpha^{p-1}(1-\alpha)^{-1} d(x_1, x_2), \end{aligned}$$

since $0 \leq \alpha < 1$. It follows that (x_n) is a Cauchy sequence in X. Since X is complete, (x_n) converges to some element $x \in X$. Now

$$\begin{aligned} d(x, f(x)) &\leq d(x, x_n) + d(x_n, f(x_n)) + d(f(x_n), f(x)) \\ &\leq d(x, x_n) + \alpha^{n-1} d(x_1, x_2) + \alpha d(x_n, x) \end{aligned} \tag{4.18}$$

Since $0 \leq \alpha < 1$ and since $d(x, x_n) \longrightarrow 0$ as $n \longrightarrow \infty$, it follows that we can make the right hand side of (4.18) as small as we like, by taking n sufficiently large. It follows that $d(x, f(x)) = 0$ which can only occur if $f(x) = x$. This completes the existence part of the proof.

There is another way of seeing this last step of the proof that is worth mentioning. We know that $x_n \longrightarrow x$ as $n \longrightarrow \infty$. Since f is a Lipschitz mapping it is continuous, so $f(x_n) \longrightarrow f(x)$. But $f(x_n) = x_{n+1}$ and $x_{n+1} \longrightarrow x$. It follows from the uniqueness of the limit that $f(x) = x$.

It remains to check that the fixed point x is unique. Suppose that y also satisfies $f(y) = y$. Then we have

$$d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y).$$

Since $0 \leq \alpha < 1$ the only way out is that $d(x, y) = 0$ which gives $x = y$. ■

EXAMPLE Here we present an example of a mapping $f : X \longrightarrow X$ of a complete space X such that

$$d_X(f(x_1), f(x_2)) < d_X(x_1, x_2) \qquad \text{for } x_1 \neq x_2 \tag{4.19}$$

but which does not have a fixed point. Let $X = \mathbb{R}$ and let

$$f(x) = x + \tfrac{1}{2}(1 - \tanh(x)).$$

Then, applying the Mean Value Theorem, we have

$$f(x_1) - f(x_2) = f'(\xi)(x_1 - x_2)$$

for ξ between x_1 and x_2. In any case,

$$\tfrac{1}{2} \leq f'(\xi) = 1 - \tfrac{1}{2}\operatorname{sech}^2 x < 1$$

so that (4.19) holds. Since $\tanh(x) < 1$ we see that f has no fixed point. □

There is an extension of the Contraction Mapping Theorem which explains how the fixed point behaves under a perturbation of the contraction.

THEOREM 53 *Let X be a complete metric space and $0 \leq \alpha < 1$. Let $c \in X$ and let $r > 0$. Let P be another metric space and suppose that $f : P \times B(c, r) \longrightarrow X$ is a mapping such that*

- *$d(f(p, x_1), f(p, x_2)) \leq \alpha d(x_1, x_2)$ for $x_1, x_2 \in B(c, r)$ and each $p \in P$.*
- *$d(c, f(p, c)) < (1 - \alpha)r$ for all $p \in P$.*

- *The mapping $p \mapsto f(p,b)$ is continuous from P to X for each fixed $b \in U(c,r)$.*

Then there is a unique continuous mapping $g : P \longrightarrow U(c,r)$ with the property that $f(p, g(p)) = g(p)$ for all $p \in P$.

Proof. We first show that for each $p \in P$, the map $f(p,\cdot)$ maps $B(c,r)$ to itself. To see this, let $b \in B(c,r)$. Then

$$\begin{aligned} d(c, f(p,b)) &\leq d(c, f(p,c)) + d(f(p,c), f(p,b)) \\ &< (1-\alpha)r + \alpha d(c,b) \\ &\leq (1-\alpha)r + \alpha r = r \end{aligned}$$

which just says that $f(p,b) \in U(c,r) \subseteq B(c,r)$. It now follows that $f(p,\cdot)$ is a contraction mapping on $B(c,r)$, and hence has a unique fixed point $g(p)$. Here we have used the fact that $B(c,r)$ being a closed subspace of a complete space is complete in its own right. Next, since $g(p) = f(p, g(p)) \in U(c,r)$ we see that g actually takes values in $U(c,r)$. Finally, we show that the mapping g is continuous. Let $p_0 \in P$, we will show that g is continuous at p_0. Applying the third hypothesis with $b = g(p_0)$ we see that for all given $\epsilon > 0$ there exists $\delta > 0$ such that

$$d(p, p_0) < \delta \Longrightarrow d(f(p_0, g(p_0)), f(p, g(p_0)) < (1-\alpha)\epsilon \tag{4.20}$$

Then for $d(p,p_0) < \delta$

$$\begin{aligned} d(g(p_0), g(p)) &= d(f(p_0, g(p_0)), f(p, g(p)) \\ &\leq d(f(p_0, g(p_0)), f(p, g(p_0)) + d(f(p, g(p_0)), f(p, g(p)) \\ &\leq (1-\alpha)\epsilon + \alpha d(g(p_0), g(p)) \end{aligned}$$

by (4.20) and the fact that $f(p,\cdot)$ is a contraction mapping. It follows that $d(g(p_0), g(p)) \leq \epsilon$ and the continuity is proved. ∎

4.4 Extension by Uniform Continuity

In this section we tackle extension by continuity as it is usually called. Actually as we shall see, this is a misnomer.

LEMMA 54 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$ be a uniformly continuous mapping. Let (x_n) be a Cauchy sequence in X. Then $(f(x_n))$ is a Cauchy sequence in Y.*

Proof. Let $\epsilon > 0$. Then there exists $\delta > 0$ such that

$$a, b \in X, d_X(a, b) < \delta \qquad \Rightarrow \qquad d_Y(f(a), f(b)) < \epsilon. \tag{4.21}$$

Since (x_n) is a Cauchy sequence, there exists $N \in \mathbb{N}$ such that

$$p, q > N \qquad \Rightarrow \qquad d_X(x_p, x_q) < \delta.$$

Combining this with (4.21) yields

$$p, q > N \qquad \Rightarrow \qquad d_Y(f(x_p), f(x_q)) < \epsilon.$$

It follows that $(f(x_n)$ is a Cauchy sequence in Y. ∎

THEOREM 55 *Let X and Y be metric spaces and suppose that Y is complete. Let A be a dense subset of X. Let $f : A \longrightarrow Y$ be a uniformly continuous map. Then there is a unique uniformly continuous mapping $\tilde{f} : X \longrightarrow Y$ such that $\tilde{f}(a) = f(a)$ for all $a \in A$.*

Proof. Let $x \in X$. Since A is dense in X there exists a sequence (a_n) in A converging to x. Now, (a_n) is a Cauchy sequence in A and so, by Lemma 4.4, $(f(a_n))$ is a Cauchy sequence in Y. Since Y is complete, this converges to some element y of Y. We will define

$$\tilde{f}(x) = y.$$

We need to prove that $\tilde{f}$ is well-defined. Suppose that (b_n) is another sequence in A converging to x. Then $(f(b_n))$ must converge to some element z of Y. We need to show that $y = z$. To see this we mix the two sequences (a_n) and (b_n) by $x_{2n-1} = a_n$, $x_{2n} = b_n$. The sequence (x_n) converges to x and so $(f(x_n))$ is convergent in Y. But since $(f(x_n))$ has one subsequence converging to y and another converging to z it follows that $y = z$ as required.

If $x \in A$, then we can always take $a_n = x$ for all $n \in \mathbb{N}$. It follows that $\tilde{f}(x) = f(x)$ for all $x \in A$.

Next we show that $\tilde{f}$ is uniformly continuous. Let $\epsilon > 0$. The there exists $\delta > 0$ such that

$$a, b \in A, d(a,b) < \delta \qquad \Rightarrow \qquad d(f(a), f(b)) < \tfrac{1}{2}\epsilon. \tag{4.22}$$

Now, let x and x' be points of X such that $d(x,x') < \frac{1}{2}\delta$. Let us find sequences (a_n) and (a'_n) in A converging to x and x' respectively. Then, there exists $N \in \mathbb{N}$ such that

$$n > N \qquad \Rightarrow \qquad d(a_n, x) < \tfrac{1}{4}\delta \text{ and } d(a'_n, x') < \tfrac{1}{4}\delta. \tag{4.23}$$

An application of the extended triangle inequality, (4.22) and (4.23) now yields

$$n > N \qquad \Rightarrow \qquad d(f(a_n), f(a'_n)) < \tfrac{1}{2}\epsilon.$$

Letting now $n \longrightarrow \infty$ and using the fact that d_Y is continuous on $Y \times Y$ we see that

$$d(\tilde{f}(x), \tilde{f}(x')) \leq \tfrac{1}{2}\epsilon < \epsilon,$$

since the sequences $(f(a_n))$ and $(f(a'_n))$ converge to $\tilde{f}(x)$ and $\tilde{f}(x')$ respectively. This establishes the uniform continuity of $\tilde{f}$.

The final step of the proof is to show that $\tilde{f}$ is unique. Let g be another continuous extension of f. Since A is dense in X and $\tilde{f}$ and g are both continuous functions that agree on A, we can use Proposition 23 to deduce that $g = \tilde{f}$. ■

EXAMPLE We show that in Theorem 55, one cannot replace uniform continuity by continuity. Let $X = [0,1]$, $A =]0,1]$ and $Y = [-1,1]$. Let

$$f(a) = \sin\frac{1}{a}$$

for all $a \in A$. Then all the hypotheses of Theorem 55 are met, except for the uniform continuity of f. The function f is of course continuous on A. We leave it as an exercise to show that f does not extend continuously to X. □

EXAMPLE Let V be a normed vector space. Suppose that T is a continuous linear map from ℓ^1 to V. Define $v_n = T(e_n) \in V$ where e_n denotes the sequence in ℓ^1 that has a 1 in the n-th place and 0 everywhere else. We see that

$$\|v_n\| \leq \|T\|_{\text{op}} \|e_n\|_{\ell^1} = \|T\|_{\text{op}}.$$

Thus

$$\sup_{n\in\mathbb{N}} \|v_n\| \le \|T\|_{\mathrm{op}}. \tag{4.24}$$

Conversely suppose that $C = \sup_{n\in\mathbb{N}} \|v_n\| < \infty$. Let F denote the set of all finitely supported sequences in ℓ^1. Then we can define a map

$$T_0 : F \longrightarrow V$$

by $T_0(\sum t_n e_n) = \sum t_n v_n$. Here the sums involved are finite sums. It is easy to check that $\|T_0(t)\| \le C\|t\|$ for all $t \in F$ and it follows that T_0 is uniformly continuous on the dense subset F of ℓ^1. Thus by Theorem 55 T_0 can be extended to a uniformly continuous mapping T on the whole of ℓ^1. In this particular example this is no big deal, because it is also straightforward to define T directly. For $t \in \ell^1$, we can take

$$T(t) = \sum_{n=1}^{\infty} t_n v_n$$

an absolutely convergent infinite sum in the complete space V. □

EXAMPLE The previous example features two possible approaches to defining an operator T, one direct and one involving extension. There do exist analogous situations where the direct approach is unavailable. For instance let $(f_j)_{j=1}^{\infty}$ be an orthonormal set in ℓ^2. Then, for $t = (t_j) \in F$, we can define $T_0(t) = T_0(\sum t_n e_n) = \sum t_n f_n$. Here we view T_0 as a mapping from F to ℓ^2. A simple calculation

$$\|\sum t_n f_n\|_2^2 = \sum_{m,n} t_m t_n \langle f_m, f_n\rangle = \sum_n t_n^2 \|f_n\|_2^2 = \sum_n t_n^2 = \|t\|_2^2$$

shows that T_0 is an isometry. It follows that T_0 extends to an isometry $T : \ell^2 \longrightarrow \ell^2$. It is important to realise that for general $t \in \ell^2$ the sum

$$\sum_{n=1}^{\infty} t_n f_n$$

does not converge absolutely in ℓ^2. It does converge in ℓ^2 norm, but taking this route is essentially repeating the argument of Theorem 55. □

4.5 Completions

DEFINITION *Let X be a metric space. Then a **completion** (Y, j) of X is a complete metric space Y, together with an isometric inclusion $j : X \longrightarrow Y$ such that $j(X)$ is dense in Y.*

The completion of a metric space is unique in the following sense.

THEOREM 56 *Let (Y, j) and (Z, k) be completions of X. Then there is a surjective isometry $\alpha : Y \longrightarrow Z$ such that $k = \alpha \circ j$.*

Proof. Let us define $\beta : j(X) \longrightarrow Z$ by $\beta(j(x)) = k(x)$. Since j is an injective mapping from X onto $j(X)$, β is well defined. Also β is an isometry since j and k are. Now apply Theorem 55 to define $\alpha : Y \longrightarrow Z$ a continuous map. Obviously we have $k = \alpha \circ j$.

It remains to show that α is an isometry. We consider two mappings from $Y \times Y$ to $\mathbb{R}^+$.

$$\begin{aligned} (y_1, y_2) &\longrightarrow d_Y(y_1, y_2) \\ (y_1, y_2) &\longrightarrow d_Z(\alpha(y_1), \alpha(y_2)) \end{aligned}$$

These mappings are continuous by Theorem 13 (page 25) and since the metric is itself continuous (page 31). Since j and k are isometries, they agree on the subset $j(X) \times j(X)$ of $Y \times Y$. By Proposition 24, $j(X) \times j(X)$ is dense in $Y \times Y$. Finally by Proposition 23 the two mappings agree everywhere on $Y \times Y$. This says that α is an isometry. ■

With the issue of uniqueness of completions out of the way, we deal with the more difficult question of existence.

THEOREM 57 *Every metric space possesses a completion.*

In the proof that we give below, we unashamedly use the completeness of $\mathbb{R}$. We take the point of view that an understanding of $\mathbb{R}$ is needed to define the metric space concept in the first place. There is another proof in the literature using equivalence classes of Cauchy sequences which avoids this issue.

Proof. We will assume that X is a bounded metric space and construct a completion. At the end of the proof we will discuss the modifications that are necessary to dispense with the boundedness hypothesis.

Let $j : X \longrightarrow C(X, \mathbb{R})$ be given by

$$(j(x_1))(x_2) = d(x_1, x_2).$$

One needs to stand back a moment to ponder this notation. Since $x_1 \in X$, $j(x_1) \in C(X, \mathbb{R})$, that is, $j(x_1)$ is itself a mapping from X to $\mathbb{R}$. For $x_2 \in X$, the notation $(j(x_1))(x_2) \in \mathbb{R}$ then stands for the image of x_2 by the mapping $j(x_1)$.

It follows from the continuity of the metric (page 31) that $j(x_1)$ is continuous. Since X is a bounded metric space, $j(x_1)$ is bounded. Next we observe

$$\|j(x_1) - j(x_2)\| = \sup_{x_3 \in X} |d(x_1, x_3) - d(x_2, x_3)|.$$

Two applications of the triangle inequality show that

$$|d(x_1, x_3) - d(x_2, x_3)| \le d(x_1, x_2) \qquad \forall x_3 \in X,$$

while taking $x_3 = x_2$ shows that

$$\sup_{x_3 \in X} |d(x_1, x_3) - d(x_2, x_3)| \ge d(x_1, x_2).$$

Thus j is an isometry. Since $C(X, \mathbb{R})$ is complete, the closed subset $\mathrm{cl}(j(X))$ is also complete by Proposition 47 (page 76). Let $Y = \mathrm{cl}(j(X))$ and consider j just as a mapping from X to Y. Then (Y, j) is a completion of X.

In the case that X is unbounded, select any point x_0 from X. Now set $(j(x_1))(x_2) = d(x_1, x_2) - d(x_0, x_2)$. The key observation is that since

$$|d(x_1, x_2) - d(x_0, x_2)| \le d(x_0, x_1)$$

the function $j(x_1)$ is actually bounded. The rest of the proof follows the same line. ■

4.6 Extension of Continuous Functions

LEMMA 58 *Let X be a metric space. Let E_0 and E_1 be disjoint closed subsets of X. Then there exists a continuous function $f : X \longrightarrow [0, 1]$ such that $f^{-1}(\{0\}) = E_0$ and $f^{-1}(\{1\}) = E_1$.*

Proof. We define

$$f(x) = \frac{\mathrm{dist}_{E_0}(x)}{\mathrm{dist}_{E_0}(x) + \mathrm{dist}_{E_1}(x)}.$$

We observe that $x \longrightarrow \mathrm{dist}_{E_0}(x)$ and $x \longrightarrow \mathrm{dist}_{E_1}(x)$ are both continuous functions on X. Furthermore $\mathrm{dist}_{E_0}(x) + \mathrm{dist}_{E_1}(x) > 0$ for all $x \in X$ by Proposition 26 (page 36) and since E_0 and E_1 are disjoint. It follows that $f : X \longrightarrow [0, 1]$ is continuous. Clearly $f(x) = 0$ if and only if $\mathrm{dist}_{E_0}(x) = 0$ which occurs if and only if $x \in E_0$ again by Proposition 26. Similarly $f(x) = 1$ if and only if $x \in E_1$. ■

We do not need it right at the moment, but there is a simple looking Corollary of Lemma 58 that is difficult to establish without using the distance to a subset function.

COROLLARY 59 *Let X be a metric space. Let E_0 and E_1 be disjoint closed subsets of X. Then there exist U_0 and U_1 disjoint open subsets of X such that $E_j \subseteq U_j$ for $j = 0, 1$.*

Proof. Let f be the function of Lemma 58 and simply set $U_0 = f^{-1}([0, \frac{1}{3}[)$ and $U_1 = f^{-1}(]\frac{2}{3}, 1])$. Of course, U_0 is open in X since $[0, \frac{1}{3}[$ is (relatively) open in $[0, 1]$. Similarly for U_1. The sets U_0 and U_1 obviously satisfy the remaining properties. ■

With this diversion out of the way, we can now continue to the main order of business.

THEOREM 60 (TIETZE EXTENSION THEOREM) *Let X be a metric space and let E be a closed subset of X. Let $g : E \longrightarrow [-1, 1]$ be a continuous mapping. Then there exists a continuous mapping $f : X \longrightarrow [-1, 1]$ extending g. Explicitly, this means that*

$$f(x) = g(x) \qquad \forall x \in E.$$

Proof. Let us denote $g_0 = g$. We start by constructing a continuous map f_0. Let $E_0 = g_0^{-1}([-1, -\frac{1}{3}])$ and let $E_1 = g_0^{-1}([\frac{1}{3}, 1])$. Since E is closed in X and E_0 and E_1 are closed in E, it follows that E_0 and E_1 are also closed in X. By a straightforward variant of Lemma 58, there is a mapping $f_0 : X \longrightarrow [-\frac{1}{3}, \frac{1}{3}]$ such that $f_0^{-1}(\{-\frac{1}{3}\}) = E_0$ and $f_0^{-1}(\{\frac{1}{3}\}) = E_1$. We claim that

$$|g_0(x) - f_0(x)| \leq \tfrac{2}{3} \qquad \forall x \in E.$$

There are three cases.

- $x \in E_0$. Then $f_0(x) = -\frac{1}{3}$ and $g_0(x) \in [-1, -\frac{1}{3}]$.
- $x \in E_1$. Then $f_0(x) = \frac{1}{3}$ and $g_0(x) \in [\frac{1}{3}, 1]$.
- $x \in E \setminus (E_0 \cup E_1)$. Then $f_0(x), g_0(x) \in]-\frac{1}{3}, \frac{1}{3}[$.

Now define

$$g_1(x) = g_0(x) - f_0(x) \qquad \forall x \in E.$$

Then $\|g_1\|_\infty \leq \frac{2}{3}$. We repeat the above argument at $\frac{2}{3}$ scale to define f_1. We then proceed inductively in the obvious way. We thus obtain continuous functions $f_n : X \longrightarrow [-2^n \cdot 3^{-(n+1)}, 2^n \cdot 3^{-(n+1)}]$ satisfying

$$\|g_n - f_n\|_\infty \leq 2^n \cdot 3^{-n},$$

where $g_n = g_{n-1} - f_{n-1}|_E$.

It is easy to see that $\|g_n\|_\infty$ tends to 0 as $n \longrightarrow \infty$. Hence we may write g as the telescoping sum

$$g = \sum_{n=1}^{\infty}(g_{n-1} - g_n) = \sum_{n=0}^{\infty} f_n|_E.$$

But the function f given by the uniformly convergent sum

$$f = \sum_{n=0}^{\infty} f_n,$$

also converges to a continuous function taking values in $[-1, 1]$ and evidently

$$f|_E = \sum_{n=0}^{\infty} f_n|_E = g,$$

as required. ■

4.7 Baire's Theorem

The following result has a number of key applications that cannot be approached in any other way.

THEOREM 61 (BAIRE'S CATEGORY THEOREM) *Let X be a complete metric space. Let A_k be a sequence of closed subsets of X with $\operatorname{int}(A_k) = \emptyset$. Then*

$$X \setminus \bigcup_{k=1}^{\infty} A_k \text{ is dense in } X. \tag{4.25}$$

In particular if X is nonempty we have

$$\bigcup_{k=1}^{\infty} A_k \neq X.$$

Proof. We suppose that (4.25) fails. Then there exist $x_0 \in X$ and $t > 0$ such that

$$U(x_0, t) \subseteq \bigcup_{k=1}^{\infty} A_k. \tag{4.26}$$

We construct a sequence (x_n) in X. Let $t_0 = \frac{1}{2}t$. We choose $x_1 \in X \setminus A_1$ such that $d(x_1, x_0) < t_0$. This is possible since otherwise we would have $U(x_0, t_0) \subseteq A_1$ contradicting the fact that $\operatorname{int}(A_1) = \emptyset$. Now define $t_1 = \min(\frac{1}{2}t_0, \frac{1}{4}\operatorname{dist}_{A_1}(x_1)) > 0$. Next find $x_2 \in X \setminus A_2$ such that $d(x_2, x_1) < t_1$. If this is not possible then $U(x_1, t_1) \subseteq A_2$ contradicting the hypothesis $\operatorname{int}(A_2) = \emptyset$. Now define $t_2 = \min(\frac{1}{2}t_1, \frac{1}{4}\operatorname{dist}_{A_2}(x_2)) > 0$ and then find $x_3 \in X \setminus A_3$ such that $d(x_3, x_2) < t_2$. Continuing in this way, we obtain a sequence (x_n) in X and a sequence (t_n) of strictly positive reals such that

$$\begin{aligned} t_n &= \min(\tfrac{1}{2}t_{n-1}, \tfrac{1}{4}\operatorname{dist}_{A_n}(x_n)) > 0 \qquad n = 1, 2, \ldots \\ x_n &\notin A_n \qquad n = 1, 2, \ldots \\ d(x_{n+1}, x_n) &< t_n \qquad n = 0, 1, 2, \ldots \end{aligned}$$

Since $d(x_{n+1}, x_n) < t_n \leq 2^{-n}t_0 = 2^{-n-1}t$, we see that (x_n) is a Cauchy sequence. The detailed justification of this is similar to one found in the proof of Theorem 52. It follows from the completeness of X that (x_n) converges to some limit point $x \in X$.

We next show that $x \notin A_k$ for all $k \in \mathbb{N}$. Indeed, by passing to the limit in the extended triangle inequality, we obtain for each $k = 1, 2, \ldots$

$$d(x, x_k) \leq \sum_{n=k}^{\infty} d(x_{n+1}, x_n) < \sum_{n=k}^{\infty} t_n \leq \sum_{n=k}^{\infty} 2^{-2-n+k} \operatorname{dist}_{A_k}(x_k) = \tfrac{1}{2} \operatorname{dist}_{A_k}(x_k).$$

It follows that $x \notin A_k$, as required.

Finally, following the same line, we find

$$d(x, x_0) \leq \sum_{n=0}^{\infty} d(x_{n+1}, x_n) < \sum_{n=0}^{\infty} t_n \leq \sum_{n=0}^{\infty} 2^{-1-n} t = t,$$

so that $x \in U(x_0, t)$. This contradicts (4.26) and completes the proof of the Theorem. ■

4.8 Complete Normed Spaces

Complete normed spaces are also called ***Banach Spaces*** . The following result is very basic and could have been left as an exercise for the reader.

THEOREM 62 *Let V and W be normed vector spaces and suppose that W is complete. Then $\mathcal{CL}(V, W)$ is a complete normed vector space with the operator norm.*

Proof. Let (T_n) be a cauchy sequence in $\mathcal{CL}(V, W)$. First consider a fixed point v of V. Then we have

$$\|T_p(v) - T_q(v)\| \leq \|T_p - T_q\|_{\text{op}} \|v\|.$$

It follows easily that $(T_n(v))$ is a Cauchy sequence in W. Since W is complete, there is an element $w \in W$ such that $T_n(v) \longrightarrow w$ as $n \longrightarrow \infty$. We now allow v to vary and define a mapping T by $T(v) = w$. We leave the reader to check that the mapping T is a *linear* mapping from V to W. Now let $\epsilon > 0$ then by the Cauchy condition there exist $N \in \mathbb{N}$ such that

$$p, q > N \qquad \Rightarrow \qquad \|T_p - T_q\|_{\text{op}} \leq \epsilon$$

or equivalently that

$$p, q > N, v \in V \qquad \Rightarrow \qquad \|T_p(v) - T_q(v)\| \leq \epsilon \|v\|. \tag{4.27}$$

Now let q tend to infinity in (4.27). We find that

$$p > N, v \in V \qquad \Rightarrow \qquad \|T_p(v) - T(v)\| \le \epsilon \|v\|,$$

or equivalently that (T_p) converges to T in $\mathcal{CL}(V, W)$. This also shows that $T \in \mathcal{CL}(V, W)$. ■

Rather more interesting is the following Proposition.

PROPOSITION 63 *Let V be a complete normed vector space and N a closed linear subspace. Then the quotient space $Q = V/N$ is a complete normed space with the norm defined by (3.38).*

Proof. This Proposition is included because it illustrates a method of proof not seen elsewhere in these notes. The key idea is the use of ***rapidly convergent subsequences*** . We denote by π the canonical projection mapping $\pi : V \longrightarrow Q$.

Let (q_n) be a Cauchy sequence in Q. Applying the Cauchy condition with $\epsilon = 2^{-k}$, we find n_k such that

$$\ell, m \ge n_k \quad \Rightarrow \quad \|q_\ell - q_m\| < 2^{-k}. \tag{4.28}$$

In particular, taking $l = n_k$, $m = n_{k+1}$, we have

$$\|q_{n_k} - q_{n_{k+1}}\| < 2^{-k}. \tag{4.29}$$

We now proceed to find lifts of the q_{n_k}. Let v_1 be any element of V with $\pi(v_1) = q_{n_1}$. Now by (4.29) and the definition of the quotient norm, there exists $u_{k+1} \in V$ such that $\|u_{k+1}\| < 2^{-k}$ and $\pi(u_{k+1}) = q_{n_{k+1}} - q_{n_k}$. We now define $v_2 = v_1 + u_2$, $v_3 = v_2 + u_3$, etc., so that we now have

$$\|v_k - v_{k+1}\| < 2^{-k} \tag{4.30}$$

and $\pi(v_k) = q_{n_k}$ for $k \in \mathbb{N}$. It is easy see to that (4.30) forces (v_k) to be a Cauchy sequence in V as in the proof of the Contraction Mapping Theorem. Since V is complete, we can infer the existence of $v \in V$ such that (v_k) converges to v.

Since π is continuous, it follows that the subsequence (q_{n_k}) converges to the element $\pi(v)$ of Q. Furthermore, we can have the estimate

$$\|q_{n_k} - \pi(v)\| \le 2^{-(k-1)}. \tag{4.31}$$

Combining this with (4.28) we find

$$\ell \ge n_k \quad \Rightarrow \quad \|q_\ell - \pi(v)\| < 3 \cdot 2^{-k},$$

from which it follows that the original sequence (q_n) converges to $\pi(v)$. ■

THEOREM 64 (OPEN MAPPING THEOREM) *Let U and V be complete normed spaces and let $T : U \longrightarrow V$ be a continuous surjective linear map. Then there is a constant $\epsilon > 0$ such that for every $v \in V$ with $\|v\| \leq 1$, there exist $u \in U$ with $\|u\| \leq \epsilon$ such that $T(u) = v$.*

The reason for the terminology is that the statement that T is an open mapping (see page 142 for the definition) is equivalent to the conclusion of the Theorem.

Proof. There are two separate ideas in the proof. The first is to use the Baire Category Theorem and the second involves iteration.

Let B_n denote $\{u : u \in U, \|u\| \leq n\}$, the closed n-ball in U. Then, since T is onto, we have

$$V = \bigcup_{n \in \mathbb{N}} T(B_n).$$

We can't use this directly in the Baire Category Theorem because we don't know that the $T(B_n)$ are closed. We take the easiest way around this difficulty and write simply

$$V = \bigcup_{n \in \mathbb{N}} \mathrm{cl}(T(B_n)).$$

By the Baire Category Theorem (page 90), there exists $n \in \mathbb{N}$ such that $\mathrm{cl}(T(B_n))$ has nonempty interior. This means that there exists $v \in V$ and $t > 0$ such that $U_V(v, t) \subseteq \mathrm{cl}(T(B_n))$. By symmetry, it follows that $U_V(-v, t) \subseteq \mathrm{cl}(T(B_n))$. We claim that $U_V(0_V, t) \subseteq \mathrm{cl}(T(B_n))$. Let $w \in U_V(0_V, t)$. Then, we can find two sequences (x_k) and (y_k) in B_n such that $(T(x_k))$ converges to $w + v$ and $(T(y_k))$ converges to $w - v$. It follows that the sequence $(T(\frac{1}{2}(x_k + y_k)))$ converges w. This establishes the claim.

Now, let v be a generic element of V with $\|v\| < t$. Then $v \in \mathrm{cl}(T(B_n))$. Hence, there exists $u_0 \in B_n$ such that $\|v - T(u_0)\| < \frac{1}{2}t$. We repeat the argument, but rescaled by a factor of $\frac{1}{2}$ and applied to $v - T(u_0)$. Thus, there is an element $u_1 \in U$ with $\|u_1\| < \frac{1}{2}n$ and such that $\|v - T(u_0) - T(u_1)\| < \frac{1}{4}t$. Continuing in this way leads to elements $u_k \in U$ with $\|u_k\| < n\,2^{-k}$ such that

$$\|v - \sum_{k=0}^{\ell} T(u_k)\| < t\,2^{-\ell-1}.$$

Using now the fact that U is complete (the completeness of V is needed to apply Baire's Theorem), we find that $T(u) = v$ where

$$u = \sum_{k=0}^{\infty} u_k \in U$$

is given by an absolutely convergent series and has norm bounded by $2n$. Rescaling gives the required result. ■

COROLLARY 65 *Let V be a vector space with two norms $\| \ \|_1$ and $\| \ \|_2$, both of which make V complete. Suppose that there is a constant C such that*

$$\|v\|_2 \leq C\|v\|_1 \qquad \forall v \in V.$$

Then $\| \ \|_1$ and $\| \ \|_2$ are equivalent norms.

Proof. Apply the Open Mapping Theorem in case that T is the identity mapping from $(V, \| \ \|_1)$ to $(V, \| \ \|_2)$. ■

It is possible to construct an infinite dimensional vector space with two *incomparable* norms both of which render it complete.

PROPOSITION 66 *Let V_1 and V_2 be vector spaces and suppose that $\| \ \|$ is a norm on $V = V_1 \oplus V_2$ which renders V complete. Let P_1 and P_2 be the linear projection operators corresponding to the direct sum. Then the following are equivalent:*

- *P_1 and P_2 are continuous.*
- *V_1 and V_2 are closed in V.*

Proof. Suppose that P_1 is continuous and that (v_n) is a sequence in V_1 which converges to some element v of V. Then $(P_1(v_n))$ converges to $P_1(v)$. But, since $P_1(v_n) = v_n$ and by the uniqueness of limits (Proposition 6), $v = P_1(v)$ or equivalently $v \in V_1$. This shows that V_1 is closed in V. Of course since $P_2 = I - P_1$, if P_1 is continuous, so is P_2 and similarly we find that V_2 is closed.

The converse is much harder. We can identify V_2 with the quotient space V/V_1. Since V_1 is closed, we have a natural quotient norm

$$\|v_2\|_Q = \inf_{v_1 \in V_1} \|v_1 + v_2\|$$

on V_2 as well as the restriction of the given norm. We see that V_2 is complete in both norms, by Proposition 63 and since V_2 is closed in V. Clearly, $\|v_2\|_Q \leq \|v_2\|$

for all $v_2 \in V_2$ so that the two norms are comparable. Hence, by Corollary 65, we see that there exists a finite constant C such that

$$\|v_2\| \leq C \inf_{v_1 \in V_1} \|v_1 + v_2\| \qquad \forall v_2 \in V_2. \tag{4.32}$$

A moment's thought shows that (4.32) is equivalent to

$$\|P_2(v)\| \leq C\|v\| \qquad \forall v \in V,$$

so that P_2 is continuous as required. One shows that P_1 is continuous by a similar argument, or by applying the formula $P_1 = I - P_2$. ■

PROPOSITION 67 *Let V_1 and V_2 be complete normed spaces and suppose that $\| \ \|$ is a norm on $V = V_1 \oplus V_2$ which renders V complete and agrees with the original norms given on V_1 and V_2. Then $\| \ \|$ is equivalent with any of the following p-standard norms on V*

$$\|v_1 + v_2\|_p = \left(\|v_1\|_{V_1}^p + \|v_2\|_{V_2}^p\right)^{\frac{1}{p}} \qquad 1 \leq p < \infty$$

or

$$\|v_1 + v_2\|_\infty = \max(\|v_1\|_{V_1}, \|v_2\|_{V_2}) \qquad p = \infty.$$

Proof. Clearly, all the given p-standard norms on V are mutually equivalent, so it suffices to work with the $p = 1$ norm. Then obviously

$$\|v_1 + v_2\| \leq \|v_1\|_{V_1} + \|v_2\|_{V_2} = \|v_1 + v_2\|_1.$$

But both $\| \ \|$ and $\| \ \|_1$ are complete norms on V. Since they are comparable, an application of Corollary 65 shows that they are equivalent. ■

We remark that Propositions 66 and 67 extend in an obvious way to finite direct sums.

5

Compactness

Compactness is one of the most important concepts in mathematical analysis. It is a topological form of finiteness. The formal definition is quite involved.

DEFINITION *A metric space X is said to be* ***compact*** *if, whenever V_α are open sets for every α in some index set I such that $\cup_{\alpha\in I} V_\alpha = X$, then there exists a finite subset $F \subseteq I$ such that $\cup_{\alpha\in F} V_\alpha = X$.*

PROPOSITION 68 *Every compact metric space is bounded.*

Proof. Let X be a compact metric space. If X is empty, then it is bounded. If not, select a point $x_0 \in X$. Now we have

$$X = \bigcup_{n\in\mathbb{N}} U(x_0, n).$$

By compactness, there is a finite subset $F \subseteq \mathbb{N}$ such that

$$X = \bigcup_{n\in F} U(x_0, n).$$

Let N be the largest integer in F. Such an integer exists because F is finite and non-empty. Then $X = U(x_0, N)$ and it follows that X is bounded. ■

PROPOSITION 69 *Every finite metric space is compact.*

Proof. Let V_α be open subsets of X for every α in some index set I such that $\cup_{\alpha\in I}V_\alpha = X$. Since X is finite, let us enumerate it as $X = \{x_1, x_2, \ldots, x_n\}$. For each $j = 1, \ldots, n$ there exists $\alpha_j \in I$ such that $x_j \in V_{\alpha_j}$. Let us set $F = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. Then clearly $\cup_{\alpha\in F}V_\alpha = X$. ■

5.1 Compact Subsets

We also use the term compact to describe subsets of a metric space.

DEFINITION *A subset K of a metric space X is **compact** iff it is compact when viewed as a metric space in the restriction metric.*

It follows immediately from the definition and Proposition 68 that compact subsets are necessarily bounded.

We need a direct way to describe which subsets are compact.

PROPOSITION 70 *Let X be a metric space and let $K \subseteq X$. Then the following two conditions are equivalent.*

- *K is a compact subset of X.*
- *Whenever V_α are open sets of X for every α in some index set I such that $\cup_{\alpha\in I}V_\alpha \supseteq K$, then there exists a finite subset $F \subseteq I$ such that $\cup_{\alpha\in F}V_\alpha \supseteq K$.*

Proof. Suppose that the first statement holds. Let us assume that I is an index set which labels open subsets V_α of X such that $\cup_{\alpha\in I}V_\alpha \supseteq K$. Then, by Theorem 30 (page 41) the subsets $K \cap V_\alpha$ are open subsets of K. We clearly have $\cup_{\alpha\in I}(K \cap V_\alpha) = K$. Since K is compact, by the definition, there exists a finite subset $F \subseteq I$ such that $\cup_{\alpha\in F}(K \cap V_\alpha) = K$. It follows immediately that $\cup_{\alpha\in F}V_\alpha \supseteq K$.

For the converse, we suppose that the second condition holds. We need to show that K is compact. Let U_α be open subsets of K for every $\alpha \in I$ such that $\cup_{\alpha\in I}U_\alpha = K$. The, according to Theorem 30 there exist open subsets V_α of X such that $K \cap V_\alpha = U_\alpha$. We clearly have $\cup_{\alpha\in I}V_\alpha \supseteq K$ so that we may apply the second condition to infer the existence of F finite $F \subseteq I$ such that $\cup_{\alpha\in F}V_\alpha \supseteq K$. But it is easy to verify that

$$K \cap \left(\bigcup_{\alpha\in F} V_\alpha\right) = \bigcup_{\alpha\in F}(K \cap V_\alpha) = \bigcup_{\alpha\in F} U_\alpha,$$

and it follows that $\cup_{\alpha \in F} U_\alpha = K$ as required. We have just verified the compactness of K as a metric space. ■

PROPOSITION 71 *Compact subsets are necessarily closed.*

Proof. Let K be a compact subset of a metric space X. Let us suppose that K is not closed. We will provide a contradiction. Let $x \in \mathrm{cl}(K) \setminus K$. Then $x \notin \mathrm{int}(X \setminus K)$. Thus for every $\epsilon > 0$,

$$U(x, \epsilon) \cap K \neq \emptyset \tag{5.1}$$

holds. On the other hand, since $x \notin K$ we have

$$K \subseteq \bigcup_{n \in \mathbb{N}} \left(X \setminus B(x, \tfrac{1}{n})\right),$$

since for every $y \in K$, $d(x, y) > 0$. Since $B(x, t)$ is closed, $X \setminus B(x, t)$ is open. We therefore apply the compactness criterion for subsets to find F a finite subset of $\mathbb{N}$, such that

$$K \subseteq \bigcup_{n \in F} \left(X \setminus B(x, \tfrac{1}{n})\right).$$

If K is empty then K is closed and we are done. If not, F is a non-empty finite subset of $\mathbb{N}$ which therefore possesses a maximal element N. It follows that K is disjoint from $B(x, \frac{1}{N})$ contradicting (5.1). ■

THEOREM 72 *Every closed subset of a compact metric space is compact.*

Proof. Let X be a compact metric space. Let K be a closed subset of X. Let V_α be open sets of X for every α in some index set I such that $\cup_{\alpha \in I} V_\alpha \supseteq K$. We will show the existence of a finite subset F of I such that $\cup_{\alpha \in F} V_\alpha \supseteq K$.

To do this, we extend the index set by one index. Let $J = I \cup \{\beta\}$. Define $V_\beta = X \setminus K$ an open subset of X since K is closed and by Theorem 8 (page 18). We have

$$\begin{aligned} \bigcup_{\alpha \in J} V_\alpha &= \left(\bigcup_{\alpha \in I} V_\alpha\right) \cup (X \setminus K) \\ &\supseteq K \cup (X \setminus K) = X. \end{aligned}$$

By the compactness of X there is a finite subset G of J such that $\cup_{\alpha \in G} V_\alpha = X$. We can assume without loss of generality that $\beta \in G$ and define $F = G \cap I$. It is then clear that F is a finite subset of I and that

$$\bigcup_{\alpha \in F} V_\alpha \supseteq K.$$

■

5.2 The Finite Intersection Property

DEFINITION *Let X be a set and let C_α be a subset of X for every α in some index set I. Then we say that the family $(C_\alpha)_{\alpha \in I}$ has the* ***finite intersection property*** *iff*

$$\bigcap_{\alpha \in F} C_\alpha \neq \emptyset$$

for every finite subset F of I.

The following Proposition is an immediate consequence of the definition of compactness and Theorem 8 (page 18).

PROPOSITION 73 *Let X be a metric space. Then the following two statements are equivalent.*

- *X is compact.*
- *Whenever $(C_\alpha)_{\alpha \in I}$ is a family of closed subsets of X having the finite intersection property, then $\cap_{\alpha \in I} C_\alpha \neq \emptyset$.*

This reformulation of compactness is often very useful.

5.3 Other Formulations of Compactness

In this section we look at some conditions which are equivalent to or very nearly equivalent to compactness. The first of these is countable compactness. Countable compactness is a technical device and is never used in practice.

DEFINITION *A metric space X is said to be* ***countably compact*** *if, whenever V_n are open sets for every $n \in \mathbb{N}$ such that $\cup_{n\in\mathbb{N}} V_n = X$, then there exists a finite subset $F \subseteq \mathbb{N}$ such that $\cup_{n\in F} V_n = X$.*

We say that a subset K of a metric space is countably compact if it is countably compact when viewed as a metric space in its own right. There is a formulation of the concept of countably compact subset entirely analogous to that given for compact subset in Proposition 70 (page 97).

Clearly, a metric space that is compact is also countably compact. The converse is true in the context of metric spaces, but false in the setting of topological spaces. Towards the converse, we have the following Proposition.

PROPOSITION 74 *Let X be a separable, countably compact metric space. Then X is compact.*

Proof. Let V_α be open sets of X for every α in some index set I satisfying $\cup_{\alpha\in I} V_\alpha = X$. We aim to find a finite subset F of I such that $\cup_{\alpha\in F} V_\alpha = X$. Observe that if for some $\alpha \in I$, we have $V_\alpha = X$, then we are done. Thus we may assume that for each $\alpha \in I$, the set $X \setminus V_\alpha$ is non-empty.

Let S be a countable dense subset of X. We observe that $S \cap V_\alpha$ is dense in V_α. We now use the proof of Theorem 29 (page 38), which shows that

$$V_\alpha = \bigcup_{s\in S\cap V_\alpha} U(s, \tfrac{1}{2}\operatorname{dist}_{X\setminus V_\alpha}(s)). \tag{5.2}$$

Now, let us define the subset Q_α of $S \times \mathbb{Q}$ by

$$Q_\alpha = \{(s,t); s \in S, t \in \mathbb{Q}, t > 0, U(s,t) \subseteq V_\alpha\} \tag{5.3}$$

Then, by (5.2) we have

$$V_\alpha = \bigcup_{(s,t)\in Q_\alpha} U(s,t). \tag{5.4}$$

We also define $Q = \cup_{\alpha\in I} Q_\alpha$. Since Q is a subset of $S \times \mathbb{Q}$ it is necessarily countable. Then we have by (5.4)

$$X = \bigcup_{(s,t)\in Q} U(s,t). \tag{5.5}$$

The key idea of the proof is to replace the union $\cup_{\alpha \in I} V_\alpha$ with the countable union $\cup_{(s,t)\in Q} U(s,t)$. We now apply the countable compactness of X to (5.5). We obtain a finite subset $R \subset Q$ such that

$$X = \bigcup_{(s,t)\in R} U(s,t). \tag{5.6}$$

For each $(s,t) \in R$ there exists $\alpha \in I$ such that $(s,t) \in Q_\alpha$. Let F be the finite subset of I having these α as members. Then by (5.3) we have

$$\bigcup_{(s,t)\in R} U(s,t) \subseteq \bigcup_{\alpha \in F} V_\alpha. \tag{5.7}$$

Finally, combining (5.6) and (5.7) gives the desired conclusion. ∎

While countable compactness is merely a means to an end, sequential compactness is a very useful tool. It is equivalent to compactness in metric spaces (but not in topological spaces) and can be used as a replacement for compactness in nearly all situations.

DEFINITION *Let X be a metric space. Then X is **sequentially compact** iff every sequence (x_n) in X possesses a convergent subsequence. A subset of a metric space is sequentially compact iff it is a sequentially compact metric space in the restriction metric.*

PROPOSITION 75 *Every compact metric space is sequentially compact.*

Proof. Let X be a compact metric space and let (x_n) be a sequence of points of X. Let $T_m = \{x_n; n \geq m\}$ for $m \in \mathbb{N}$. Then the closed sets $\mathrm{cl}(T_m)$ clearly have the finite intersection property. Hence $\cap_{m\in\mathbb{N}} \mathrm{cl}(T_m) \neq \emptyset$. Let x be a member of this set. We construct a subsequence of (x_n) that converges to x. Let (ϵ_n) be a sequence of strictly positive real numbers decreasing to zero. Then we define the natural subsequence (n_k) inductively. Since $x \in \mathrm{cl}(T_1)$, we choose $n_1 \in \mathbb{N}$ such that $d(x, x_{n_1}) < \epsilon_1$. Now assuming that n_k is already defined, we use the fact that $x \in \mathrm{cl}(T_{n_k+1})$ to find $n_{k+1} > n_k$ and such that $d(x, x_{n_{k+1}}) < \epsilon_{k+1}$. It is easy to see that the subsequence (x_{n_k}) converges to x. Since (x_n) was an arbitrary sequence of points of X it follows that X is sequentially compact. ∎

LEMMA 76 *Let (x_n) be a sequence in the cell $[a,b]^d$ in $\mathbb{R}^d$. Then (x_n) possesses a subsequence which converges to some point of $\mathbb{R}^d$.*

Proof. We define the natural subsequence (n_k) inductively. Let $n_1 = 1$. Let $C_1 = [a,b]^d$ the original cell. Let $c = \frac{1}{2}(a+b)$. The we divide up $[a,b]$ as $[a,c] \cup [c,b]$. Taking the n-th product, this divides the cell C_1 up into 2^d cells with half the linear size of C_1. We select one of these cells C_2 with the property that the set

$$R_2 = \{n; n \in \mathbb{N}, n > n_1, x_n \in C_2\}$$

is infinite. It cannot happen that all of the 2^d cells fail to have this property, for then the set $\{n; n \in \mathbb{N}, n > n_1\}$ would be finite. We choose $n_2 \in R_2$.

To understand the general step of the inductive process, suppose that C_k, R_k and $n_k \in R_k$ have been chosen. Then as before we divide C_k into 2^d cells of half the linear size. We select one of these cells C_{k+1} with the property that

$$R_{k+1} = \{n; n \in R_k, n > n_k, x_n \in C_{k+1}\}$$

is infinite. We choose $n_{k+1} \in R_{k+1}$.

Let $\epsilon > 0$. Then there exists $K \in \mathbb{N}$ such that $\operatorname{diam}(C_K) < \epsilon$. It follows that

$$\begin{aligned} p, q \geq K \qquad &\Rightarrow \qquad n_p, n_q \in R_K, \\ &\Rightarrow \qquad d(x_{n_p}, x_{n_q}) < \epsilon. \end{aligned}$$

In words, this just says that the subsequence (x_{n_k}) is a Cauchy sequence. But since $\mathbb{R}^d$ is complete, it will converge to some point of $\mathbb{R}^d$. ■

THEOREM 77 (BOLZANO–WEIERSTRASS THEOREM) *Every closed bounded subset of $\mathbb{R}^d$ is sequentially compact.*

Proof. Let K be a closed bounded subset of $\mathbb{R}^d$ and suppose that (x_n) is a sequence of points of K. Since K is bounded it is contained in some cell $[a,b]^d$ of $\mathbb{R}^d$. Then according to Lemma 76, (x_n) possesses a subsequence (x_{n_k}) convergent to some point x of $\mathbb{R}^d$. But since K is closed, it follows that $x \in K$. This shows that K is sequentially compact. ■

THEOREM 78 (HEINE–BOREL THEOREM) *A subset K of $\mathbb{R}^d$ is compact iff it is closed and bounded.*

Proof. We have already seen that a compact subset of a metric space is necessarily closed and bounded. It therefore remains only to show that if K is closed and bounded then it is compact. Since $\mathbb{R}^d$ is separable, it follows from Theorem 27 (page 38) that K is also. Hence by Proposition 74 it is enough to show that K is countably compact. We will establish the condition for a subset to be countably compact analogous to that given by Proposition 70 (page 97). Thus, let V_k be open subsets of $\mathbb{R}^d$ for $k \in \mathbb{N}$. Suppose that $\cup_{k\in\mathbb{N}} V_k \supseteq K$. We will show that there exists a finite subset F of $\mathbb{N}$ such that

$$\bigcup_{k\in F} V_k \supseteq K. \tag{5.8}$$

Suppose not. Then, for every k, the set $F = \{1, 2, \ldots, k\}$ fails to satisfy (5.8). So, there is a point $x_k \in K$ with

$$x_k \notin \bigcup_{\ell=1}^{k} V_\ell. \tag{5.9}$$

Now by the Bolzano–Weierstrass Theorem, the sequence (x_k) has a subsequence which converges to some element $x \in K$. By hypothesis, there exists $m \in \mathbb{N}$ such that $x \in V_m$. Since V_m is open, some tail of the subsequence lies entirely inside V_m. This follows from Proposition 7 (page 18). Therefore, there exists $k \in \mathbb{N}$ with $k \geq m$ and $x_k \in V_m$. This contradiction with (5.9) establishes the result. ■

COROLLARY 79 *Every closed bounded cell*

$$\prod_{j=1}^{d} [a_j, b_j]$$

in $\mathbb{R}^d$ is compact.

EXAMPLE The star space X based on an *infinite* set S provides an example of a complete bounded metric space that is *not* compact. See page 59 for the definition of the star space and page 78 for the proof of completeness. Let (s_n) be a sequence of distinct elements of S. Then define a sequence (x_n) of X by $x_n = 0\langle c\rangle + 1\langle s_n\rangle$ where c denotes the centre of X. Then it is immediate that

$$d_X(x_m, x_n) = \begin{cases} 2 & \text{if } m \neq n \\ 0 & \text{if } m = n \end{cases}$$

It follows that (x_n) possesses no convergent subsequence. Thus X is not sequentially compact. □

EXAMPLE Another example of a bounded complete space that is not compact is the unit ball of ℓ^1. The sequence of coordinate vectors (e_n) does not possess a convergent subsequence. □

EXAMPLE The orthogonal groups provide examples of interesting compact spaces. An $n \times n$ matrix U is said to be ***orthogonal*** iff $U'U = I$. Here we have denoted U' the transpose of U. The set of all orthogonal $n \times n$ matrices is usually denoted $O(n)$. It is well known to be a group under matrix multiplication. We view $O(n)$ as a subset of the vector space $M(n, n, \mathbb{R})$ of all $n \times n$ matrices which is a n^2 dimensional real vector space. The equations $U'U = I$ can be rewritten as

$$\sum_{\ell=1}^{n} u_{\ell j} u_{\ell k} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases} \tag{5.10}$$

showing that $O(n)$ is a closed subset of $M(n, n, \mathbb{R})$. The first case in (5.10) can again be rewritten as

$$\sum_{\ell=1}^{n} u_{\ell j}^2 = 1 \qquad j = 1, 2, \ldots, n$$

showing that $|u_{\ell j}| \leq 1$ for all $j = 1, 2, \ldots, n$ and all $\ell = 1, 2, \ldots, n$. Thus $O(n)$ is a bounded subset of $M(n, n, \mathbb{R})$. The Heine–Borel Theorem can now be used to conclude that $O(n)$ is compact. □

5.4 Preservation of Compactness by Continuous Mappings

One of the most important properties of compactness is that it is preserved by continuous mappings.

THEOREM 80 *Let X and Y be metric spaces. Suppose that X is compact. Let $f : X \longrightarrow Y$ be a continuous surjection. Then Y is also compact.*

Proof. We work directly from the definition. Let V_α be open sets of Y for every α in some index set I such that $\cup_{\alpha \in I} V_\alpha = Y$. Then, by Theorem 11 (page 24), $f^{-1}(V_\alpha)$ are open subsets of X. We have

$$X = \bigcup_{\alpha \in I} f^{-1}(V_\alpha).$$

We now apply the compactness of X to deduce the existence of a finite subset F of I such that

$$X = \bigcup_{\alpha \in F} f^{-1}(V_\alpha).$$

We wish to deduce that

$$Y = \bigcup_{\alpha \in F} V_\alpha. \tag{5.11}$$

Let $y \in Y$, then since f is surjective, there exists $x \in X$ such that $f(x) = y$. There exists $\alpha \in F$ such that $x \in f^{-1}(V_\alpha)$. It follows that $y = f(x) \in V_\alpha$. This verifies (5.11). ■

There is a formulation of this result in terms of compact subsets which is probably used more frequently.

COROLLARY 81 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$ be a continuous mapping. Let K be a compact subset of X. Then $f(K)$ is a compact subset of Y.*

Proof. By definition, K is a compact metric space in its own right. Since $f|_K$ can be regarded as a continuous mapping from K onto $f(K)$, it follows that $f(K)$ is compact, when viewed as a metric space with the metric obtained by restriction from Y. Hence by Theorem 80, $f(K)$ is a compact subset of Y. ■

Theorem 80 has important consequences because of the application to real-valued functions. We need the following result.

PROPOSITION 82 *The supremum of every non-empty compact subset K of $\mathbb{R}$ belongs to K.*

Of course, the same result applies to the infimum.

Proof. Let K be a compact non-empty subset of $\mathbb{R}$. Since K is bounded and non-empty, it possesses a supremum x. For every $n \in \mathbb{N}$, the number $x - \frac{1}{n}$ is not an upper bound for K. Thus there exists $x_n \in K$ and $x_n > x - \frac{1}{n}$. On the other hand, x is an upper bound for K so that $x_n \leq x$. It follows that $|x - x_n| < \frac{1}{n}$ so that (x_n) converges to x. Since $x_n \in K$ and K is closed, it follows that $x \in K$. ■

THEOREM 83 *Let X be a non-empty compact metric space and let $f : X \longrightarrow \mathbb{R}$ be continuous. Then f attains its maximum value.*

Proof. By Theorem 80, $f(X)$ is a non-empty compact subset of $\mathbb{R}$ which therefore contains its supremum. Hence, there exists $x_0 \in X$ such that

$$f(x_0) = \sup_{x \in X} f(x),$$

as required. ■

One of the most significant applications of this result involves norms on finite-dimensional spaces.

COROLLARY 84 *Let V be a finite dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$. Then any two norms on V are equivalent.*

Proof. We give the proof for a finite-dimensional real vector space. The complex case is similar. Let us select a basis $(e_1, e_2, \ldots, e_n)$ of V. We define a norm $\| \ \|_1$ on V by

$$\|\sum_{j=1}^{n} t_j e_j\|_1 = \sum_{j=1}^{n} |t_j|.$$

Then for any other norm $\| \ \|_V$ on V it will be shown that $\| \ \|_1$ and $\| \ \|_V$ are equivalent. We have

$$\|\sum_{j=1}^{n} t_j e_j\|_V \leq \sum_{j=1}^{n} |t_j| \|e_j\|_V \leq C \sum_{j=1}^{n} |t_j| = C \|\sum_{j=1}^{n} t_j e_j\|_1, \tag{5.12}$$

where

$$C = \max_{j=1}^{n} \|e_j\|_V.$$

For the converse inequality we will need to use the Heine–Borel Theorem which was proved with respect to the infinity norm

$$\|\sum_{j=1}^{n} t_j e_j\|_\infty = \max_{j=1}^{n} |t_j|.$$

This is not a problem because

$$\max_{j=1}^{n} |t_j| \le \sum_{j=1}^{n} |t_j| \le n \max_{j=1}^{n} |t_j|.$$

so that $\| \quad \|_1$ and $\| \quad \|_\infty$ are equivalent. It follows that the unit sphere S for the norm $\| \quad \|_1$ is compact for the metric of the $\| \quad \|_1$ norm. Explicitly we have

$$S = \{\sum_{j=1}^{n} t_j e_j ; \sum_{j=1}^{n} |t_j| = 1\}.$$

By (5.12), $v \longrightarrow \|v\|_V$ is continuous as a map

$$(V, \| \quad \|_1) \quad \longrightarrow \quad \mathbb{R}.$$

It follows that this function attains its minimum value on S. Thus, if we let

$$c = \inf_{v \in S} \|v\|_V, \tag{5.13}$$

there actually exists $u \in S$ such that $\|u\|_V = c$. Since u cannot be the zero vector, it follows that $c > 0$. Rescaling (5.12) now yields

$$c\|v\|_1 \le \|v\|_V,$$

for all $v \in V$. ∎

5.5 Compactness and Uniform Continuity

One of the most important applications of compactness is to uniform continuity. This is used heavily in all areas of approximation.

THEOREM 85 *Let X be a compact metric space and let Y be a metric space. Let $f : X \longrightarrow Y$ be a continuous mapping, then f is uniformly continuous.*

Proof. Let $\epsilon > 0$. We apply the continuity of f. At each point $x \in X$ there exist a number $\delta_x > 0$ such that

$$f(U(x, \delta_x)) \subseteq U(f(x), \tfrac{1}{2}\epsilon). \tag{5.14}$$

We can now write

$$X = \bigcup_{x \in X} U(x, \tfrac{1}{2}\delta_x).$$

Applying the compactness of X there is a finite subset $F \subseteq X$ such that

$$X = \bigcup_{x \in F} U(x, \tfrac{1}{2}\delta_x). \tag{5.15}$$

Now let $\delta = \min_{x \in F} \frac{1}{2}\delta_x$. We claim that this δ works in the definition of uniform continuity. Let z_1 and z_2 be points of X satisfying $d(z_1, z_2) < \delta$. By (5.15), there exists $x \in F$ such that $z_1 \in U(x, \frac{1}{2}\delta_x)$. Now, using the triangle inequality we have

$$d(x, z_2) \leq d(x, z_1) + d(z_1, z_2) < \tfrac{1}{2}\delta_x + \delta \leq \delta_x,$$

so that both z_1 and z_2 lie in $U(x, \delta_x)$. It now follows from (5.14) that $f(z_1)$ and $f(z_2)$ both lie in $U(f(x), \frac{1}{2}\epsilon)$. It then follows again by the triangle inequality that $d(f(z_1), f(z_2)) < \epsilon$ as required. ■

The following Theorem is a typical application of the use of uniform continuity in approximation theory.

THEOREM 86 (BERNSTEIN APPROXIMATION THEOREM) *Let $f : [0, 1] \longrightarrow \mathbb{R}$ be a continuous function. Define the nth **Bernstein polynomial** by*

$$B_n(f, x) = \sum_{k=0}^{n} {}^nC_k \, f(\tfrac{k}{n}) x^k (1 - x)^{n-k}.$$

Then $(B_n(f, \cdot))$ converges uniformly to f on $[0, 1]$.

Sketch proof. We leave the proof of the following three identities to the reader

$$1 = \sum_{k=0}^{n} {}^nC_k\, x^k(1-x)^{n-k}, \tag{5.16}$$

$$nx = \sum_{k=0}^{n} k\ {}^nC_k\, x^k(1-x)^{n-k}, \tag{5.17}$$

$$n(n-1)x^2 = \sum_{k=0}^{n} k(k-1)\ {}^nC_k\, x^k(1-x)^{n-k}. \tag{5.18}$$

Then it is easy to see that

$$\begin{aligned}
&\sum_{k=0}^{n} \left(x-\frac{k}{n}\right)^2\ {}^nC_k\, x^k(1-x)^{n-k} \\
&= \sum_{k=0}^{n} \left(x^2 - \frac{2k}{n}x + \frac{k^2-k}{n^2} + \frac{k}{n^2}\right)\ {}^nC_k\, x^k(1-x)^{n-k}, \\
&= x^2 - 2x^2 + \frac{n(n-1)}{n^2}x^2 + \frac{n}{n^2}x, \\
&= \frac{1}{n}x(1-x).
\end{aligned}$$

by applying (5.16), (5.17) and (5.18). Then for $\delta > 0$ we obtain a Tchebychev inequality

$$\begin{aligned}
\sum_{|x-\frac{k}{n}|>\delta} \delta^2\ {}^nC_k\, x^k(1-x)^{n-k} &\le \sum_{|x-\frac{k}{n}|>\delta} \left(x-\frac{k}{n}\right)^2\ {}^nC_k\, x^k(1-x)^{n-k} \\
&\le \frac{1}{n}x(1-x).
\end{aligned}$$

We are now ready to study the approximation. Since f is continuous on the compact set $[0,1]$ it is also uniformly continuous. Furthermore, by Corollary 81 (page 105), f is bounded. Thus we have

$$f(x) - B_n(f,x) = f(x) - \sum_{k=0}^{n} {}^nC_k\, f(\tfrac{k}{n})x^k(1-x)^{n-k},$$

$$= \sum_{k=0}^{n} {}^nC_k \, (f(x) - f(\tfrac{k}{n}))x^k(1-x)^{n-k},$$

and

$$\begin{aligned} |f(x) - B_n(f,x)| &\le \sum_{k=0}^{n} {}^nC_k \, |f(x) - f(\tfrac{k}{n})| x^k(1-x)^{n-k}, \\ &\le E_1 + E_2, \end{aligned}$$

where

$$\begin{aligned} E_1 &= \sum_{|x-\frac{k}{n}|>\delta} {}^nC_k \, |f(x) - f(\tfrac{k}{n})| x^k(1-x)^{n-k}, \\ &\le 2\|f\|_\infty \delta^{-2} \frac{1}{n} x(1-x), \\ &\le \frac{1}{2n} \|f\|_\infty \delta^{-2}, \end{aligned}$$

and

$$\begin{aligned} E_2 &= \sum_{|x-\frac{k}{n}|\le\delta} {}^nC_k \, |f(x) - f(\tfrac{k}{n})| x^k(1-x)^{n-k}, \\ &\le \sum_{k=0}^{n} {}^nC_k \, \omega_f(\delta) x^k(1-x)^{n-k}, \\ &= \omega_f(\delta). \end{aligned}$$

Let now $\epsilon > 0$. Then, using the uniform continuity of f, choose $\delta > 0$ so small that $\omega_f(\delta) < \frac{1}{2}\epsilon$. Then, with δ now fixed, select N so large that $\frac{1}{2N}\|f\|_\infty \delta^{-2} < \frac{1}{2}\epsilon$. It follows that

$$\sup_{0\le x\le 1} |f(x) - B_n(f,x)| \le \epsilon \qquad \forall n \ge N,$$

as required for uniform convergence of the Bernstein polynomials to f. ■

This proof does not address the question of motivation. Where does the Bernstein polynomial come from? To answer this question, we need to assume that the reader has a rudimentary knowledge of probability theory. Let X be a random variable taking values in $\{0, 1\}$. Assume that it takes the value 1 with probability

x and the value 0 with probability $1-x$. Now assume that we have n independent random variables $X_1, \ldots, X_n$ all with the same distribution as X. Let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Then it is an easy calculation to see that

$$P(S_n = \tfrac{k}{n}) = {}^nC_k\, x^k (1-x)^{n-k}$$

where $P(E)$ stands for the probability of the event E. It follows that

$$\mathbb{E}(f(S_n)) = B_n(f, x)$$

where $\mathbb{E}(Y)$ stands for the expectation of the random variable Y. By the law of averages, we should expect S_n to "converge to" x as n converges to ∞. Hence $B_n(f, x)$ should converge to $f(x)$ as n converges to ∞.

While the above argument is imprecise, it is possible to give a proof of the Bernstein Approximation Theorem using the Law of Large Numbers.

5.6 Compactness and Uniform Convergence

There are also some applications of compactness to establish uniform convergence.

PROPOSITION 87 (DINI'S THEOREM) *Let X be a compact metric space and suppose that (f_n) is a sequence of real-valued continuous functions on X decreasing to 0. That is*

$$f_n(x) \geq f_{n+1}(x) \qquad \forall n \in \mathbb{N},\ \forall x \in X$$

and for each fixed $x \in X$,

$$f_n(x) \quad \overset{n \longrightarrow \infty}{\longrightarrow} \quad 0.$$

Then (f_n) converges to 0 uniformly.

Proof. Obviously $f_n(x) \geq 0$ for all $n \in \mathbb{N}$ and $x \in X$. Let $\epsilon > 0$. Then for each $x \in X$ there exists N_x such that

$$n \geq N_x \qquad \Rightarrow \qquad f_n(x) < \tfrac{1}{2}\epsilon.$$

Now for all $x \in X$ there exists $\delta_x > 0$ such that

$$d(z,x) < \delta_x \qquad \Rightarrow \qquad |f_{N_x}(z) - f_{N_x}(x)| < \tfrac{1}{2}\epsilon.$$

We can now write

$$X = \bigcup_{x \in X} U(x, \delta_x).$$

Applying the compactness of X there is a finite subset $F \subseteq X$ such that

$$X = \bigcup_{x \in F} U(x, \delta_x).$$

Now let $N = \max_{x \in F} N_x$. We will show that $n \geq N$ implies that $f_n(z) < \epsilon$ simultaneously for all $z \in X$.

To verify this, let $z \in X$. Then we find $x \in F$ such that $d(z,x) < \delta_x$. It follows from this that $|f_{N_x}(z) - f_{N_x}(x)| < \frac{1}{2}\epsilon$. But, combining this with $f_{N_x}(x) < \frac{1}{2}\epsilon$ we obtain $f_{N_x}(z) < \epsilon$. Finally, since the sequence $(f_n(z))$ is decreasing we have $f_n(z) < \epsilon$ for all $n \geq N$. ∎

5.7 Equivalence of Compactness and Sequential Compactness

We begin with a Theorem whose proof parallels the proof of the Heine–Borel Theorem (page 103).

THEOREM 88 *Every separable sequentially compact space is compact.*

Proof. Let X be a separable sequentially compact metric space. We must show that X is compact. By Proposition 74 it is enough to show that X is countably compact. Thus, let V_k be open subsets of X for $k \in \mathbb{N}$. Suppose that $\cup_{k \in \mathbb{N}} V_k = X$. We will show that there exists a finite subset F of $\mathbb{N}$ such that

$$\bigcup_{k \in F} V_k = X. \tag{5.19}$$

Suppose not. Then, for every k, the set $F = \{1, 2, \ldots, k\}$ fails to satisfy (5.19). So, there is a point $x_k \in X$ with

$$x_k \notin \bigcup_{\ell=1}^{k} V_\ell. \tag{5.20}$$

Now by sequential compactness, the sequence (x_k) has a subsequence which converges to some element $x \in X$. By hypothesis, there exists $m \in \mathbb{N}$ such that $x \in V_m$. Since V_m is open, some tail of the subsequence lies entirely inside V_m. Therefore, there exists $k \in \mathbb{N}$ with $k \geq m$ and $x_k \in V_m$. This contradiction with (5.20) establishes the result. ∎

DEFINITION *A metric space X is said to be **totally bounded** iff for every $\epsilon > 0$ there exists a finite subset F of X such that*

$$\bigcup_{x \in F} U(x, \epsilon) = X. \tag{5.21}$$

THEOREM 89 *If X is a sequentially compact metric space, then X is totally bounded.*

Proof. Let $\epsilon > 0$ and suppose that (5.21) fails for every finite subset F of X. We will obtain a contradiction with the sequential compactness of X.

We define a sequence (x_n) inductively. Let x_1 be any point of X. We observe that X cannot be empty since then (5.21) holds with $F = \emptyset$. Now assume that $x_1, \ldots, x_n$ have been defined. We choose x_{n+1} such that

$$x_{n+1} \in X \setminus \bigcup_{k=1}^{n} U(x_k, \epsilon).$$

Once again, it is the failure of (5.21), this time with $F = \{x_1, \ldots, x_n\}$ which guarantees the existence of x_{n+1}.

Since X is sequentially compact, the sequence (x_n) possesses a subsequence convergent to some point x of X. Hence there exists N such that $x_N \in U(x, \epsilon)$. Thus $x \in U(x_N, \epsilon)$. Using the fact that $U(x_N, \epsilon)$ is open and hence a neighbourhood of x, and the convergence of the subsequence to x we see that there exists $n > N$ with $x_n \in U(x_N, \epsilon)$. But this contradicts the definition of x_n. ∎

An extension of this result will be needed later.

DEFINITION *A subset Y of a metric space X is said to be **totally bounded** iff for every $\epsilon > 0$ there exists a finite subset F of Y such that*

$$\bigcup_{y \in F} U(y, \epsilon) \supseteq Y.$$

It is almost immediate from the definition that if Y is a totally bounded subset, then so is $\mathrm{cl}(Y)$. In essence this is because of the inclusion chain

$$\bigcup_{y\in F} U(y,2\epsilon) \supseteq \bigcup_{y\in F} B(y,\epsilon) \supseteq \bigcup_{y\in F} U(y,\epsilon) \supseteq Y.$$

It follows from this that

$$\bigcup_{y\in F} U(y,2\epsilon) \supseteq \mathrm{cl}(Y).$$

The proof of the following result follows that of Theorem 89 so closely that we leave the details to the reader.

THEOREM 90 *Let Y be a subset of a metric space X. If $\mathrm{cl}(Y)$ is sequentially compact, then Y is a totally bounded subset of X*

Another very easy result is the following.

PROPOSITION 91 *Every totally bounded metric space is separable.*

Sketch proof. Choose a sequence (ϵ_n) of strictly positive reals decreasing to zero. For each $n \in \mathbb{N}$ apply the total boundedness condition to obtain a finite subset F_n of X such that

$$\bigcup_{x\in F_n} U(x,\epsilon_n) = X.$$

We leave the reader to show that

$$\bigcup_{n\in\mathbb{N}} F_n$$

is a countable dense subset of X. ■

We can now establish the converse to Proposition 75 (page 101).

COROLLARY 92 *Every sequentially compact metric space is compact.*

Proof. This is an immediate consequence of Theorem 88, Theorem 89 and Proposition 91.

5.8 Compactness and Completeness

PROPOSITION 93 *A sequentially compact metric space is complete.*

Proof. Let (x_n) be a Cauchy sequence in a sequentially compact metric space X. Then there is a subsequence (x_{n_k}) converging to some element $x \in X$. We use the convergence of this subsequence and the Cauchy condition to establish the convergence of the original sequence to x.

Let $\epsilon > 0$. Then, by the Cauchy condition, there exists N such that

$$p, q > N \qquad \Rightarrow \qquad d(x_p, x_q) < \tfrac{1}{2}\epsilon.$$

By convergence of the subsequence we also have

$$k > K \qquad \Rightarrow \qquad d(x_{n_k}, x) < \tfrac{1}{2}\epsilon.$$

Let us choose $k = \max(N, K) + 1$. Then taking $q = n_k$ and using the triangle inequality, we obtain

$$p > N \qquad \Rightarrow \qquad d(x_p, x) < \epsilon.$$

as required to establish convergence. ■

THEOREM 94 *A complete totally bounded metric space is sequentially compact.*

Proof. This proof uses the famous diagonal subsequence argument. Let X be a complete totally bounded metric space. Let (x_k) be a sequence in X. Let (ϵ_k) be a sequence of strictly positive reals decreasing to zero. For each $k \in \mathbb{N}$ we apply the total boundedness condition to obtain a finite subset F_k of X such that

$$\bigcup_{x \in F_k} U(x, \epsilon_k) = X.$$

We extract subsequences inductively.

$$\begin{array}{ccccc}
x_{n_{1,1}} & x_{n_{1,2}} & x_{n_{1,3}} & x_{n_{1,4}} & \cdots \\
x_{n_{2,1}} & x_{n_{2,2}} & x_{n_{2,3}} & x_{n_{2,4}} & \cdots \\
x_{n_{3,1}} & x_{n_{3,2}} & x_{n_{3,3}} & x_{n_{3,4}} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
x_{n_{k,1}} & x_{n_{k,2}} & x_{n_{k,3}} & x_{n_{k,4}} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots
\end{array}$$

The first subsequence $(x_{n_{1,\ell}})$ is a subsequence of (x_n) contained in a set $U(s_1, \epsilon_1)$ for some $s_1 \in F_1$. This uses the fact that the sequence (x_n) cannot meet all such $U(s_1, \epsilon_1)$ in a finite set. The same reasoning allows us to extract the second subsequence $(x_{n_{2,\ell}})$ from $(x_{n_{1,\ell}})$ so that $(x_{n_{2,\ell}})$ is contained in a set $U(s_2, \epsilon_2)$ for some $s_2 \in F_2$. We continue in this way.

We now consider the ***diagonal subsequence*** (x_{m_k}) defined by

$$x_{m_k} = x_{n_{k,k}}.$$

The crucial observation is that, for each k, the tail sequence $x_{m_k}, x_{m_{k+1}}, x_{m_{k+2}}, \ldots$ is itself a subsequence of $x_{n_{k,1}}, x_{n_{k,2}}, x_{n_{k,3}}, x_{n_{k,4}}, \ldots$. Thus, for each $k \in \mathbb{N}$ the tail sequence $x_{m_k}, x_{m_{k+1}}, x_{m_{k+2}}, \ldots$ lies in $U(s_k, \epsilon_k)$ and hence has diameter less than $2\epsilon_k$. It follows immediately that (x_{m_k}) is a Cauchy sequence and hence convergent in X. ■

5.9 Equicontinuous Sets

Throughout this section, K denotes a compact metric space. We denote by $C(K)$ the space of bounded real-valued continuous functions on K. All the proofs presented here also work for complex valued functions. We consider $C(K)$ as a normed space with the uniform norm. We have already observed that $C(K)$ is complete with this norm — see the example following Proposition 50 (page 77).

DEFINITION *Let $F \subseteq C(K)$. We say that F is* ***equicontinuous*** *iff for all $\epsilon > 0$, there exists $\delta > 0$ such that*

$$d_K(x, y) < \delta \qquad \Rightarrow \qquad |f(x) - f(y)| < \epsilon \qquad \forall f \in F.$$

It is more or less clear that a set F is equicontinuous iff there is a "modulus of continuity function" that works simultaneously for all functions in the set F. Explicitly we have the following Lemma, the proof of which is left as an exercise to the reader.

LEMMA 95 *Let $F \subseteq C(K)$. Then F is equicontinuous iff there is a function $\omega : [0, \infty[\longrightarrow [0, \infty[$ satisfying $\omega(0) = 0$ and continuous at 0, such that*

$$|f(x) - f(y)| \leq \omega(d_K(x, y)) \qquad \forall x, y \in K, \ \forall f \in F.$$

The key result of this section is the following.

THEOREM 96 (ASCOLI–ARZELA THEOREM) *Let $F \subseteq C(K)$. Then the following are equivalent statements.*

- *F has compact closure in $C(K)$.*
- *F is bounded in $C(K)$ and F is equicontinuous.*

Proof. We assume first that F has compact closure in $C(K)$. Then according to Proposition 68 (page 96), F is bounded in $C(K)$. We show that F is equicontinuous. Suppose not. Then there exists $\epsilon > 0$, two sequences (x_n) and (y_n) in K and a sequence (f_n) in F such that $(d(x_n, y_n))$ converges to 0 and $|f_n(x_n) - f_n(y_n)| \geq \epsilon$. Now using the hypothesis that F has compact closure, we see that (f_n) has a subsequence convergent in $C(K)$. We denote the limit function by f. Then there exists $n \in \mathbb{N}$ such that

$$\|f - f_n\|_\infty \leq \tfrac{1}{3}\epsilon.$$

It now follows that $|f(x_n) - f(y_n)| \geq \frac{1}{3}\epsilon$, contradicting the uniform continuity of f. The function f is uniformly continuous by virtue of Theorem 85 (page 108).

The real work of the proof is contained in the converse. Let us assume that F is bounded and equicontinuous. It suffices to show that F is totally bounded. For then we will have that $\mathrm{cl}(F)$ is totally bounded (by a remark on page 114) and complete, since $C(K)$ is complete. It is then enough to apply Theorem 94 and Corollary 92 to deduce that $\mathrm{cl}(F)$ is compact in $C(K)$.

Let $\epsilon > 0$. Then using the equicontinuity, we can find $\delta > 0$ such that

$$d(x, y) < \delta, f \in F \qquad \Rightarrow \qquad |f(x) - f(y)| < \tfrac{1}{3}\epsilon. \tag{5.22}$$

Now using the total boundedness of K, we can find $N \in \mathbb{N}$ and $x_1, x_2, \ldots x_N \in K$ such that

$$\bigcup_{n=1}^{N} U(x_n, \delta) = K. \tag{5.23}$$

Since F is bounded, the set

$$\{(f(x_1), f(x_2), \ldots, f(x_N)); f \in F\}$$

is a bounded subset of $\mathbb{R}^N$ and hence is totally bounded in $\mathbb{R}^N$ by the Heine–Borel Theorem and Theorem 90. Hence there exists $M \in \mathbb{N}$ and functions

$f_1, f_2, \ldots, f_M$ in F such that for all $f \in F$ there exists m with $1 \leq m \leq M$ such that

$$\sup_{n=1}^{N} |f(x_n) - f_m(x_n)| \leq \tfrac{1}{3}\epsilon. \tag{5.24}$$

Combining now (5.22), (5.23) and (5.24) we have, for an arbitrary point $x \in K$ and n chosen such that $d(x, x_n) < \delta$,

$$|f(x) - f_m(x)| \leq |f(x) - f(x_n)| + |f(x_n) - f_m(x_n)| + |f_m(x) - f_m(x_n)| < \epsilon.$$

It follows that F is totally bounded in $C(K)$. ■

5.10 The Stone–Weierstrass Theorem

One of the key approximation theorems in analysis is the Stone–Weierstrass Theorem. Let K be a compact metric space and recall that the notation $C(K)$ stands for the space of real-valued continuous functions on K. The space $C(K)$ is a vector space under pointwise operations, a fact that we have already heavily used. It is also a linear associative algebra — this means that $C(K)$ is closed under pointwise multiplication and that it satisfies the axioms for a ring.

DEFINITION *A subset $A \subseteq C(K)$ is said to be a* ***subalgebra*** *of $C(K)$ if it is closed under both the linear and multiplicative operations. Explicitly, this means*

- $f_1, f_2 \in A, t_1, t_2 \in \mathbb{R} \Longrightarrow t_1 f_1 + t_2 f_2 \in A.$
- $f_1, f_2 \in A \Longrightarrow f_1 \cdot f_2 \in A.$

where the function combinations are defined by

$$(t_1 f_1 + t_2 f_2)(x) = t_1 f_1(x) + t_2 f_2(x) \qquad \forall x \in K$$

and

$$(f_1 \cdot f_2)(x) = f_1(x) f_2(x) \qquad \forall x \in K.$$

DEFINITION *A subalgebra A of $C(K)$ is said to be* ***unital*** *iff the constant function* $\mathbf{1}$ *which is identically equal to* 1 *belongs to A. A subalgebra A is said to* ***separating*** *iff whenever x_1 and x_2 are two distinct points of K, there exists a function $f \in A$ such that $f(x_1) \neq f(x_2)$.*

If A is both unital and separating, then whenever x_1 and x_2 are two distinct points of K and a_1 and a_2 are given real numbers, there exists $f \in A$ such that $f(x_1) = a_1$ and $f(x_2) = a_2$.

THEOREM 97 (STONE–WEIERSTRASS THEOREM) *Let K be a compact metric space and suppose that A be a unital separating subalgebra of $C(K)$. Then A is dense in $C(K)$ for the standard uniform metric on $C(K)$.*

Before we can prove this result, we need to develop some preliminary ideas.

LEMMA 98 *If A is a unital separating subalgebra of $C(K)$, then so is its uniform closure* $\mathrm{cl}(A)$.

Proof. Obviously $\mathrm{cl}(A)$ is unital and separating because $A \subseteq \mathrm{cl}(A)$. It remains to check that $\mathrm{cl}(A)$ is a subalgebra. This is routine. For instance, to show that $f \cdot g \in \mathrm{cl}(A)$ whenever $f, g \in \mathrm{cl}(A)$, we find a sequence (f_n) in A converging uniformly to f and a sequence (g_n) converging uniformly to g. Clearly

$$\begin{aligned} \|f \cdot g - f_n \cdot g_n\| &= \|(f - f_n) \cdot g + f_n \cdot (g - g_n)\| \\ &\leq \|(f - f_n) \cdot g\| + \|f_n \cdot (g - g_n)\| \\ &= \|f - f_n\|\|g\| + \|f_n\|\|g - g_n\| \\ &\longrightarrow 0 \end{aligned}$$

so that $(f_n \cdot g_n)$ converges uniformly to $f \cdot g$. The proof that $\mathrm{cl}(A)$ is closed under linear operations is similar. ∎

The upthrust of Lemma 98 is that the Stone–Weierstrass Theorem can be reformulated in the following way.

THEOREM 99 *Let K be a compact metric space and suppose that A be a uniformly closed unital separating subalgebra of $C(K)$. Then $A = C(K)$.*

LEMMA 100 *Let $a > 0$. Then there exists a sequence (p_n) of real polynomials such that $p_n(x) \longrightarrow |x|$ uniformly on $[-a, a]$.*

This is a consequence of Theorem 86 after some elementary rescaling, but it is also possible to give a proof *à la main*.

Proof. There are various pitfalls in designing a strategy for the proof. For instance, taking the p_n to be the partial sums of a fixed power series is doomed to failure. The proof has to be fairly subtle.

Without loss of generality one may take $a = 1$. Let us define the function

$$f_n(x) = \sqrt{\frac{2}{\pi}} n e^{-\frac{1}{2}n^2x^2}.$$

The key facts about this function are that it is positive, that

$$\int_{-\infty}^{\infty} f_n(x)dx = 2$$

and that for large values of n, the graph of f_n has a "spike" near 0. Let ϵ_n be a sequence of strictly positive numbers converging to 0. Let r_n be an even polynomial such that

$$\sup_{-1 \le x \le 1} |f_n(x) - r_n(x)| \le \epsilon_n.$$

We can easily construct r_n by truncating the power series expansion of f_n. If ϵ_n is suitably small, r_n will have this same spiky behaviour. Following this philosophy, we can expect the polynomial q_n given by

$$q_n(s) = \int_0^s r_n(t)dt \tag{5.25}$$

to approximate the "signum" function, and the second primitive p_n

$$p_n(x) = \int_0^x q_n(s)ds \tag{5.26}$$

should approximate the "modulus" function. One can actually obtain p_n directly from r_n by

$$p_n(x) = \int_0^x (x - t)r_n(t)dt.$$

Clearly, from (5.25) and (5.26) p_n is an even polynomial function. Since both $|x|$ and $p_n(x)$ are even in x we need only estimate

$$\sup_{0\le x\le 1} |x - p_n(x)|. \tag{5.27}$$

Clearly

$$\left| \int_0^x (x-t) r_n(t)dt - \int_0^x (x-t) f_n(t)dt \right| \le \epsilon_n$$

for all $x \in [0,1]$ so that it is enough to show that

$$\sup_{0\le x\le 1} \left| x - \int_0^x (x-t) f_n(t)dt \right|$$

tends to zero as n tends to infinity. We have

$$\begin{aligned} \sup_{0\le x\le 1} \left| x - \int_0^x (x-t) f_n(t)dt \right| &= \sup_{0\le x\le 1} \left| \int_0^\infty x f_n(t)dt - \int_0^x (x-t) f_n(t)dt \right| \\ &\le \left\{ \sup_{0\le x\le 1} x \int_x^\infty f_n(t)dt \right\} + \int_0^\infty t f_n(t)dt \end{aligned} \tag{5.28}$$

The second term in (5.28) is independent of x and tends to zero like $\frac{1}{n}$ so we concentrate on the first term which, after making a change of variables in the integral, can be rewritten

$$\sqrt{\frac{2}{\pi}} \sup_{0\le x\le 1} x \int_{nx}^\infty e^{-\frac{1}{2}t^2} dt \tag{5.29}$$

But (5.29) also tends to 0 as n tends to ∞ since

$$x \int_{nx}^\infty e^{-\frac{1}{2}t^2} dt \le \begin{cases} n^{-\frac{1}{2}} \int_0^\infty e^{-\frac{1}{2}t^2} dt & \text{if } 0 \le x \le n^{-\frac{1}{2}}, \\ \int_{\sqrt{n}}^\infty e^{-\frac{1}{2}t^2} dt & \text{if } n^{-\frac{1}{2}} \le x \le 1. \end{cases}$$

These estimates show that (5.27) tends to 0 as n tends to ∞ as required. ■

LEMMA 101 *Let A be a uniformly closed unital subalgebra of $C(K)$. Let $f, g \in A$. Then the functions* $\max(f,g)$ *and* $\min(f,g)$ *are also in A.*

Proof. Since we have the identities

$$\max(f,g) = \tfrac{1}{2}(f+g+|f-g|)$$

and

$$\min(f,g) = \tfrac{1}{2}(f+g-|f-g|)$$

it is enough to establish that if $h \in A$ then $|h| \in A$. For then, taking $h = f - g$ the result follows. Since h is a continuous function defined on a compact space, it is bounded and hence it takes values in $[-a, a]$ for some $a > 0$. It now follows from Lemma 100 that the sequence of functions $(p_n \circ h)$ converges uniformly to $|h|$. Each function $p_n \circ h$ is in A since A is a unital subalgebra of $C(K)$. Hence, since A is also uniformly closed it follows that $|h| \in A$. ■

Proof of the Stone–Weierstrass Theorem. We start with a function $f \in C(K)$ that we wish to approximate and a positive number ϵ which is the allowed uniform error. Let x be an arbitrary point of K which we fix for the moment. Now let y be another arbitrary point of K which we allow to vary. Since A is both unital and separating, we can find a function $h_{x,y} \in A$ such that

$$h_{x,y}(x) = f(x)$$

and

$$h_{x,y}(y) = f(y).$$

Since both f and $h_{x,y}$ are continuous at y there is an open neighbourhood $V_{x,y}$ of y such that $h_{x,y}(z) - f(z) < \epsilon$ for all $z \in V_{x,y}$. Clearly we have for each fixed x

$$K = \bigcup_{y \in K} V_{x,y},$$

and hence by the compactness of K, there exist $m \in \mathbb{N}$ and $y_1, \ldots, y_m \in K$ such that

$$K = \bigcup_{k=1}^{m} V_{x,y_k}.$$

It is worth pointing out that m and the points $y_1, \ldots, y_m$ depend on x, but it would be too cumbersome to express this fact notationally.

The function

$$g_x = \min_{k=1}^{m} h_{x,y_k}$$

is in $\mathrm{cl}(A)$ because of Lemma 101 and has the following properties

$$g_x(x) = f(x)$$

and

$$g_x(z) < f(z) + \epsilon \qquad \forall z \in K. \tag{5.30}$$

We note that (5.30) holds since, for all $z \in K$ there exists k with $1 \leq k \leq m$ such that $z \in V_{x,y_k}$. We then have

$$g_x(z) \leq h_{x,y_k}(z) < f(z) + \epsilon.$$

Dependence on y has now been eliminated, and we now allow x to vary. Since f and g_x are continuous at x, there is an open neighbourhood U_x of x such that

$$g_x(z) > f(z) - \epsilon \qquad \forall z \in U_x \tag{5.31}$$

Clearly we have

$$K = \bigcup_{x \in K} U_x,$$

and hence by the compactness of K, there exist $\ell \in \mathbb{N}$ and $x_1, \ldots, x_\ell \in K$ such that

$$K = \bigcup_{j=1}^{\ell} U_{x_j}.$$

The function

$$g = \max_{j=1}^{\ell} g_{x_j}$$

is in $\mathrm{cl}(A)$ applying Lemma 101 again. We check that both the inequalities

$$f(z) - \epsilon < g(z) < f(z) + \epsilon \qquad \forall z \in K$$

hold. The inequality on the right holds because of (5.30). For the inequality on the left, let $z \in K$. Then there exists j with $1 \leq j \leq \ell$ such that $z \in U_{x_j}$. We have, using (5.31)

$$g(z) \geq g_{x_j}(z) > f(z) - \epsilon.$$

We have shown that $\mathrm{cl}(A)$ is dense in $C(K)$ and hence we conclude that $\mathrm{cl}(A) = C(K)$ as required. ■

EXAMPLE Let $K = [-1, 1]$ and let A be the algebra of (restrictions of) polynomial functions. Then A is clearly a unital separating subalgebra of $C(K)$ and is therefore (uniformly) dense in $C(K)$. Of course this example contains Lemma 100 as a special case. □

EXAMPLE Let $K = [0, 1] \times [0, 1]$ the unit square. If f and g are continuous functions on $[0, 1]$, we can make a new function on K by

$$(f \otimes g)(s, t) = f(s) \cdot g(t). \tag{5.32}$$

It is easy to see that the set A of all finite sums of such functions is a unital separating subalgebra of $C(K)$. In fact, as linear spaces, we have

$$A \cong C([0, 1]) \otimes C([0, 1])$$

the tensor product of $C([0, 1])$ with itself. This is the reason for using the $\otimes$ notation in (5.32). The Stone–Weierstrass Theorem shows that A is dense in $C(K)$. There are many quite difficult problems associated with this example, for instance it is true, but not immediately obvious that A is a proper subalgebra of $C(K)$. □

The Stone–Weierstrass Theorem allows a number of extensions. First of all, there is an extension to complex-valued continuous functions.

THEOREM 102 *Let K be a compact metric space and suppose that A be a unital separating self-adjoint subalgebra of $C(K, \mathbb{C})$. Then A is dense in $C(K, \mathbb{C})$ for the uniform metric.*

Here, the condition that A is ***self-adjoint*** means that $\overline{f} \in A$ whenever $f \in A$. The function $\overline{f}$ is defined by

$$\overline{f}(z) = \overline{f(z)} \qquad \forall z \in K.$$

Proof. The key observation is that if $f \in A$ then $\Re f = \frac{1}{2}(f + \overline{f})$ is also in A. It is now easy to see that

$$\Re A = \{\Re f; f \in A\} = A \cap C(K, \mathbb{R})$$

is a unital separating subalgebra of $C(K, \mathbb{R})$. Thus applying the standard Stone–Weierstrass Theorem we see that $\Re A$ is dense in $C(K, \mathbb{R})$. Since $\Re A \subseteq A$ the result follows immediately. ■

EXAMPLE Perhaps the most interesting application is to trigonometric polynomials. Let $\mathbb{T}$ denote the quotient group $\mathbb{R}/2\pi\mathbb{Z}$. Here we are viewing $2\pi\mathbb{Z}$ as an additive subgroup of $\mathbb{R}$ considered as an abelian group. We can think of $\mathbb{T}$ as the reals modulo 2π. Topologically, $\mathbb{T}$ is a circle, so it is called the circle group. In particular, $\mathbb{T}$ is compact, if for instance it is given the metric

$$d_{\mathbb{T}}(\dot{t}, \dot{s}) = \inf_{n\in\mathbb{Z}} |2n\pi + t - s|$$

for $t, s \in \mathbb{R}$ and $\dot{t}, \dot{s}$ denoting the corresponding points of $\mathbb{T}$. A ***trigonometric polynomial*** is a complex-valued function p on $\mathbb{R}$ given by a finite sum

$$p(t) = \sum_{n=-N}^{N} a_n e^{int} \qquad t \in \mathbb{R}$$

where $a_n \in \mathbb{C}$. The function p is 2π-periodic when viewed as a function on $\mathbb{R}$ and hence it may be considered as a function on $\mathbb{T}$. It is straightforward to see that the set A of all trigonometric polynomials is a unital self-adjoint separating subalgebra of $C(\mathbb{T}, \mathbb{C})$. It follows that A is dense in $C(\mathbb{T}, \mathbb{C})$. This result is very important in the theory of Fourier series, although it is normally approached from a rather different angle. □

The second major extension of the Stone–Weierstrass Theorem involves ***one-point compactifications***. A complete treatment is outside the scope of these notes. It is however possible to give the general idea.

Given a space such as $\mathbb{R}$ it is possible to add a ***point at infinity*** designated ∞, and define a new metric $\tilde{d}$ on the resulting space. Consider the map $f : \mathbb{R} \longrightarrow \mathbb{R}^2$ defined by

$$f(u) = (\frac{1-u^2}{1+u^2}, \frac{2u}{1+u^2})$$

which actually maps into the unit circle S^1 in $\mathbb{R}^2$. The only point of the unit circle which is not in the image of f is the point $(-1, 0)$. We treat this point as if it were $f(\infty)$. We define

$$\tilde{d}(u, v) = \|f(u) - f(v)\|$$

where $\|\ \ \|$ is the standard Euclidean norm on $\mathbb{R}^2$ and

$$\tilde{d}(\infty, v) = \tilde{d}(v, \infty) = \|(-1, 0) - f(v)\|.$$

Together with the required $\tilde{d}(\infty, \infty) = 0$, this clearly defines a metric on $\mathbb{R}\cup\{\infty\}$, because it is really just the Euclidean distance on the unit circle. If we restrict the metric $\tilde{d}$ to $\mathbb{R}$ we obtain a metric which is topologically equivalent to, but not uniformly equivalent to the standard metric on $\mathbb{R}$. Clearly, $(\mathbb{R} \cup \{\infty\}, \tilde{d})$ is a compact metric space because the unit circle is a compact subset of $\mathbb{R}^2$. This space is called the one-point compactification of $\mathbb{R}$. Similar constructions lead to one-point compactifications of many other spaces. For instance, the one-point compactification of $\mathbb{R}^n$ can be identified to the n-sphere S^n in $\mathbb{R}^{n+1}$. There is also a very natural two-point compactification of $\mathbb{R}$ which can be denoted $[-\infty, \infty]$.

DEFINITION *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$. We say that f* ***possesses a limit*** *a* ***at infinity*** *if and only if for all $\epsilon > 0$, there exists $A > 0$ such that $|f(x) - a| < \epsilon$ whenever $|x| > A$. We also say that f* ***vanishes at infinity*** *if and only if f possesses the limit 0 at infinity.*

The set of all continuous real-valued functions on $\mathbb{R}$ vanishing at infinity will be denoted $C_0(\mathbb{R})$. It is easy to see that $C_0(\mathbb{R})$ is a uniformly closed subalgebra of $C(\mathbb{R})$. The following Proposition is left as an exercise.

PROPOSITION 103 *A continuous function $f : \mathbb{R} \longrightarrow \mathbb{R}$ extends to a continuous function $\tilde{f} : \mathbb{R} \cup \{\infty\} \longrightarrow \mathbb{R}$ if and only if f possesses a limit at infinity.*

THEOREM 104 *Let A be a separating subalgebra of $C_0(\mathbb{R})$ that separates from infinity. Explicitly, this last statement means that for all $x \in \mathbb{R}$ there exists $f \in A$ such that $f(x) \neq 0$. Then A is uniformly dense in $C_0(\mathbb{R})$.*

Proof. Let B denote the set $\{f + \lambda\mathbf{1}; f \in A, \lambda \in \mathbb{R}\}$, an algebra of continuous functions on $\mathbb{R}$ which possess a limit at infinity. The set of lifts $\tilde{B} = \{\tilde{g}; g \in B\}$ is then a unital subalgebra of $C(\mathbb{R} \cup \{\infty\})$. The hypotheses on A guarantee that $\tilde{B}$ separates the points of $\mathbb{R} \cup \{\infty\}$. Thus by the Stone–Weierstrass Theorem, $\tilde{B}$ is uniformly dense in $C(\mathbb{R} \cup \{\infty\})$. It follows easily that A is uniformly dense in $C_0(\mathbb{R})$. ■

EXAMPLE This example is on $[0, \infty[$ rather than $\mathbb{R}$. Let A consist of all functions $f : [0, \infty[\longrightarrow \mathbb{R}$ of the type

$$f(x) = \sum_{j=1}^{n} a_j e^{-\lambda_j x} \qquad (x \geq 0)$$

where $a_j \in \mathbb{R}$ and $\lambda_j > 0$. Such a function is clearly continuous and vanishes at infinity. The set A is clearly a separating algebra of functions which also separates from infinity. Hence A is uniformly dense in $C_0([0,\infty[)$. This result can be used to yield the Uniqueness Theorem for Laplace transforms. □

6

Connectedness

From the intuitive point of view, a metric space is connected if it is in one piece.

DEFINITION *A **splitting** of a metric space X is a partition*

$$X = X_1 \cup X_2, \qquad (6.1)$$
$$\emptyset = X_1 \cap X_2,$$

*where X_1 and X_2 are open subsets of X. The splitting (6.1) is said to be **trivial** iff either $X_1 = X$ and $X_2 = \emptyset$ or $X_1 = \emptyset$ and $X_2 = X$. A metric space X is **connected** iff every splitting of X is trivial.*

In a splitting, the subsets X_1 and X_2, being complements of each other, are also closed.

PROPOSITION 105 *Every closed interval $[a, b]$ of $\mathbb{R}$ is connected.*

Proof. If $b < a$ then $[a, b] = \emptyset$. If $a = b$ then $[a, b]$ is a singleton. In either case, all partitions of $[a, b]$, whether they are splittings or not, are trivial.

Let us suppose that $a < b$ and that

$$[a, b] = X_1 \cup X_2, \qquad (6.2)$$
$$\emptyset = X_1 \cap X_2,$$

is a splitting of $[a, b]$. We may assume without loss of generality that $a \in X_1$, for if not, it suffices to interchange the sets X_1 and X_2. Let us suppose that $X_2 \neq \emptyset$. Then we define

$$c = \inf X_2.$$

We claim that $a < c$. Since X_1 is open, there exists $\epsilon > 0$ such that $U(a, \epsilon) \subseteq X_1$. This means that $[a, a + \epsilon[\subseteq X_1$ or equivalently that $X_2 \subseteq [a + \epsilon, b]$. It follows that $c \geq a + \epsilon$.

Exactly the same argument shows that if $c \in X_1$ and $c < b$ then there exists $\epsilon > 0$ such that $[c, c + \epsilon[\subseteq X_1$. But since by definition of c, we have $[a, c[\subseteq X_1$, we conclude that $[a, c + \epsilon[\subseteq X_1$ or equivalently $X_2 \subseteq [c + \epsilon, b]$ contradicting the definition of c. On the other hand, if $c = b$ and $c \in X_1$, then X_2 must be empty.

Hence, it must be the case that $c \in X_2$. But then there exists $\epsilon > 0$ such that $]c - \epsilon, c + \epsilon[\subseteq X_2$ which also contradicts the definition of c.

We are therefore forced to concluded that the supposition $X_2 \neq \emptyset$ is false. It follows that the splitting (6.2) is trivial. ■

6.1 Connected Subsets

So far we have discussed connectedness for metric spaces. We now extend the concept to subsets in the usual way.

DEFINITION *Let A be a subset of a metric space X. Then A is a* ***connected subset*** *iff A is a connected metric space in the restriction metric inherited from X.*

When we disentangle this definition using Theorem 30 (page 41) we obtain the following complicated Proposition.

PROPOSITION 106 *Let X be a metric space and let $A \subseteq X$. Then the following two conditions are equivalent.*

- *A is a connected subset of X.*
- *Whenever V_1 and V_2 are open subsets of X such that*

$$A \subseteq V_1 \cup V_2 \tag{6.3}$$

  and

$$\emptyset = A \cap V_1 \cap V_2, \tag{6.4}$$

  then either $A \subseteq V_1$ or $A \subseteq V_2$.

We leave the proof of this Proposition to the reader.

LEMMA 107 *Let X be a metric space and suppose that A is a connected subset of X. Then for every splitting*

$$X = V_1 \cup V_2, \tag{6.5}$$

$$\emptyset = V_1 \cap V_2, \tag{6.6}$$

of X we must have either $A \subseteq V_1$ or $A \subseteq V_2$.

Proof. It is immediate that (6.3) follows from (6.5) and (6.4) follows from (6.6). The conclusion follows immediately from the connectivity of A. ■

Another result we can obtain from Proposition 106 is the following.

PROPOSITION 108 *Let A be a connected subset of a metric space X. Then $\mathrm{cl}(A)$ is also connected.*

Proof. We suppose that V_1 and V_2 are open subsets of X such that

$$\mathrm{cl}(A) \qquad \subseteq V_1 \cup V_2 \tag{6.7}$$

and

$$\emptyset = \mathrm{cl}(A) \cap V_1 \cap V_2$$

hold. Then *a fortiori* (6.3) and (6.4) also hold. Since A is connected, we deduce that either $A \subseteq V_1$ or $A \subseteq V_2$. Let us suppose that $A \subseteq V_1$ without loss of generality. We claim that

$$\mathrm{cl}(A) \cap V_2 = \emptyset. \tag{6.8}$$

We establish the claim by contradiction. If $x \in \mathrm{cl}(A) \cap V_2$, then we can find a sequence (x_j) in A converging to x. Since V_2 is open and $x \in V_2$, for j large enough, we will have

$$x_j \in A \cap V_2 \subseteq A \cap V_1 \cap V_2 = \emptyset,$$

a contradiction. The claim is established. But now by (6.7) and (6.8) we find that $\mathrm{cl}(A) \subseteq V_1$. Similarly, supposing that $A \subseteq V_2$ will lead to $\mathrm{cl}(A) \subseteq V_2$. ■

6.2 Connectivity of the Real Line

PROPOSITION 109 *Every interval in $\mathbb{R}$ is connected.*

Proof. Let I be an interval of $\mathbb{R}$. We view I as a metric space in its own right and show that it is connected. Suppose not. Then there is a non-trivial splitting of I. Let $a, b \in I$ be points of I on different sides of the splitting. Without loss of generality we may suppose that $a \leq b$. But, by Proposition 105 (page 128) the closed interval $[a, b]$ is a connected subset of I containing both a and b. Hence, by Lemma 107 a and b must lie on the same side of any splitting – a contradiction.■

The converse is also true.

THEOREM 110 *Every non-empty connected subset of $\mathbb{R}$ is an interval.*

Proof. Let I be a connected subset of $\mathbb{R}$. Let $a = \inf I$ and $b = \sup I$ with the understanding that $a = -\infty$ if I is unbounded below and $b = \infty$ if I is unbounded above. It is the order completeness axiom that guarantees the existence of a and b. We claim that $]a, b[\subseteq I$. For if not, there exists $c \notin I$ satisfying $a < c < b$. But then taking $X = \mathbb{R}$, $A = I$, $V_1 =]-\infty, c[$ and $V_2 =]c, \infty[$ in Proposition 106 (page 129) shows that I is not connected. If $a \in \mathbb{R}$ then a may or may not be in I. If $b \in \mathbb{R}$ then b may or may not be in I. But in any event, I is an interval. ■

6.3 Connected Components

THEOREM 111 *Let X be a metric space. Let A and B be connected subsets of X with $A \cap B \neq \emptyset$. Then $A \cup B$ is connected.*

Proof. Let V_1 and V_2 are open subsets of X such that

$$A \cup B \subseteq V_1 \cup V_2$$

and

$$\emptyset = (A \cup B) \cap V_1 \cap V_2, \tag{6.9}$$

then we must show that either $A \cup B \subseteq V_1$ or $A \cup B \subseteq V_2$. It is clear from Lemma 107 and the fact that A is connected that either $A \subseteq V_1$ or $A \subseteq V_2$. Similarly, since B is connected, either $B \subseteq V_1$ or $B \subseteq V_2$. There are then 4 possibilities.

- $A \subseteq V_1$ and $B \subseteq V_1$.
- $A \subseteq V_1$ and $B \subseteq V_2$.
- $A \subseteq V_2$ and $B \subseteq V_1$.
- $A \subseteq V_2$ and $B \subseteq V_2$.

We show that the second and third cases are impossible. Suppose for instance that the second case holds. Let $x \in A \cap B$. Then $x \in V_1$ and $x \in V_2$. From this it follows that $x \in (A \cup B) \cap V_1 \cap V_2$ which contradicts (6.9). The third case is impossible by similar reasoning. It follows that either the first case holds, so that $A \cup B \subseteq V_1$, or the fourth case holds and $A \cup B \subseteq V_2$. ■

The next step is to discuss whether two points can be separated one from the other in a metric space. This turns out to be a key notion.

DEFINITION *Two elements x_1 and x_2 of a metric space X are* ***connected through*** *X iff there is a connected subset C of X such that $x_1, x_2 \in C$. In this circumstance we will write $x_1 \underset{X}{\sim} x_2$.*

THEOREM 112 *The relation $\underset{X}{\sim}$ is an equivalence relation on X.*

Proof. The symmetry condition is obvious because the definition of the relation is symmetric in x_1 and x_2. For the reflexivity, it suffices to take $C = \{x\}$ if $x_1 = x_2 = x$. All the work is in establishing the transitivity. Let $x_1, x_2, x_3 \in X$ and suppose that $x_1 \underset{X}{\sim} x_2$ and $x_2 \underset{X}{\sim} x_3$. Then by definition, there exist connected subsets A and B of X such that $x_1, x_2 \in A$ and $x_2, x_3 \in B$. Clearly, $x_2 \in A \cap B$. An application of Theorem 111 shows that $A \cup B$ is connected. Of course, $x_1, x_3 \in A \cup B$ so that $x_1 \underset{X}{\sim} x_3$. ■

DEFINITION *Let X be a metric space. The equivalence classes of the relation $\underset{X}{\sim}$ are called* ***components*** *. For an element $x \in X$, the* ***component of*** *x means the equivalence class containing x. In an obvious way, the components of X are the maximal connected subsets of X.*

The following is an immediate consequence of Proposition 108.

PROPOSITION 113 *Let X be a metric space and let C be a component of X. Then C is closed in X.*

This is a good opportunity to prove the following Proposition.

PROPOSITION 114 *Every open subset of $\mathbb{R}$ is a countable disjoint union of open intervals.*

Proof. Let V be an open subset of $\mathbb{R}$. We consider V as a metric space in its own right. We can write V as a disjoint union of its components. Let U be a typical component of V. Then by the previous result, U is an interval. We claim that U is open. Let $x \in U$. Then $x \in V$ and, since V is open in $\mathbb{R}$, there exists $\epsilon > 0$ such that $]x-\epsilon, x+\epsilon[\subseteq V$. But $]x-\epsilon, x+\epsilon[$ is a connected set and hence must lie in the same component as x. This shows that $]x-\epsilon, x+\epsilon[\subseteq U$. Hence U is open. Finally, since each open interval must contain a rational number, select in each component a rational. Since $\mathbb{Q}$ is countably infinite, it is clear that the number of components of V is countable. ∎

How can we recognize components? In general it is not always easy. The following Lemma is sometimes useful.

LEMMA 115 *Let C be a nonempty subset of a metric space X which is simultaneously open, closed and connected. Then C is a component of X.*

Proof. Let C be a nonempty connected open closed subset of X. Let Y be the component of X containing C. Then, since $X = C \cup (X \setminus C)$ is a splitting of X and Y is connected we find that either $Y \subseteq C$ or $Y \subseteq X \setminus C$). Since C is nonempty, and $C \subseteq Y$ the second alternative is not possible. Hence $C = Y$ and C is a component. ∎

EXAMPLE Consider the subset $X = \{0\} \cup \{\frac{1}{n}; n \in \mathbb{N}\}$ of $\mathbb{R}$. Clearly each of the singletons $\{\frac{1}{n}\}$ is relatively open and relatively closed in X and also connected. Hence each set $\{\frac{1}{n}\}$ is a component of X. When these are removed from X we are left just with $\{0\}$ which is clearly connected. Hence $\{0\}$ is also a component. □

EXAMPLE A variation on the preceding example is

$$X = \bigcup_{j=0}^{\infty} I_j$$

where $I_0 = [-1, 0]$ and $I_j = [\frac{1}{2j+1}, \frac{1}{2j}]$ for $j \in \mathbb{N}$. Each of the intervals I_j with $j \in \mathbb{N}$ is open, closed and connected in X and hence a component. The remaining set I_0 is clearly connected and hence it too must be a component. □

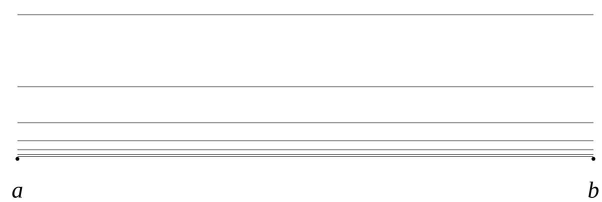

Figure 6.1: Example of distinct components that cannot be split.

EXAMPLE In $\mathbb{R}^2$ let $a = (-1, 0)$, $b = (1, 0)$. Let I_k be the closed line segment joining $(-1, 2^{-k})$ to $(1, 2^{-k})$. Finally, let

$$A = \{a, b\} \cup (\bigcup_{k=1}^{\infty} I_k).$$

The components of A are $\{a\}$, $\{b\}$ and I_k for $k \in \mathbb{N}$. What are the splittings of A? Suppose that one of the splitting sets A_1 contains a. Then, since A_1 is open (in A), it also contains a tail of the sequence $(-1, 2^{-k})$. But, since each I_k is connected, it also contains a tail of the I_k and in particular a tail of the sequence $(1, 2^{-k})$. Finally, since A_1 is closed (in A), $b \in A_1$. It is now not difficult to see that

$$A_2 = \bigcup_{k \in F} I_k$$

and

$$A_1 = A \setminus A_2,$$

where F is some finite subset of $\mathbb{N}$. Certainly, the points a and b lie on the same side of every splitting, despite the fact that $\{a\}$ and $\{b\}$ are distinct components. This shows that the converse of Lemma 107 is false. The subset $\{a\} \cup \{b\}$ of A lies on the same side of every splitting of A, but is not a connected subset of A. □

EXAMPLE Let E be the Cantor set in $\mathbb{R}$ (page 20). The components of the Cantor set are all singletons. Let a and b be distinct points of the Cantor set. Suppose without loss of generality that $a < b$. Let $\epsilon = b - a > 0$. Now select n such that $3^{-n} < \frac{1}{2}\epsilon$. The intervals of E_n have length 3^{-n} so clearly, a and b must belong to different constituent subintervals of E_n. Thus we may select $c \notin E_n$ with $a < c < b$. Now apply Proposition 106 with $V_1 =]-\infty, c[$ and $V_2 =]c, \infty[$. Since $a \in V_1$ and $b \in V_2$ it follows that there can be no connected subset of E that contains both a and b. □

DEFINITION *A metric space is said to be **totally disconnected** iff every component is a singleton.*

6.4 Compactness and Connectedness

There is a subtle interplay between compactness and connectedness.

PROPOSITION 116 *Let K be a compact metric space and let C be a component in K. Let $\mathcal{V}$ be the collection of all simultaneously open and closed subsets of K containing C. Then*

$$C = \bigcap_{V \in \mathcal{V}} V$$

Proof. Let us define $D = \cap_{V \in \mathcal{V}} V$. We will show that D is a connected subset of K. Certainly D is a closed subset of K because it is an intersection of closed sets. If it is not connected we can write

$$D = D_1 \cup D_2, \qquad \emptyset = D_1 \cap D_2 \tag{6.10}$$

where D_1 and D_2 are non-empty subsets of D simultaneously open and closed in D. Since D is closed in K and D_j is closed in D it follows that D_j is closed in K for $j = 1, 2$. Since D_1 and D_2 are also disjoint, it is therefore possible to separate them with sets U_1 and U_2 open in K (see Corollary 59). We have

$$U_1 \cap U_2 = \emptyset \qquad \text{and} \qquad D_j \subseteq U_j,\ j = 1, 2.$$

Then

$$K \setminus (U_1 \cup U_2) \subseteq K \setminus (D_1 \cup D_2)$$

$$
\begin{aligned}
&= K \setminus D \\
&= \bigcup_{V \in \mathcal{V}} (K \setminus V)
\end{aligned}
$$

Since $K \setminus (U_1 \cup U_2)$ is a closed subset of K it is compact and hence we can find a finite subset $\mathcal{F} \subseteq \mathcal{V}$ such that

$$
K \setminus (U_1 \cup U_2) \subseteq \bigcup_{V \in \mathcal{F}} (K \setminus V). \tag{6.11}
$$

Let us define

$$
W = \bigcap_{V \in \mathcal{F}} V.
$$

A moment's thought convinces us that $W \in \mathcal{V}$. We can restate (6.11) as $W \subseteq U_1 \cup U_2$. Clearly $W \cap U_1 = W \cap (K \setminus U_2)$ is the intersection of two closed sets and hence closed (as well as open). Similarly, $W \cap U_2$ is simultaneously open and closed. Thus the three sets $W \cap U_1$, $W \cap U_2$ and $K \setminus W$ form a three way splitting of K. Since C is connected it lies entirely in one of the sets of the splitting. Since $C \subseteq W$ we can assume without loss of generality that $C \subseteq W \cap U_1$. But then $W \cap U_1 \in \mathcal{V}$ and it follows that $D \subseteq U_1$. It follows that in the splitting (6.10), $D_2 = \emptyset$. The contradiction shows that D is connected. Finally by the maximality of C among connected subsets of K we see that $D = C$. ■

6.5 Preservation of Connectedness by Continuous Mappings

One of the most important properties of connectedness is that it is preserved by continuous mappings.

THEOREM 117 *Let X and Y be metric spaces. Suppose that X is connected. Let $f : X \longrightarrow Y$ be a continuous surjection. Then Y is also connected.*

Proof. Let

$$
\begin{aligned}
Y &= Y_1 \cup Y_2, \\
\emptyset &= Y_1 \cap Y_2,
\end{aligned}
$$

be a splitting of Y. Then it is easy to see that

$$
\begin{aligned}
X &= f^{-1}(Y_1) \cup f^{-1}(Y_2), \\
\emptyset &= f^{-1}(Y_1) \cap f^{-1}(Y_2),
\end{aligned}
$$

is a splitting of X. But since X is connected, the second splitting must be trivial. It follows that the first splitting is also trivial. ■

COROLLARY 118 *Let X and Y be metric spaces and let $f : X \longrightarrow Y$ be a continuous mapping. Let A be a connected subset of X. Then the direct image $f(A)$ is a connected subset of Y.*

COROLLARY 119 (INTERMEDIATE VALUE THEOREM) *Let $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous mapping. Let x be between $f(a)$ and $f(b)$. Then $x \in f([a, b])$.*

Proof. Let us suppose without loss of generality that $f(a) \leq f(b)$. Then the statement that x is between $f(a)$ and $f(b)$ implies that $f(a) \leq x \leq f(b)$. By the previous Corollary, $f([a, b])$ is a connected subset of $\mathbb{R}$. By Theorem 110 (page 131), $f([a, b])$ is an interval. Since this set contains $f(a)$ and $f(b)$, it also must contain the interval $[f(a), f(b)]$. It follows that x is in the direct image of f. ■

The following definition is long overdue.

DEFINITION *A metric space X is **discrete** iff every subset of X is open. A subset A of a metric space X is discrete iff it is discrete as a metric space with the restriction metric.*

Because of Theorem 5 a metric space is discrete iff its singletons are open.

EXAMPLE The subset $\mathbb{Z}$ is a discrete subset of $\mathbb{R}$. Indeed, given $n \in \mathbb{Z}$ we have $\{n\} = \mathbb{Z} \cap]n - \frac{1}{2}, n + \frac{1}{2}[$ showing that $\{n\}$ is open in $\mathbb{Z}$. □

Obviously a discrete metric space is totally disconnected. The Cantor set is a space that is totally disconnected but not discrete. The following Proposition involves an important technique.

PROPOSITION 120 *Let X be a connected metric space. Let Y be a discrete metric space. Let $f : X \longrightarrow Y$ be a continuous mapping. Then f is a constant mapping.*

Proof. By Theorem 117, the direct image $f(X)$ is connected and hence contained in a single component of Y. But the only components of Y are singletons. Hence f is a constant mapping. ■

It will be noted that the same conclusion would hold if Y we replace the hypothesis Y discrete by Y totally disconnected. However this extension is seldom used in practice.

6.6 Path Connectedness

There is a strong form of connectedness called path connectedness which is sometimes useful.

DEFINITION *A metric space X is **path connected** if for every pair of points $x_0, x_1 \in X$ there is a path from x_0 to x_1. Such a **path** is a continuous mapping $f : [0,1] \longrightarrow X$ such that $f(0) = x_0$ and $f(1) = x_1$.*

THEOREM 121 *Every path connected space is connected.*

Proof. Let X be a path connected metric space. We show that X has but one component. Let $x_0, x_1 \in X$, we will show that x_0 and x_1 lie in the same component of X. Let f be a path joining x_0 to x_1. Then, by Theorem 117, the underlying set $f([0,1])$ of the path is a connected set to which both x_0 and x_1 belong. Thus x_0 and x_1 belong to the same component of X. ■

We leave the following Lemma as an exercise for the reader. It is an easy application of the Glueing Theorem (page 43).

LEMMA 122 *Let X be a metric space and let $x_0, x_1, x_2 \in X$. Suppose that there is a path joining x_0 and x_1. Suppose that there is a path joining x_1 and x_2. Then there is a path joining x_0 and x_2.*

THEOREM 123 *Let V be a connected open subset of $\mathbb{R}^d$. Then V is path connected.*

Proof. If V is empty, there is nothing to show. Let $x_0 \in V$. Then, by Lemma 122, it is enough to show that for every $x \in V$ there is a path in V joining x_0 to x. Let W be the set of all such points x. Then by considering the constant path with value x_0 we see that $x_0 \in W$.

Claim: W ***is open in*** V Let $x_1 \in W$. Then, since V is open, there exists $\epsilon > 0$ such that $U(x_1, \epsilon) \subseteq V$. In particular, for every point x of $U(x_1, \epsilon)$ the line segment joining x_1 to x lies in V. The function $g : [0, 1] \longrightarrow V$ given by

$$g(t) = (1 - t)x_1 + tx \qquad \forall t \in [0, 1]$$

is a path from x_1 to x lying entirely in V. But, since $x_1 \in W$ there is a path in V from x_0 to x_1 so by another application of Lemma 122, there is a path in V from x_0 to x. Hence $U(x_1, \epsilon) \subseteq W$. Thus W is open in V.
Claim: W ***is closed in*** V Let $x_1 \in V \setminus W$. Then, repeating the previous argument, there exists $\epsilon > 0$ such that $U(x_1, \epsilon) \subseteq V$ and furthermore for every point x of $U(x_1, \epsilon)$ there is a path from x_1 to x lying entirely in V. It again follows from Lemma 122 that if $x \in W$ then $x_1 \in W$. But $x_1 \notin W$, so it follows that $U(x_1, \epsilon) \subseteq V \setminus W$. Hence $V \setminus W$ is open in V. It follows that W is closed in V.

Since V is connected and W is a non-empty open closed subset of V it follows that $W = V$. ∎

We remark that the proof actually shows that V is connected then any two points x_0 and x of V can be joined by a path consisting of finitely many line segments — that is a ***piecewise linear*** path. Also, there is nothing special about $\mathbb{R}^d$ here, the same proof would work in any real normed vector space.

EXAMPLE A standard example of a space that is connected but not path connected is the subset $A = Y \cup S$ of $\mathbb{R}^2$ where Y is the line segment

$$Y = \{(0, y); -1 \leq y \leq 1\}$$

lying in the y-axis and S is the union of the following line segments.

- From $(2^{-n}, -1)$ to $(2^{-n}, 1)$ for $n \in \mathbb{Z}^+$.
- From $(2^{-(n+1)}, -1)$ to $(2^{-n}, -1)$ for $n \in \mathbb{Z}^+$, n even.
- From $(2^{-(n+1)}, 1)$ to $(2^{-n}, 1)$ for $n \in \mathbb{Z}^+$, n odd.

The sets Y and S are shown on the left in Figure 6.2. It is clear that both Y and S are path connected and hence connected. Thus, the only possible non-trivial splitting of A is $A = Y \cup S$ (or its reversal) and it is easy to see that S is not closed in A since for example the sequence of points $((2^{-n}, 0))$ of S converges to the point $(0, 0)$ of Y. It follows that A is connected.

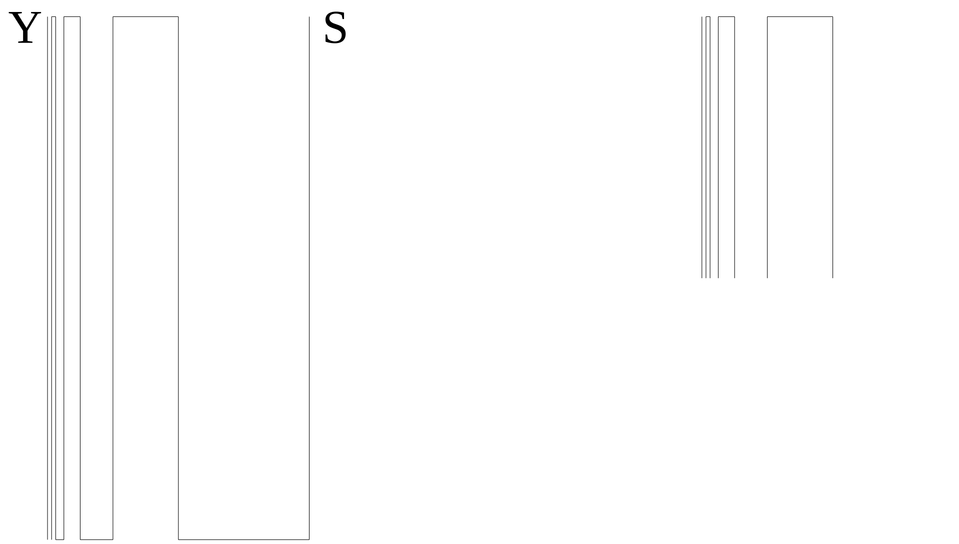

Figure 6.2: A connected space that is not path connected (left) and a horizontal swath cut through the same space (right).

To see that A is not path connected is harder. We first study the connectivity of the intersection of A with a narrow horizontal strip. Let $a < b$ and suppose that $[-1, 1]$ is not contained in $[a, b]$. Let H be the horizontal strip

$$H = \{(x, y); a \leq y \leq b\}.$$

In case $a = 0$, $b = 1$ the set $A \cap H$ is shown on the right in Figure 6.2. We leave the reader to show that $Y \cap H$ is a component in $A \cap H$. The method is similar to that used in the two examples following Lemma 115.

Let now $f : [0, 1] \longrightarrow A$ be a path from $(0, 0)$ to $(1, 0)$ lying in A. Define

$$t = \inf_{f(s) \in S} s.$$

Informally, t is the *first time* that the path jumps from Y to S. If $t = 0$ then $f(t) \in Y$. On the other hand, if $t > 0$ then $f(s) \in Y$ for $0 \leq s < t$ and it follows by continuity of f and the fact that Y is closed that $f(t) \in Y$. Since f is continuous at t there exists $\delta > 0$ such that

$$|s - t| \leq \delta \qquad \Rightarrow \qquad d(f(s), f(t)) \leq \tfrac{1}{2}.$$

Thus taking η to be the y-coordinate of $f(t)$ and setting $a = \eta - \frac{1}{2}$ and $b = \eta + \frac{1}{2}$ we see that the restriction of f to $[t - \delta, t + \delta]$ is a continuous mapping taking values in $A \cap H$. But since $f(t) \in Y \cap H$ and $Y \cap H$ is a component of $A \cap H$ we see that $f(s) \in Y \cap H$ for all s in $[t - \delta, t + \delta]$. This contradicts the definition of t. □

6.7 Separation Theorem for Convex Sets

In this section we tackle the Separation Theorem for Convex Sets stated on page 63. At first glance, there does not seem to be much of a connection between this topic and connectedness. To show the hidden connection we start with the following proposition.

PROPOSITION 124 *Let C be an open convex subset of V a finite dimensional real normed vector space of dimension at least 2. Suppose that $0_V \notin C$. Then there exists a line L through 0_V (that is a one-dimensional linear subspace L of V) such that $L \cap C = \emptyset$.*

Before tackling the proof we need the following Lemma.

LEMMA 125 *Let V be a real normed vector space of dimension at least 2 and let $S = \{v; v \in V, \|v\| = 1\}$ be its unit ball. Then S is a connected subset of V.*

Proof. Let $v_0, v_1 \in S$. Then consider

$$v(t) = \|(1-t)v_0 + tv_1\|^{-1}((1-t)v_0 + tv_1)$$

for $0 \leq t \leq 1$. If the vector $(1-t)v_0 + tv_1$ is non-zero for all $t \in [0,1]$ then $t \longrightarrow v(t)$ is a continuous path from v_0 to v_1 lying in S. On the other hand, if $(1-t)v_0 + tv_1 = 0_V$, then $(1-t)v_0 = -tv_1$, and taking the norm of both sides leads to $(1-t) = t$ so that $t = \frac{1}{2}$ and it follows that $v_1 = -v_0$. Thus, either there is a continuous path in joining v_0 to v_1 or v_0 and v_1 are antipodal points of S.

Now suppose that v_0, v_1 and v_2 are three *distinct* points of S. Then at most one of the pairs $\{v_0, v_1\}$, $\{v_1, v_2\}$ and $\{v_2, v_0\}$ is antipodal and it follows that we can connect v_0 and v_1 by a continuous path in S either directly, or by using Lemma 122. Finally it is easy to see that if $\dim(V) \geq 2$ then S possesses at least 3 distinct points and the proof is complete. Of course, in case $\dim(V) = 1$, then S is a 2 point disconnected set. ■

Proof of Proposition 124. We define the following subset A of S.

$$A = \{u; u \in S, \exists t > 0 \text{ such that } tu \in C\}.$$

Let $u \in A$, and let $t > 0$ be such that $tu \in C$. Since C is open in V there exists $s > 0$ such that $U(tu, s) \subseteq C$. Now let $v \in S$ be such that $\|v - u\| < st^{-1}$. Then clearly $\|tv - tu\| < s$, so that $tv \in C$ and $v \in A$. We have just shown that A is relatively open in S.

Now define

$$B = \{u; u \in S, -u \in A\}.$$

Then B is relatively open in S since A is. Also $A \cap B = \emptyset$ since if there exist $t_1 > 0$ and $t_2 < 0$ such that $t_1 u, t_2 u \in C$ then it follows from the convexity of C that $0_V \in C$ which is contrary to hypothesis.

If C is empty then the result is obvious. Hence we may assume that $C \neq \emptyset$ and it follows that both A and B are nonempty. The scenario $S = A \cup B$ is now ruled out by the connectivity of S. Hence there exists $u \in S \setminus (A \cup B)$, and the line L through 0_V and u does not meet C. The point tu is not in C since

- $u \notin A$ if $t > 0$,
- $u \notin B$ if $t < 0$,
- $tu = 0_V \notin C$ if $t = 0$.

■

DEFINITION *Let X and Y be metric spaces and let $\varphi : X \longrightarrow Y$. Then φ is an* ***open mapping*** *iff the direct image $\varphi(\Omega)$ is an open subset of Y for every open subset Ω of X.*

Proof of Theorem 43. Without loss of generality, we may suppose that $v = 0_V$. For this it suffices to apply a translation.

The proof is by induction on the dimension of V. If V is one-dimensional then the result is obvious, since an open convex set in $\mathbb{R}$ is just an open interval. Thus, we may suppose that $n \geq 2$, that the result is proved for vector spaces of dimension $n - 1$ and establish it in case $\dim(V) = n$. Since $n \geq 2$ we can apply Proposition 124 to find a one-dimensional subspace L of V that does not meet C. Consider now the quotient vector space $Q = V/L$ and let π be the canonical projection $\pi : V \longrightarrow Q$.

The direct image $\pi(C)$ is clearly a convex subset of Q. To see this, let $q_1, q_2 \in \pi(C)$. Then we can find lifts $v_1, v_2 \in C$ such that $\pi(v_j) = q_j$ for $j = 1, 2$. Let t_1, t_2 be nonnegative real numbers such that $t_1 + t_2 = 1$. Then, by the convexity of C we find that $t_1 v_1 + t_2 v_2 \in C$ and it follows that $t_1 q_1 + t_2 q_2 = \pi(t_1 v_1 + t_2 v_2) \in \pi(C)$.

We next claim that $\pi(C)$ is an open subset of Q. In fact, π is an open mapping. Let Ω be an arbitrary open subset of V and let $q_0 \in \pi(\Omega)$. Then, there exist $v_0 \in V$ such that $v_0 \in \Omega$ and $\pi(v_0) = q_0$. Since Ω is open in V, there exists $t > 0$ such that $U(v_0, t) \subseteq \Omega$. Now, let $q \in U(q_0, t)$. Since $\|q - q_0\|_Q < t$, and by the definition of the quotient norm (page 68), there exist $w \in V$ such that $\|w\|_V < t$ and $\pi(w) = q - q_0$. It now follows that $v = v_0 + w \in \Omega$ and that $\pi(v) = q$. Thus we have shown that $U(q_0, t) \subseteq \pi(\Omega)$. Since q_0 was an arbitrary point of $\pi(\Omega)$, it follows that $\pi(\Omega)$ is open in Q.

Finally, since $0_Q \notin \pi(C)$ (because $L \cap C = \emptyset$), we can apply the inductive hypothesis to obtain the existence of a linear form φ on Q such that

$$\varphi(q) < \varphi(0_Q) = 0 \qquad \forall q \in \pi(C).$$

It follows immediately that

$$\varphi \circ \pi(v) < 0 = \varphi \circ \pi(0_V) \qquad \forall v \in C.$$

Since $\varphi \circ \pi$ is a linear form on V, this completes the inductive step. ■

7

The Differential

In a single variable, differential calculus is seen as the study of limits of quotients of the type
$$\frac{f(v) - f(v_0)}{v - v_0}.$$
In several variables this approach no longer works. We need to view the derivative at v_0 as a linear map df_{v_0} such that we have

$$f(v) = f(v_0) + df_{v_0}(v - v_0) + \text{ error term.}$$

Here, the quantity $f(v)$ has been written as the sum of three terms. The term $f(v_0)$ is the constant term. It does not depend on v. The second term $df_{v_0}(v - v_0)$ is a linear function df_{v_0} of $v - v_0$. Finally the third term is the error term. The linear map df_{v_0} is called the ***differential*** of f at v_0. Sometimes we collect together the first and second terms as an ***affine*** function of v. A function is affine if and only if it is a constant function plus a linear function. This then is the key idea of differential calculus. We attempt to approximate a given function f at a given point v_0 by an affine function within an admissible error. Which functions are admissible errors for this purpose? We answer this question in the next section.

There are two settings that we can use to describe the theory. We start out using abstract real normed vector spaces. However as soon as one is faced with real problems in finitely many dimensions one is going to introduce coordinates — i.e. one selects bases in the vector spaces and works with the coordinate vectors. This leads to the second concrete setting which interprets differentials by Jacobian matrices.

7.1 The Little "o" of the Norm Class

Let V and W be real normed vector spaces.

DEFINITION *Let $\Omega \subseteq V$ be an open set and let $v_0 \in \Omega$. Then a function $\varphi : \Omega \longrightarrow W$ is in the class $\mathcal{E}_{\Omega,v_0}$ called* ***little "o" of the norm*** *at v_0 iff for all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\|\varphi(v)\| \leq \epsilon\|v - v_0\|$$

for all $v \in \Omega$ with $\|v - v_0\| < \delta$.

It is clear from the definition that if $\varphi \in \mathcal{E}_{\Omega,v_0}$ then $\varphi(v_0) = 0$.

If we replace the norms on V and W by equivalent norms then it is clear that the class of functions $\mathcal{E}_{\Omega,v_0}$ does not change. This has an important corollary in the finite dimensional case. Since all norms on a finite dimensional real vector space are equivalent (see Corollary 84 on page 106), we see that if V and W are finite dimensional then the class $\mathcal{E}_{\Omega,v_0}$ is completely independent of the norms on V and W. In other words, the class $\mathcal{E}_{\Omega,v_0}$ is an invariant of the linear space structure of V and W.

The following Lemma is very important for the definition of the differential. It tells us that we can distinguish between a linear function of $v - v_0$ and an admissible error function.

LEMMA 126 *Let $\Omega \subseteq V$ be an open set and let $v_0 \in \Omega$. Let $\varphi : \Omega \longrightarrow W$ be given by*

$$\varphi(v) = \lambda(v - v_0) \qquad \forall v \in \Omega$$

where $\lambda : V \longrightarrow W$ is a linear mapping. Suppose that $\varphi \in \mathcal{E}_{\Omega,v_0}$. Then $\varphi(v) = 0$ for all $v \in \Omega$.

Proof. Let $u \in V$. Then for all $\epsilon > 0$ we have

$$\|\varphi(v_0 + tu)\| \leq \epsilon\|tu\|$$

for all values of t such that $|t|$ is small enough. Using the specific form of φ we obtain

$$\|\lambda(tu)\| \leq \epsilon\|tu\|.$$

Using the linearity and the definition of the norm, this leads to

$$|t|\|\lambda(u)\| \leq \epsilon|t|\|u\|.$$

Choosing now t small and non-zero, we find that

$$\|\lambda(u)\| \leq \epsilon\|u\|.$$

Since this is true for all $\epsilon > 0$ we have $\lambda(u) = 0$. But this holds for all $u \in V$ and the result follows. ■

The next Proposition is routine and will be used heavily in these notes.

PROPOSITION 127 *Let $\Omega \subseteq V$ be an open set and let $v_0 \in \Omega$. Then $\mathcal{E}_{\Omega,v_0}$ is a vector space under pointwise addition and scalar multiplication.*

We leave the proof to the reader.

7.2 The Differential

In this section, U, V and W are real normed vector spaces.

DEFINITION *Let $\Omega \subseteq V$ be an open set and let $v_0 \in \Omega$. Then a function $f : \Omega \longrightarrow W$ is* ***differentiable at*** *v_0 with differential df_{v_0} (a continuous linear map from V to W) iff there exists a function $\varphi : \Omega \longrightarrow W$ in the class $\mathcal{E}_{\Omega,v_0}$ such that*

$$f(v) = f(v_0) + df_{v_0}(v - v_0) + \varphi(v) \qquad \forall v \in \Omega. \tag{7.1}$$

In this situation, the quantity df_{v_0} is called the ***differential*** of f at v_0. Notice that we insist that df_{v_0} is a *continuous* linear map. At this point that does not seem very important, but it would be impossible to progress much further without that important assumption. Of course, in the finite dimensional setting, all linear functions are continuous and this additional condition is irrelevant.

It should be pointed out that in classical variational calculus (which embraces the infinite dimensional setting) the function df_{v_0} is also called the ***first variation.*** The modern approach is to subsume variational calculus into the mainstream subject.

It is an immediate consequence of Lemma 126 that if the derivative df_{v_0} exists then it is unique.

EXAMPLE If f is a continuous linear mapping from V to W, then it is everywhere differentiable and its derivative is given by

$$df_{v_0}(v) = f(v).$$

The error term is zero. □

EXAMPLE If α is a bilinear mapping $\alpha : \mathbb{R}^a \oplus \mathbb{R}^b \longrightarrow \mathbb{R}^k$, then we have

$$\begin{aligned} \alpha(x, y) &= \alpha(x_0 + (x - x_0), y_0 + (y - y_0)) \\ &= \alpha(x_0, y_0) + \alpha(x_0, y - y_0) + \alpha(x - x_0, y_0) + \alpha(x - x_0, y - y_0) \end{aligned} \tag{7.2}$$

The first term in (7.2) is the constant term, the second and third terms are linear. The last term is little "o" of the norm since

$$\|\alpha(x - x_0, y - y_0)\| \leq \|\alpha\|_{\text{op}} \|x - x_0\| \|y - y_0\|.$$

Here $\| \ \|_{\text{op}}$ stands for the bilinear operator norm. □

PROPOSITION 128 *Let $\Omega \subseteq V$ be an open set and let $f : \Omega \longrightarrow W$ be a function differentiable at $v_0 \in \Omega$. Then f is **Lipschitz at** v_0 in the sense that there exists $\delta > 0$ and $0 < C < \infty$ such that*

$$\|f(v) - f(v_0)\| \leq C\|v - v_0\|$$

whenever $v \in \Omega \cap U(v_0, \delta)$. In particular, f is continuous at v_0.

Proof. Using the notation of (7.1), we have

$$\|df_{v_0}(v - v_0)\| \leq \|df_{v_0}\|_{\text{op}} \|v - v_0\| \tag{7.3}$$

and for $\epsilon = 1$, there exists $\delta > 0$ such that

$$\|\varphi(v)\| \leq \|v - v_0\|. \tag{7.4}$$

for $v \in \Omega \cap U(v_0, \delta)$. Combining (7.3) and (7.4) with (7.1) we find

$$\|f(v) - f(v_0)\| \leq (\|df_{v_0}\|_{\text{op}} + 1)\|v - v_0\|$$

for $v \in \Omega \cap U(v_0, \delta)$ as required. ■

Proposition 128 has a partial converse.

PROPOSITION 129 *Let $\Omega \subseteq V$ be an open set and let $f : \Omega \longrightarrow W$ be a function differentiable at $v_0 \in \Omega$. Suppose that there exists $\delta > 0$ and $0 < C < \infty$ such that*

$$\|f(v) - f(v_0)\| \le C\|v - v_0\|$$

whenever $v \in \Omega \cap U(v_0, \delta)$. Then $\|df_{v_0}\|_{\mathrm{op}} \le C$.

Proof. We write

$$f(v) = f(v_0) + df_{v_0}(v - v_0) + \varphi(v - v_0)$$

where $\varphi \in \mathcal{E}_{\Omega,v_0}$. Let $\epsilon > 0$. The, there exists δ_1 with $0 < \delta_1 < \delta$ such that

$$v \in \Omega, \|v - v_0\|_V < \delta_1 \Longrightarrow \|\varphi(v - v_0)\|_W \le \epsilon\|v - v_0\|_V$$

and consequently, for $v \in \Omega$ with $\|v - v_0\|_V < \delta_1$ we find

$$\|df_{v_0}(v - v_0)\|_W \le (C + \epsilon)\|v - v_0\|_V.$$

Since $v - v_0$ is free to roam in a ball centered at 0_V, it follows that $\|df_{v_0}\|_{\mathrm{op}} \le C + \epsilon$. Finally, since ϵ is an arbitrary positive number, we have the desired conclusion. ∎

The following technical Lemma will be needed for the Chain Rule.

LEMMA 130 *Let $\Omega \subseteq V$ be an open set, Δ an open subset of W, and let $f : \Omega \longrightarrow \Delta$ be a function Lipschitz at $v_0 \in \Omega$. Let $\psi : \Delta \longrightarrow U$ be in $\mathcal{E}_{\Delta,f(v_0)}$. Then the composed function $\psi \circ f$ is in $\mathcal{E}_{\Omega,v_0}$.*

Proof. There exists $\delta_1 > 0$ and $0 < C < \infty$ such that

$$\|f(v) - f(v_0)\| \le C\|v - v_0\| \tag{7.5}$$

whenever $v \in \Omega \cap U(v_0, \delta_1)$. Let $\epsilon > 0$. Define $\epsilon_1 = C^{-1}\epsilon > 0$. Then since ψ is little "o" of the norm, there exists $\delta_2 > 0$ such that we have

$$\|\psi(w)\| \le \epsilon_1\|w - f(v_0)\|$$

provided $w \in \Delta$ and $\|w - f(v_0)\| < \delta_2$. Now define $\delta = \min(\delta_1, C^{-1}\delta_2) > 0$. Then, using (7.5), $v \in \Omega$ and $\|v - v_0\| < \delta$ together imply that $\|f(v) - f(v_0)\| < \delta_2$ and hence also

$$\|\psi(f(v))\| \le \epsilon_1\|f(v) - f(v_0)\| \le C\epsilon_1\|v - v_0\|.$$

Since $\epsilon = C\epsilon_1$, this completes the proof. ∎

THEOREM 131 (THE CHAIN RULE) *Let $\Omega \subseteq V$ be an open set, Δ an open subset of W, let $f : \Omega \longrightarrow \Delta$ be a function differentiable at $v_0 \in \Omega$ and let $g : \Delta \longrightarrow U$ be differentiable at $f(v_0)$. Then the composed function $g \circ f$ is differentiable at v_0 and*

$$d(g \circ f)_{v_0} = dg_{f(v_0)} \circ df_{v_0}.$$

Proof. We use the differentiability hypotheses to write

$$f(v) = f(v_0) + df_{v_0}(v - v_0) + \varphi(v) \qquad \forall v \in \Omega \tag{7.6}$$

and

$$g(w) = g(f(v_0)) + dg_{f(v_0)}(w - f(v_0)) + \psi(w) \qquad \forall w \in \Delta \tag{7.7}$$

where $\varphi : \Omega \longrightarrow W$ is in the class $\mathcal{E}_{\Omega,v_0}$ and $\psi : \Delta \longrightarrow U$ is in the class $\mathcal{E}_{\Delta,f(v_0)}$. Combining (7.6) and (7.7) yields

$$g(f(v)) = g(f(v_0)) + dg_{f(v_0)}(df_{v_0}(v - v_0) + \varphi(v)) + \psi(f(v)) \qquad \forall v \in \Omega.$$

Using the linearity of $dg_{f(v_0)}$ we can rewrite this in the form

$$g \circ f(v) = g \circ f(v_0) + (dg_{f(v_0)} \circ df_{v_0})(v - v_0) + dg_{f(v_0)}(\varphi(v)) + \psi(f(v)), \tag{7.8}$$

for all $v \in \Omega$. The first term on the right of (7.8) is constant and the second term is continuous linear because it is the composition of two continuous linear functions. Since $\mathcal{E}_{\Omega,v_0}$ is a vector space, it suffices to show that the third and fourth terms on the right of (7.8) are in $\mathcal{E}_{\Omega,v_0}$. For $dg_{f(v_0)}(\varphi(v))$ this is a consequence of the continuity of $dg_{f(v_0)}$, and for $\psi(f(v))$ it is a consequence of Lemma 130. ■

There is no product rule as such in the multivariable calculus, because it is not clear which product one should take.

EXAMPLE For the most general case of the product rule, α is a bilinear mapping $\alpha : \mathbb{R}^a \times \mathbb{R}^b \longrightarrow \mathbb{R}^k$. Let now Ω be open in V and let $x_0 \in \Omega$. Let f and g be mappings from Ω into $\mathbb{R}^a$ and $\mathbb{R}^b$ respectively differentiable at x_0. Then let

$$h(x) = \alpha(f(x), g(x)) \qquad \forall x \in \Omega.$$

Applying the chain rule and using the derivative of α found earlier, we find that h is differentiable at x_0 and the derivative is given by

$$dh_{x_0}v = \alpha(f(x_0), dg_{x_0}v) + \alpha(df_{x_0}v, g(x_0)).$$

□

7.3 Derivatives, Differentials and Directional Derivatives

We have already seen how to define the derivative of a vector valued function on page 36. How does this definition square with the concept of differential given in the last chapter? Let V be a general normed vector space, $g :]a, b[\longrightarrow V$ and t a point of $]a, b[$. Then, it follows directly from the definitions of derivative and differential that the existence of one of $f'(t)$ and df_t implies the existence of the other, and

$$df_t(1) = f'(t).$$

This formula reconciles the fundamental difference between $f'(t)$ and df_t, namely that $f'(t)$ is a vector and df_t is a linear transformation. In effect, the existence of the limit

$$f'(t) = \lim_{s\to t}(s-t)^{-1}(f(s)-f(t))$$

as an element of V, is the same as showing that the quantity

$$f(s) - (f(t) + (s-t)f'(t))$$

is little "o" of $s - t$. Thus, $df_t(s-t) = (s-t)f'(t)$ or equivalently $df_t(1) = f'(t)$.

For a one-dimensional domain, the concepts of derivative and differential are closely related. We can attempt to understand the case in which the domain is multidimensional by restricting the function to lines. Let us suppose that Ω be an open subset of a normed vector space U and that $u_0 \in \Omega, u_1 \in U$. We can then define a function $g : \mathbb{R} \longrightarrow U$ by $g(t) = u_0 + tu_1$. The function g parametrizes a line through u_0. We think of u_1 as the ***direction vector*** , but this term is a misnomer because the magnitude of u_1 will play a role. For $|t|$ small enough, $g(t) \in \Omega$. Hence, if $f : \Omega \longrightarrow V$ is a differentiable function, the composition $f \circ g$ will be differentiable in some neighbourhood of 0 and

$$(f\circ g)'(0) = d(f\circ g)_0(1) = df_{u_0}dg_0(1) = df_{u_0}g'(0) = df_{u_0}(u_1). \tag{7.9}$$

since both g and $f \circ g$ are defined on a one-dimensional space. Equation (7.9) allows us to understand what $df_{u_0}(u_1)$ means, but unfortunately it cannot be used to define the differential.

DEFINITION *The **directional derivative** $D_{u_1}f(u_0)$ of the function f at the point u_0 in the direction u_1 is defined as the value of $(f\circ g)'(0)$ if this exists. In symbols*

$$D_{u_1}f(u_0) = \lim_{s\to 0} s^{-1}(f(u_0+su_1) - f(u_0)). \tag{7.10}$$

Clearly, in case f is differentiable, we can combine (7.9) and (7.10) to obtain

$$df_{u_0}(u_1) = D_{u_1} f(u_0). \tag{7.11}$$

EXAMPLE Consider the function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

It is easy to check that f is linear on every line passing through the origin $(0, 0)$. Hence the directional derivative $D_{(\xi,\eta)} f(0, 0)$ exists for every direction vector $(\xi, \eta) \in \mathbb{R}^2$. In fact, it comes as no surprise that

$$D_{(\xi,\eta)} f(0, 0) = \begin{cases} \frac{\xi^2 \eta}{\xi^2+\eta^2} & \text{if } (\xi, \eta) \neq (0, 0), \\ 0 & \text{if } (\xi, \eta) = (0, 0). \end{cases}$$

and this is *not* a linear function of (ξ, η) and therefore cannot possibly be equal to $df_{(0,0)}(\xi, \eta)$ which would necessarily have to be linear in (ξ, η). It follows from (7.9) that $df_{(0,0)}$ cannot exist. □

7.4 The Mean Value Theorem

We now return to the one-dimensional case to discuss the Mean Value Theorem which is of central importance.

DEFINITION *Let X be a metric space and let $f : X \longrightarrow \mathbb{R}$. Let $x_0 \in X$. Then x_0 is a* ***local maximum point*** *for f iff there exists $t > 0$ such that $f(x) \leq f(x_0)$ for all $x \in U(x_0, t)$. The concept* ***local minimum point*** *is defined similarly.*

LEMMA 132 *Let a and b be real numbers such that $a < b$. Let $f :]a, b[\longrightarrow \mathbb{R}$ be a differentiable mapping. Let $\xi \in]a, b[$ be a local maximum point for f. Then $f'(\xi) = 0$.*

Proof. Suppose not, we will provide a contradiction. Without loss of generality we can assume that $f'(\xi) > 0$. we leave the case $f'(\xi) < 0$ which is similar, to the reader.

Suppose that $f'(\xi) > 0$. We can write

$$f(x) = f(\xi) + f'(\xi)(x - \xi) + \phi(x)$$

where $\phi \in \mathcal{E}_{]a,b[,\xi}$. Choose $\epsilon = \frac{1}{2}f'(\xi)$, then there exist δ such that $0 < \delta < \min(b - \xi, \xi - a)$ such that $|\phi(x)| \le \epsilon|x - \xi|$ for $|x - \xi| < \delta$. It follow that

$$f(x) \ge f(\xi) + \tfrac{1}{2}f'(\xi)(x - \xi) \tag{7.12}$$

for $\xi \le x < \xi + \delta$. On the other hand, since ξ is a local maximum point for f there exists $t > 0$ such that

$$f(x) \le f(\xi) \tag{7.13}$$

for $\xi \le x < \xi + t$. It is easy to see that (7.12) and (7.13) are contradictory. ■

PROPOSITION 133 (ROLLE'S THEOREM) *Let a and b be real numbers such that $a < b$. Let $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous map. Suppose that f is differentiable at every point of $]a, b[$. Suppose that $f(a) = f(b)$. Then there exists ξ such that $a < \xi < b$ and $f'(\xi) = 0$.*

Proof. If f is constant, then any ξ such that $a < \xi < b$ will satisfy $f'(\xi) = 0$. Hence there is some point $x \in]a, b[$ such that $f(x) \ne f(a)$. There are two cases depending on whether $f(x) > f(a)$ or $f(x) < f(a)$. We deal only with the first case and leave the second, which is entirely similar to the reader.

Suppose that $f(x) > f(a)$. Then since f is a continuous real-valued function on a compact space it attains its maximum value by Theorem 83 (page 106). Suppose that this maximum is attained at $\xi \in [a, b]$. Then clearly

$$f(\xi) \ge f(x) > f(a) = f(b),$$

so that $\xi \ne a$ and $\xi \ne b$. It follows that $\xi \in]a, b[$ and hence $f'(\xi) = 0$ by Lemma 132. ■

THEOREM 134 (THE MEAN VALUE THEOREM) *Let a and b be real numbers such that $a < b$. Let $g : [a, b] \longrightarrow \mathbb{R}$ be a continuous map. Suppose that g is differentiable at every point of $]a, b[$. Then there exists ξ such that $a < \xi < b$ and*

$$g(b) - g(a) = g'(\xi)(b - a).$$

Proof. It suffices to apply Rolle's Theorem to the function given by

$$f(x) = (b-a)g(x) - (g(b) - g(a))(x-a).$$

It should be observed that $f(a) = (b-a)g(a) = f(b)$. ∎

The Mean Value Theorem has many Corollaries. For convenience we collect them together here.

COROLLARY 135 *Let a and b be real numbers such that $a < b$. Let $g :]a,b[\longrightarrow \mathbb{R}$ be a mapping differentiable at every point of $]a,b[$.*

- *If $g'(x) = 0$ for all $x \in]a,b[$, then g is constant on $]a,b[$.*
- *If $g'(x) \geq 0$ for all $x \in]a,b[$, then g is increasing on $]a,b[$ in the wide sense, (i.e. $g(x_2) \geq g(x_1)$ for $a < x_1 \leq x_2 < b$).*
- *If $g'(x) > 0$ for all $x \in]a,b[$, then g is strictly increasing on $]a,b[$, (i.e. $g(x_2) > g(x_1)$ for $a < x_1 < x_2 < b$).*
- *If $g'(x) \leq 0$ for all $x \in]a,b[$, then g is decreasing on $]a,b[$ in the wide sense, (i.e. $g(x_2) \leq g(x_1)$ for $a < x_1 \leq x_2 < b$).*
- *If $g'(x) < 0$ for all $x \in]a,b[$, then g is strictly decreasing on $]a,b[$, (i.e. $g(x_2) < g(x_1)$ for $a < x_1 < x_2 < b$).*

COROLLARY 136 *Let $\Omega \subseteq V$ be a connected open set of a real normed vector space V and let f be a differentiable function $f : \Omega \longrightarrow \mathbb{R}$. Suppose that $df_v = 0$ for all $v \in \Omega$. Then f is constant on Ω.*

Proof. Let v_0 and v be two points of Ω. Suppose that v_0 and v can be joined by a differentiable path in Ω. Then we claim that $f(v) = f(v_0)$. Let us denote the path by $\alpha : [0,1] \longrightarrow \Omega$. Then, applying the Chain rule to $f \circ \alpha$ we find that $(f \circ \alpha)'(t) = d(f \circ \alpha)_t(1) = df_{\alpha(t)} \circ d\alpha_t(1) = 0$. It now suffice to apply the Mean Value Theorem to verify the claim,

$$f(v_0) = f \circ \alpha(0) = f \circ \alpha(1) = f(v).$$

Theorem 123 tells us that there is a continuous path between v_0 and v. Unfortunately, this is not enough. In fact, it is always possible to have a differentiable path, but it is easier to use a piecewise linear path — see the remarks following the proof of Theorem 123. If v_0 and v can be joined by a line segment in Ω, then we have a differentiable path between the points and the argument outlined above shows that $f(v) = f(v_0)$. But, in any case there is a path consisting of finitely many line segments which joins v_0 to v. It then suffices to apply the above argument to each of these line segments in turn. ■

7.5 A Lipschitz Type Estimate

In this section we establish a Lipschitz type estimate for differentiable functions often associated with the Mean Value Theorem.

THEOREM 137 *Let $a < t_1 \leq t_2 < b$. Let V be a normed vector space and let $f :]a, b[\longrightarrow V$ be a differentiable function. Then*

$$\|f(t_2) - f(t_1)\| \leq (t_2 - t_1) \sup_{t \in [t_1, t_2]} \|f'(t)\|_V. \tag{7.14}$$

We offer one proof of this result and an almost-proof. Yet another almost-proof can be made using the Fundamental Theorem of Calculus to be established in the next chapter.

First Proof. Observe that the case $t_1 = t_2$ is obvious since both sides of (7.14) vanish. Thus we can assume that $t_1 < t_2$. Now it is easy to see that we can normalize so that $t_1 = 0$, $t_2 = 1$, $a < 0$ and $b > 1$.

Let

$$C = \sup_{t \in [0,1]} \|f'(t)\|_V,$$

and fix $\epsilon > 0$. Consider the set

$$A = \{t; t \in [0, 1], \|f(t) - f(0)\| \leq (C + \epsilon)t\}.$$

Clearly, A is a closed subset of $[0, 1]$ and $0 \in A$. Now define

$$B = \{s; s \in [0, 1], 0 \leq t \leq s \Rightarrow t \in A\}.$$

Then it is an exercise to see that B is also a closed set in $[0, 1]$. Furthermore $0 \in B$. We will show that B is also open in $[0, 1]$. It will then follow from the connectivity of $[0, 1]$ (by Proposition 105 on page 128) that $B = [0, 1]$. Thus $1 \in A$ yielding

$$\|f(1) - f(0)\| \leq C + \epsilon.$$

Since this holds for each strictly positive real number ϵ, it now follows that

$$\|f(1) - f(0)\| \le C,$$

which is just (7.14) normalized to the situation at hand.

Thus, all rests on showing that B is open in $[0, 1]$. Towards this, let $s \in B$. Then since $s \in A$ we have

$$\|f(s) - f(0)\| \le (C + \epsilon)s. \tag{7.15}$$

Now apply the definition of differentiability at s to find that

$$f(s + r) = f(s) + rf'(s) + \phi(r)$$

where $\phi(r)/r$ tends to 0_V as $r \longrightarrow 0$. Hence, for $\delta > 0$ chosen suitably small, we have

$$\|f(s + r) - f(s)\| \le (C + \epsilon)r. \tag{7.16}$$

for $0 \le r < \delta$. Combining (7.15) and (7.16) we obtain

$$\|f(s + r) - f(0)\| \le (C + \epsilon)(s + r).$$

It follows that $s + r \in A$ for $0 \le r < \delta$. Since $s \in B$ already implies that $t \in A$ for $0 \le t \le s$, it now follows that $t \in A$ for $0 \le t < s + \delta$. This implies that s is an interior point of B. ■

Second Proof.

Here we make the additional assumption that V is finite dimensional. This assumption can be circumvented, but to do so would take us too far afield. The idea is to use duality to reduce the problem to the one-dimensional case.

Let $\varphi \in V'$ with $\|\varphi\|_{V'} \le 1$. Here the norm taken is the dual norm (see (3.33) on page 63). Then let $h = \varphi \circ f$. Then $h :]a, b[\longrightarrow \mathbb{R}$. Applying the Mean Value Theorem (page 152) and taking absolute values, we find

$$|h(t_2) - h(t_1)| \le (t_2 - t_1) \sup_{t \in [t_1, t_2]} |h'(t)|$$

which can be rewritten as

$$|\varphi(f(t_2) - f(t_1))| \le (t_2 - t_1) \sup_{t \in [t_1, t_2]} |\varphi(f'(t))|.$$

Now take the supremum of both sides over all possible φ. Clearly, the two suprema on the right are independent and can be interchanged. Thus we get

$$\sup_{\|\varphi\|_{V'}\leq 1} |\varphi(f(t_2) - f(t_1))| \leq (t_2 - t_1) \sup_{t\in[t_1,t_2]} \sup_{\|\varphi\|_{V'}\leq 1} |\varphi(f'(t))|. \tag{7.17}$$

Finally, apply Proposition 42 to both sides of (7.17) to obtain (7.14). ■

A third "proof" uses the formula

$$f(t_2) - f(t_1) = \int_{t_1}^{t_2} f'(t)dt$$

from the third part of Lemma 158 in the next chapter. The additional assumptions that are needed to make this proof work are the completeness of V (an assumption which can be removed with a little additional work) and the continuity of f'.

The first Corolloary is the following.

COROLLARY 138 *Let $a < b$. Let V be a normed vector space and let $f :]a, b[\longrightarrow V$ be a differentiable function with everywhere vanishing drivative. Then f is constant on $]a, b[$.*

We now use Theorem 137 to establish the following Corollary.

COROLLARY 139 *Let U and V be normed vector spaces. Let $\Omega \subseteq U$ be a convex open set and let $f : \Omega \longrightarrow V$ be a function differentiable on Ω. Let u_1 and u_2 be points of Ω. Then*

$$\|f(u_2) - f(u_1)\| \leq \|u_2 - u_1\| \sup_{u\in L(u_1,u_2)} \|df_u\|_{\text{op}} \tag{7.18}$$

where $\| \ \|_{\text{op}}$ denotes the corresponding operator norm for linear endomorphisms from U to V. Here $L(u_1, u_2)$ denotes the line segment joining u_1 to u_2. This set must lie in Ω since Ω is a convex set.

Proof. Let $\theta(t) = (1 - t)u_1 + tu_2$ a mapping from an interval $]a, b[$ to Ω where a and b are suitably chosen real numbers with $a < 0$ and $b > 1$. Then, according to Theorem 137, we have

$$\|f \circ \theta(1) - f \circ \theta(0)\| \leq \sup_{t\in[0,1]} \|(f \circ \theta)'(t)\|_V. \tag{7.19}$$

Comparing with (7.9), we see that $(f\circ\theta)'(t)$ can be viewed as a directional derivative

$$(f \circ \theta)'(t) = D_{u_2-u_1} f(\theta(t)) = df_{\theta(t)}(u_2 - u_1). \tag{7.20}$$

We clearly have the operator norm estimate

$$\|df_{\theta(t)}(u_2 - u_1)\|_V \le \|u_2 - u_1)\|_U \|df_{\theta(t)}\|_{\text{op}}, \tag{7.21}$$

so that combining (7.20) and (7.21) yields

$$\|(f \circ \theta)'(t)\|_V \le \|u_2 - u_1)\|_U \|df_{\theta(t)}\|_{\text{op}} \tag{7.22}$$

Clearly, $\theta(0) = u_1$, $\theta(1) = u_2$ and $\theta(t) \in L(u_1, u_2)$. Hence, further combining (7.19) and (7.22) gives (7.18). ■

This in turn can be used to establish yet another result characterizing mappings with zero differential. The proof follows that of Corollary 136

COROLLARY 140 *Let U and V be real normed vector spaces. Let $\Omega \subseteq U$ be a connected open set of U and let f be a differentiable function $f : \Omega \longrightarrow V$. Suppose that $df_u = 0$ for all $u \in \Omega$. Then f is constant on Ω.*

7.6 One-sided derivatives and limited differentials

In general, we try to avoid discussing limited differentials, but there are times when their introduction is essential.

DEFINITION *Let V be a normed space and let $f : [0, \infty[\longrightarrow V$. If the limit*

$$\lim_{x \longrightarrow 0} x^{-1}(f(x) - f(0))$$

*exists as a limit in V over the metric space $[0, \infty[$ then we say f has a **one-sided derivative** $f'(0)$ at 0. Explicitly, this means that for all $\epsilon > 0$, there exists $\delta > 0$ such that $\|x^{-1}(f(x) - f(0)) - f'(0)\| < \epsilon$ whenever $0 < x < \delta$.*

We can make a similar definition at any "closed" endpoint of an interval. Obviously, there is an analogous definition of one-sided differential. In several variables, we might have a real-valued function defined for instance on a closed ball and want to define the differential at a boundary point of the ball. The most general concept is that of a limited differential. We need first to extend the concept of *little "o"*.

DEFINITION *Let $\Omega \subseteq V$ be any subset of a real normed vector space V and let $v_0 \in \Omega$. A function $\varphi : \Omega \longrightarrow V$ is in the class $\mathcal{E}_{\Omega,v_0}$ if and only if for all $\epsilon > 0$, there exist $\delta > 0$ such that $v \in \Omega$ and $\|v - v_0\| < \delta$ implies that*

$$\|\varphi(v)\| \leq \epsilon\|v - v_0\|.$$

Obviously, if v_0 is an isolated point of Ω then this condition is vacuous. Therefore the definition only has meaning if v_0 is an accumulation point of Ω.

DEFINITION *Let $\Omega \subseteq V$ be any set and let $v_0 \in \Omega$. Then a function $f : \Omega \longrightarrow W$ has a* ***limited differential*** *df_{v_0} (a continuous linear map from V to W) at v_0 if and only if there exists a function $\varphi : \Omega \longrightarrow W$ in the class $\mathcal{E}_{\Omega,v_0}$ such that*

$$f(v) = f(v_0) + df_{v_0}(v - v_0) + \varphi(v) \qquad \forall v \in \Omega. \tag{7.23}$$

Again, this definition is only useful if the geometry of v_0 and Ω is such that the differential df_{v_0} is necessarily unique. This is certainly true in most cases of interest, but not in general.

7.7 The Differential and Direct Sums

In this section we look briefly at the differential in the situation where the underlying vector spaces are direct sums. Let us suppose that

$$W = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

where W_j are complete normed spaces for $1 \leq j \leq k$. It is an extension of Proposition 67 that any norm on W that renders W complete and agrees with the norm of each W_j is necessarily equivalent to any of the standard p-norm combinations. Thus, from the point of view of differential calculus, provided that we insist on the completeness of W, the precise norm that is used on W is immaterial.

Let V be another normed space and let Ω be an open subset of V. Let $F^j : \Omega \longrightarrow W_j$ and let $F = \prec F^1, F^2, \ldots, F^k \succ$ be the mapping given by

$$F(v) = F^1(v) \oplus F^2(v) \oplus \cdots \oplus F^k(v).$$

The following Theorem is routine, and we leave the proof to the reader.

THEOREM 141 *Let $v_0 \in \Omega$. Then F is differentiable at v_0 if and only if each of the functions F^j is differentiable at v_0 and in either case*

$$dF_{v_0}(v) = dF^1_{v_0}(v) \oplus dF^2_{v_0}(v) \oplus \cdots \oplus dF^k_{v_0}(v),$$

or, more succinctly

$$dF_{v_0} = \prec dF^1_{v_0}, dF^2_{v_0}, \ldots, dF^k_{v_0} \succ$$

When the domain space is a direct sum however the situation becomes more complicated. Let now W be a fixed complete normed space and let

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

be a complete direct sum. The remarks above about equivalence of norms continue to apply.

Let again $F : \Omega \longrightarrow W$. Unfortunately the function F can no longer be considered as a k-tuple of functions. We can however define ***partial differentials*** that describe the behaviour of F on each V_j.

Let θ_j denote the map $\theta_j : V_j \longrightarrow V$ given by

$$\theta_j(u) = v_0 + u \qquad \forall u \in V_j.$$

Then θ_j is a continuous affine mapping. The continuity results from the fact that the norm of V when restricted to V_j agrees with the norm of V_j. In particular, θ_j is also an isometry. Thus $\Omega_j = \theta_j^{-1}(\Omega)$ is a neighbourhood of the zero vector of V_j. Considering θ_j as a mapping

$$\theta_j : \Omega_j \longrightarrow \Omega$$

we can form the composition $F \circ \theta_j : \Omega_j \longrightarrow W$.

DEFINITION *The **partial differential** $dF^j_{v_0}$ of F with respect to the j-th component V_j is precisely the differential $d(F \circ \theta_j)_{0_{V_j}}$.*

Assuming that F is differentiable at v_0, an immediate application of the chain rule (Theorem 131) gives

$$dF^j_{v_0} = d(F \circ \theta_j)_{0_{V_j}} = dF_{v_0} \circ \varphi_j,$$

where $\varphi_j = (d\theta_j)_{0_{V_j}}$ is just the canonical inclusion of V_j into V.

The question arises as to whether the existence of all the partial differentials $dF^j_{v_0}$ at a *fixed point* v_0 necessarily imply the existence of dF_{v_0}. The answer is negative. In fact the example with directional derivatives on page 151 is easily adapted to show this. However, partial differentials would be all but useless if something of this nature failed to be true. The key result is the following.

THEOREM 142 *Let W be a complete normed space and let*

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

be a complete direct sum. Let Ω be an open subset of V. Suppose that for each j with $1 \le j \le k$ and for all $v \in \Omega$ the partial differential dF^j_v exists. Suppose further that $v \longrightarrow dF^j_v$ is a continuous function from Ω into the space of continuous linear mappings from V_j to W with the operator norm. Then the differential dF_v exists at each point $v \in \Omega$ and the mapping $v \longrightarrow dF_v$ is a continuous function from Ω into the space of continuous linear mappings from V to W.

Proof. We prove the Theorem in the special case $k = 2$ and leave the proof in the general case to the reader. In fact, it suffices to apply a rather straightforward induction argument.

Since the precise norm used on V is immaterial, we can use the 1-standard norm $\|v_1 \oplus v_2\| = \|v_1\| + \|v_2\|$. We write $v = v_1 \oplus v_2$ with $v_j \in V_j$ for $j = 1, 2$. we denote a typical increment by $u = u_1 \oplus u_2$. Since $v \in \Omega$, there exists $\delta_1 > 0$ such that $v + u \in \Omega$ whenever $\|u\| < \delta_1$. We temporarily fix u_2 and define a function G from a suitable neighbourhood of 0_{V_1} in V_1 to W by

$$G(u_1) = F((v_1 + u_1) \oplus (v_2 + u_2)) - dF^1_v(u_1).$$

Then, by hypothesis we have

$$dG_{u_1} = dF^1_{v+(u_1 \oplus u_2)} - dF^1_v.$$

Now let $\epsilon > 0$. Then by continuity of dF^1, there exists $\delta_2 > 0$ such that $\|dF^1_{v+(u_1\oplus u_2)} - dF^1_v\|_{\text{op}} < \epsilon$ whenever $\|u_1 \oplus u_2\| < \delta_2$. We may as well assume here that $\delta_2 < \delta_1$. Since $\|dG_{u_1}\|_{\text{op}} < \epsilon$ we obtain from Corollary 139 applied to the line segment joining 0_{V_1} to u_1 in V_1 that

$$\|G(u_1) - G(0_{V_1})\| \le \epsilon \|u_1\|,$$

provided that $\|u_1 \oplus u_2\| = \|u_1\| + \|u_2\| < \delta_2$. Rewriting this in terms of F yields

$$\|F((v_1 + u_1) \oplus (v_2 + u_2)) - F(v_1 \oplus (v_2 + u_2)) - dF_v^1(u_1)\| \leq \epsilon \|u_1\| \tag{7.24}$$

whenever $\|u_1 \oplus u_2\| < \delta_2$. On the other hand, since dF_v^2 exists, there exists $\delta_3 > 0$ such that $\|u_2\| < \delta_3$ implies that

$$\|F(v_1 \oplus (v_2 + u_2)) - F(v_1 \oplus v_2) - dF_v^2(u_2)\| \leq \epsilon \|u_2\|. \tag{7.25}$$

We may as well assume that $0 < \delta_3 < \delta_2$. Then combining (7.24) and (7.25) leads to

$$\|F(v + (u_1 \oplus u_2)) - F(v) - dF_v^1(u_1) - dF_v^2(u_2)\| \leq \epsilon(\|u_1\| + \|u_2\|) = \epsilon \|u\|,$$

provided that $\|u_1 \oplus u_2\| < \delta_3$. This establishes the existence of the differential dF_v and shows that

$$dF_v(u_1 \oplus u_2) = dF_v^1(u_1) + dF_v^2(u_2). \tag{7.26}$$

The continuity of dF follows directly from (7.26) and the continuity of dF^1 and dF^2. ■

7.8 Partial Derivatives

Definition *Let $\Omega \subseteq \mathbb{R}^m$ be an open set and let $x \in \Omega$. Let $(e_1, e_2, \ldots, e_m)$ be the standard ordered basis of $\mathbb{R}^m$. Let $f : \Omega \longrightarrow W$ be a mapping into a real normed space W. Then the **partial derivative***

$$\frac{\partial f}{\partial x_j}(x) \tag{7.27}$$

is defined to be the directional derivative $D_{e_j} f(x)$ of f in the direction e_j. It is an element of W. If f is differentiable, we can restate (7.11) in the form

$$\frac{\partial f}{\partial x_j}(x) = df_x(e_j) \tag{7.28}$$

If now ξ is a vector in $\mathbb{R}^m$ we can expand in terms of the usual basis to write

$$\xi = \sum_{j=1}^{m} \xi_j e_j.$$

Combining this with (7.28) and the linearity of df_x we obtain

$$df_x(\xi) = \sum_{j=1}^{m} \xi_j \frac{\partial f}{\partial x_j}(x) \tag{7.29}$$

There is another way of thinking of (7.29). In the notation of the previous section, we can let $V = \mathbb{R}^m$ and let the V_j be the one-dimensional subspace spanned by the single vector e_j. Then of course we have

$$V = \bigoplus_{j=1}^{m} V_j.$$

It is clear that the existence of the partial derivative $\partial f/\partial x_j(x)$ is entirely equivalent to the existence of the partial differential df_x^j. The partial derivative is an element of W while the partial differential is a linear map from the one-dimensional space spanned by e_j into the space W. The two are related by

$$df_x^j(e_j) = \frac{\partial f}{\partial x_j}(x).$$

This approach allows us to state without proof the following Corollary to Theorem 142.

COROLLARY 143 *Let $\Omega \subseteq \mathbb{R}^m$ be an open set. Let W be a real normed space and $f : \Omega \longrightarrow W$ a mapping for which the partial derivatives $\partial f/\partial x_j$ exist everywhere on Ω for $1 \le j \le m$ and are continuous as functions from Ω into W. Then the differential df_x exists at each point x of Ω and (7.29) holds. Furthermore, the map $x \longrightarrow df_x$ is a continuous mapping.*

These are still not the partial derivatives of elementary calculus because they are in general vector valued. In case that W is finite dimensional we can identify W to $\mathbb{R}^k$ and write it as a direct sum of k one-dimensional subspaces. We denote by f_1 through f_k the corresponding coordinate functions. Then, the existence of all the partials $\partial f_i/\partial x_j$ as i runs over 1 to k is equivalent to the existence of the vector-valued partial derivative $\partial f/\partial x_j$. It further follows that the existence and continuity of all the partials $\partial f_i/\partial x_j$ for $(1 \le i \le k, 1 \le j \le m)$ implies the existence and continuity of the differential df_x for all $x \in \Omega$. In this case, the $k \times m$ ***Jacobian matrix***

$$\left(\frac{\partial f_i}{\partial x_j}(x)\right)_{ij}$$

is precisely the matrix representing the linear transformation df_x with respect to the usual bases in $\mathbb{R}^m$ and $\mathbb{R}^k$. Symbolically we have

$$\begin{pmatrix} f_1(x+\xi) \\ f_2(x+\xi) \\ \vdots \\ f_k(x+\xi) \end{pmatrix} = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_k(x) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_m}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \cdots & \frac{\partial f_2}{\partial x_m}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(x) & \cdots & \frac{\partial f_k}{\partial x_m}(x) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix} + \text{ error term,}$$

where the error term is little "o" of $\{\sum_{j=1}^m \xi_j^2\}^{\frac{1}{2}}$.

7.9 The Second Differential

The following result from linear algebra is well known.

PROPOSITION 144 *Let V_1, V_2 and V be vector spaces. Then there is a linear isomorphism between the space $\mathcal{L}(V_1, \mathcal{L}(V_2, V))$ of linear maps from V_1 into the space of linear maps from V_2 into V and $\mathcal{B}(V_1 \times V_2, V)$ of bilinear maps from $V_1 \times V_2$ to V. If $T \in \mathcal{L}(V_1, \mathcal{L}(V_2, V))$ the corresponding element $\theta \in \mathcal{B}(V_1 \times V_2, V)$ is defined by*

$$\theta(v_1, v_2) = (T(v_1))(v_2). \tag{7.30}$$

Effectively, the left hand side of (7.30) is linear in v_1 because T is linear and it is linear in v_2 because $T(v_1) \in \mathcal{L}(V_2, V)$ is linear. Conversely, given θ one can use (7.30) to define T.

Our next order of business is to consider the normed space version of Proposition 144. Let V_1, V_2 and V now be normed vector spaces. We will denote by $\mathcal{CL}(V_2, V)$ the normed space of continuous linear maps from V_2 to V. The norm of $\mathcal{CL}(V_2, V)$ is the operator norm. Similarly, we denote by $\mathcal{CB}(V_1 \times V_2, V)$ the space of continuous bilinear maps from $V_1 \times V_2$ to V. The proof of Theorem 40 extends easily to give the following Lemma.

LEMMA 145 *Let $\theta \in \mathcal{CB}(V_1 \times V_2, V)$. Then the quantity*

$$\|\theta\|_{\text{op}} = \sup_{\substack{\|v_1\|_{V_1} \le 1 \\ \|v_2\|_{V_2} \le 1}} \|\theta(v_1, v_2)\|_V$$

is necessarily finite.

We call $\|\theta\|_{\text{op}}$ the ***bilinear operator norm*** of θ. The following result is now routine.

PROPOSITION 146 *We can identify the spaces $\mathcal{CL}(V_1, \mathcal{CL}(V_2, V))$ and $\mathcal{CB}(V_1 \times V_2, V)$ using the correspondence of Proposition 144.*

Proof. This really boils down to

$$\begin{aligned}
\|\theta\|_{\text{op}} &= \sup_{\substack{\|v_1\|_{V_1} \le 1 \\ \|v_2\|_{V_2} \le 1}} \|\theta(v_1, v_2)\|_V \\
&= \sup_{\|v_1\|_{V_1} \le 1} \left\{ \sup_{\|v_2\|_{V_2} \le 1} \|(T(v_1))(v_2)\|_V \right\} \\
&= \sup_{\|v_1\|_{V_1} \le 1} \|T(v_1)\|_{\mathcal{CL}(V_2, V))} \\
&= \|T\|_{\mathcal{CL}(V_1, \mathcal{CL}(V_2, V))}.
\end{aligned}$$

Thus, if θ has finite norm, so does T and vice-versa. ■

Now we can approach the subject matter of this section. Let U and V be real normed vector spaces and let Ω be an open subset of U. Let $f : \Omega \longrightarrow V$ be a nice mapping and let u be a point of Ω. By the differential of f at u, if it exists, we understand an element df_u of $\mathcal{CL}(U, V)$. Let us now define a new mapping $g : \Omega \longrightarrow \mathcal{CL}(U, V)$, by $g(u) = df_u$. Clearly, g is also a mapping of an open subset of a normed vector space into a normed vector space. We can therefore contemplate the possibility that g has a differential. If this exists at $u \in \Omega$, dg_u will be a continuous linear mapping of U into $\mathcal{CL}(U, V)$, that is, an element of $\mathcal{CL}(U, \mathcal{CL}(U, V))$. By Proposition 146, we may equally well view dg_u as an element of $\mathcal{CB}(U \times U, V)$, that is as a continuous bilinear mapping of $U \times U$ to V.

Of course, dg_u is really the ***second differential*** of f, because g was already the first differential, and it is natural to use the notation $d^2 f_u$ to denote dg_u.

We can understand the second differential in terms of iterated directional derivatives.

THEOREM 147 *Let U and V be real normed vector spaces, Ω be an open subset of U and $f : \Omega \longrightarrow V$ be a continuously differentiable mapping. Let u be a point* *of Ω, and suppose that $d^2 f_u$ exists. Then, whenever $u_1, u_2 \in U$, the iterated directional derivative of f exists and we have*

$$D_{u_1}(D_{u_2} f)(u) = (d^2 f_u(u_1))(u_2).$$

Proof. The key to the proof is the ***evaluation map***. For $u \in U$ we denote by eval_u the mapping

$$\mathrm{eval}_u : \mathcal{CL}(U, V) \longrightarrow V$$

given by $\mathrm{eval}_u(T) = T(u)$. It is easy to see that eval_u is a continuous linear mapping. Maintaining the notation $g(u) = df_u$ introduced above, we now have

$$\begin{aligned} D_{u_2} f(u) &= df_u(u_2) \\ &= \mathrm{eval}_{u_2}(df_u) \\ &= \mathrm{eval}_{u_2} \circ g(u). \end{aligned}$$

Thus, by the chain rule, and since eval_{u_2} is linear we find

$$\begin{aligned} D_{u_1}(D_{u_2} f)(u) &= D_{u_1}(\mathrm{eval}_{u_2} \circ g)(u) \\ &= \mathrm{eval}_{u_2} \circ D_{u_1} g(u) \\ &= \mathrm{eval}_{u_2} \circ dg_u(u_1) \\ &= \mathrm{eval}_{u_2}(d^2 f_u(u_1)) \\ &= (d^2 f_u(u_1))(u_2), \end{aligned}$$

as required. ■

COROLLARY 148 *In the special case that $U = \mathbb{R}^m$, $x \in \Omega$ we can obtain*

$$(df_x^2(\xi))(\eta) = \sum_{i=1}^m \xi_i \frac{\partial}{\partial x_i} \left(\sum_{j=1}^m \eta_j \frac{\partial f}{\partial x_j} \right) (x)$$

or, more succinctly, viewing $d^2 f_x$ as a bilinear form

$$d^2 f_x(\xi, \eta) = \sum_{i=1}^m \sum_{j=1}^m \xi_i \eta_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

We leave the proof to the reader. It follows the same lines as that of (7.29).

We now tackle the ***symmetry of the second differential***.

THEOREM 149 *Let U and V be real normed vector spaces, Ω be an open subset of U and $f : \Omega \longrightarrow V$ be a differentiable mapping. Let u be a point of Ω, and suppose that $d^2 f_u$ exists. Then, whenever $u_1, u_2 \in U$ we have*

$$(d^2 f_u(u_1))(u_2) = (d^2 f_u(u_2))(u_1).$$

Notice that we are assuming here that $d^2 f$ exists only at the point u. There is no assumption that $d^2 f$ exists elsewhere nor any assumption that it is continuous. The proof is involved and difficult to understand. We will need the following technical Lemma which uses the same notations as in Theorem 149.

LEMMA 150 *Given $\epsilon > 0$, there exists $\delta > 0$ such that the inequality*

$$\|f(u + u_1 + u_2) - f(u + u_1) - f(u + u_2) + f(u) - d^2 f_u(u_2)(u_1)\| \\ \leq \epsilon \left(\|u_1\| + \|u_2\|\right) \|u_1\| \tag{7.31}$$

holds whenever $\|u_1\| < \delta$ and $\|u_2\| < \delta$.

It is implicit in the Lemma that δ will be chosen so small that the arguments $u + u_1 + u_2$, $u + u_1$ and $u + u_2$ are all points of Ω. Once the Lemma is established it is easy to prove the Theorem.

Proof of Theorem 149. Reverse the roles of u_1 and u_2 in (7.31) to obtain

$$\|f(u + u_1 + u_2) - f(u + u_1) - f(u + u_2) + f(u) - d^2 f_u(u_1)(u_2)\| \\ \leq \epsilon \left(\|u_1\| + \|u_2\|\right) \|u_2\| \tag{7.32}$$

Now, using (7.31), (7.32) and the triangle inequality we find

$$\|d^2 f_u(u_2)(u_1) - d^2 f_u(u_1)(u_2)\| \leq \epsilon \left(\|u_1\| + \|u_2\|\right)^2 \tag{7.33}$$

However, (7.33) is unchanged if u_1 and u_2 are simultaneously replaced by tu_1 and tu_2 where $t > 0$. Thus, by making a suitable choice of t, we can dispense with the conditions $\|u_1\| < \delta$ and $\|u_2\| < \delta$. We are now free to let ϵ tend to zero in (7.33), yielding the required conclusion. ∎

Thus everything hinges on Lemma 150.

Proof of Lemma 150. Using the definition of the differential, we can find $\delta > 0$ such that

$$\|df_w - df_u - (d^2f)_u(w-u)\| \leq \tfrac{1}{8}\epsilon\|w-u\| \tag{7.34}$$

for $\|w-u\| < 4\delta$. We will also assume that δ is chosen so small that all arguments of f occuring in this proof necessarily lie in Ω. Now, replace w by w' in (7.34) and use the triangle inequality to obtain

$$\|df_{w'} - df_w - (d^2f)_u(w'-w)\| \leq \tfrac{1}{8}\epsilon\left(\|w'-u\| + \|w-u\|\right). \tag{7.35}$$

We now replace w' by $w+z$ in (7.35) and use $\|w+z-u\| \leq \|z\| + \|w-u\|$ to obtain

$$\|df_{w+z} - df_w - (d^2f)_u(z)\| \leq \tfrac{1}{4}\epsilon\left(\|z\| + \|w-u\|\right). \tag{7.36}$$

We now fix z temporarily and introduce the auxilliary function h defined in a suitable neighbourhood of u in Ω by

$$h(w) = f(w+z) - f(w) - ((d^2f)_u(z))(w).$$

Since $w \mapsto ((d^2f)_u(z))(w)$ is linear, the differential of h is given by

$$dh_w = df_{w+z} - df_w - (d^2f)_u(z)$$

which, when combined with (7.36) yields

$$\|dh_w\|_{\mathrm{op}} \leq \tfrac{1}{4}\epsilon\left(\|z\| + \|w-u\|\right) \tag{7.37}$$

for $\|w-u\| < 2\delta$ and $\|z\| < 2\delta$. Since the set $\{w; \|w-u\| < 2\delta\}$ is a convex subset of U, we may apply Corollary 139 to obtain the Lipschitz estimate

$$\|h(w_1) - h(w_2)\| \leq \tfrac{1}{4}\epsilon\left(\|z\| + \max_{\tilde{w}\in L(w_1,w_2)}\left(\|\tilde{w}-u\|\right)\right)\|w_1-w_2\|.$$

But a moment's thought convinces us that the maximum norm is taken at one or other end of the line segment (this is just convexity of the norm). Hence we may write

$$\begin{aligned}\|h(w_1) - h(w_2)\|& \\ &\leq \tfrac{1}{4}\epsilon\left(\|z\| + \max\left(\|w_1-u\|, \|w_2-u\|\right)\right)\|w_1-w_2\|. \end{aligned}\tag{7.38}$$

Substituting back the definition of h into (7.38) we find

$$\|f(w_1+z)-f(w_1)-f(w_2+z)+f(w_2)-(d^2f_u(z))(w_1-w_2)\| \\ \leq \epsilon C(z,w_1,w_2)\|w_1-w_2\|$$

for $\|w_1-u\|<2\delta$, $\|w_2-u\|<2\delta$ and $\|z\|<2\delta$ and where $C(z,w_1,w_2)$ denotes

$$\tfrac{1}{4}\left(\|z\|+\max\left(\|w_1-u\|,\|w_2-u\|\right)\right).$$

We now make the following substitutions

$$\begin{aligned} w_1 &= u+u_1+u_2 \\ w_2 &= u+u_2 \\ z &= -u_2 \end{aligned}$$

to obtain

$$\|f(u+u_1)-f(u+u_1+u_2)-f(u)+f(u+u_2)-(d^2f_u(-u_2))(u_1)\| \\ \leq \epsilon C(z,w_1,w_2)\|u_1\|$$

which leads to (7.31), since

$$\begin{aligned} C(z,w_1,w_2) &= \tfrac{1}{4}\Big(\|-u_2\|+\max(\|u_1+u_2\|,\|u_2\|)\Big) \\ &\leq \tfrac{1}{4}\Big(\|u_2\|+\|u_1\|+\|u_2\|\Big) \\ &\leq \|u_1\|+\|u_2\|. \end{aligned}$$

■

In linear algebra, there are no canonical forms for general bilinear mappings. However, once one imposes symmetry of the bilinear mapping, the situation becomes very well understood and a canonical form is available. Theorem 149 therefore allows one the possibility of understanding the second differential at a point from a qualitative point of view — it makes sense to say that it is positive definite, or negative definite or indefinite.

In the finite dimensional case, the symmetry of the second derivative can be understood purely on the level of partial derivatives. If the second differential of a function f exists, then its representing matrix is given in terms of second order partial derivatives by

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m}(x) \end{pmatrix}$$

and is called the ***Hessian matrix*** of f at x.

THEOREM 151 *Let Ω be an open subset of $\mathbb{R}^m$. Let $f : \Omega \longrightarrow \mathbb{R}$. Let the partial derivatives*

$$\frac{\partial f}{\partial x_j} \text{ and } \frac{\partial^2 f}{\partial x_j \partial x_k}$$

all exist in Ω and suppose that the second order derivatives are continuous in Ω. Then

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}$$

in Ω.

Proof. This result can actually be obtained as a corollary of Theorem 149, but it is more instructive to give an independent proof, much simpler than that of Theorem 149. For simplicity, we will assume that $m = 2$ and use the notation $x = x_1, y = x_2$.

For k and h sufficiently small we have

$$\begin{aligned} f(x+k, y+h) - f(x+k, y) - f(x, y+h) + f(x, y) \\ = k\left(\frac{\partial f}{\partial x}(\xi, y+h) - \frac{\partial f}{\partial x}(\xi, y)\right) \end{aligned} \tag{7.39}$$

from applying the Mean Value Theorem (page 152) to the function g given by

$$g(t) = f(t, y+h) - f(t, y).$$

The point ξ lies between x and $x+k$. Applying the Mean Value Theorem a second time gives

$$f(x+k, y+h) - f(x+k, y) - f(x, y+h) + f(x, y) = kh\left(\frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)\right).$$

where η lies between y and $y+h$. An exactly similar argument yields

$$f(x+k,y+h)-f(x+k,y)-f(x,y+h)+f(x,y)=kh(\frac{\partial^2 f}{\partial x\partial y}(\xi_1,\eta_1)),$$

where ξ_1 lies between x and $x+k$ and η_1 lies between y and $y+h$. For $kh\neq 0$ we now get

$$\frac{\partial^2 f}{\partial x\partial y}(\xi_1,\eta_1)=\frac{\partial^2 f}{\partial y\partial x}(\xi,\eta).$$

Using the continuity of both second partials at (x,y), it suffices to let k and h tend to zero to conclude that

$$\frac{\partial^2 f}{\partial x\partial y}(x,y)=\frac{\partial^2 f}{\partial y\partial x}(x,y).$$

■

This entire theory extends in an obvious way to the case in which the function f is vector valued and also to the third and higher order derivatives.

7.10 Local Extrema

The concepts of local maximum and local minimum point were already defined on page 151. The situation was discussed for differentiable functions of a single variable. The situation extends naturally to the case where the function is defined on an open subset of a normed vector space. .

THEOREM 152 *Let $\Omega\subseteq U$ be an open set of a normed real vector space U. Let $f:\Omega\longrightarrow\mathbb{R}$ be a differentiable function. Suppose also that $u\in\Omega$ is a local minimum point of f. Then*

- $df_u=0$.
- *If it exists, d^2f_u is a positive semidefinite bilinear form. Explicitly, this means that*

$$d^2f_u(w,w)\geq 0\qquad\forall w\in U$$

Proof. The problem is reduced to the scalar case. Let $\varphi(t)=f(u+tw)$ where t is real and w is a fixed vector in U. Since $u\in\Omega$ open, φ is defined in some interval $]-a, a[$ with $a > 0$. Also it is clear that 0 is a local minimum point for φ hence, according to Lemma 132 (page 151) we find that $\varphi'(0) = 0$. The Chain Rule (Theorem 131, page 149) now shows that $\varphi'(0) = df_u(w)$ so that

$$df_u(w) = 0 \qquad \forall w \in U.$$

It follows that $df_u = 0$.

For the second part of the proof, we see from Theorem 147 that $\varphi''(0)$ exists. Let us suppose that $\varphi''(0) < 0$ then we will produce a contradiction. Let $\epsilon = -\frac{1}{2}\varphi''(0) > 0$ then because $\varphi''(0)$ exists, we can find $\delta > 0$ such that

$$\left| \frac{\varphi'(t) - \varphi'(0)}{t} - \varphi''(0) \right| < \epsilon \tag{7.40}$$

whenever $|t| < \delta$. Since $\varphi'(0) = 0$ we can rewrite (7.40) in the form

$$\frac{\varphi'(t)}{t} < -\tfrac{1}{2}\epsilon.$$

This shows that $\varphi'(t) < 0$ if $0 < t < \delta$ and that $\varphi'(t) > 0$ if $-\delta < t < 0$. On the other hand, since 0 is a local minimum point for φ we can find $\delta_1 > 0$ such that

$$\varphi(t) \geq \varphi(0)$$

for $|t| < \delta_1$. But an application of the Mean Value Theorem now gives say for $s = \frac{1}{2}\min(\delta, \delta_1)$ that

$$0 \leq \varphi(s) - \varphi(0) = s\varphi'(t) < 0,$$

since $0 < t < s$. This contradiction shows that $\varphi''(0) \geq 0$. Finally, we have $d^2 f_u(w, w) = \varphi''(0)$ by Theorem 147. ■

In the opposite direction we have the following result.

PROPOSITION 153 *Let $\Omega \subseteq U$ be an open set of a normed real vector space U and that $f : \Omega \longrightarrow \mathbb{R}$ be a twice continuously differentiable function. Suppose also that $df_u = 0$ and that $d^2 f_u$ is positive definite. Explicitly, this last statement means that there exists $\epsilon > 0$ such that*

$$d^2 f_u(w, w) \geq \epsilon \|w\|^2 \qquad \forall w \in U$$

Then u is a local minimum point for f.

In the case that U is finite dimensional, the positive definiteness condition given above is just the usual positive definiteness of bilinear forms. The proof of Proposition 153 is put off until page 184.

8

Integrals and Derivatives

In order to understand differential calculus better, we need to have some knowledge of the integral. We describe here a very simplistic version of the Riemann Integral. The treatment given here is rather sketchy, because we are not particularly interested in developing integration for its own sake in this text, but rather because it can be used to clarify some of the issues arising in differential calculus.

8.1 A Riemann type integration theory

Let V be a complete normed vector space, typically $\mathbb{R}^n$ with any given norm. Let $[a, b]$ be any closed bounded interval in $\mathbb{R}$. Let $f : [a, b] \longrightarrow V$ be a continuous mapping. We show how to define the integral of f as an element of V.

All integration theories need to encompass a notion of length or measure of sets. In the Lebesgue theory this task is accomplished for certain subsets the real line. Once successfully completed, the corresponding integration theory is really very easy. The difficulties associated with Lebesgue integration are with the construction of the concept of length for as general a class of subsets of $\mathbb{R}$ as possible. In what we do here we avoid these difficulties by considering only the length of bounded intervals of $\mathbb{R}$. For J a bounded interval, we denote length(J) its length. If $t \geq s$ and J is one of the four intervals $]s, t[$, $[s, t[$, $]s, t]$ and $[s, t]$, then we define length$(j) = t - s$. We also define length$(\emptyset) = 0$.

Towards this we need the following combinatorial Lemma which we leave as an exercise.

LEMMA 154 *Let I be a bounded interval be written as the disjoint union*

$$\dot{\bigcup_{j=1}^{m}} J_j = I,$$

of intervals J_j. Then

$$\sum_{j=1}^{m} \operatorname{length}(J_j) = \operatorname{length}(I).$$

DEFINITION *Let I be a bounded interval. Then a* ***Riemann partition*** *$\mathcal{P}$ of I is a finite family of disjoint intervals J_j, $j = 1, \ldots, m$ whose union is I, together with* ***representative points*** *x_j, $j = 1, \ldots, m$ with $x_j \in J_j$, $j = 1, \ldots, m$. The* ***step*** *of such a Riemann partition is defined by*

$$\operatorname{step}(\mathcal{P}) = \max_{j=1}^{m} \operatorname{length}(J_j)$$

DEFINITION *Let I be a closed bounded interval and f be a continuous mapping $f : I \longrightarrow V$ where V is a complete normed space. Let $\mathcal{P}$ be a Riemann partition of I, then the corresponding* ***Riemann sum*** *is defined by*

$$S(\mathcal{P}, f) = \sum_{j=1}^{m} \operatorname{length}(J_j) f(x_j).$$

Here, the lengths $\operatorname{length}(J_j)$ *are scalars, while $f(x_j)$ is an element of V. Thus the Riemann sum $S(\mathcal{P}, f)$ is an element of V.*

We need a Lemma which relates two Riemann sums.

LEMMA 155 *Let $\mathcal{P} = ((J_j)_{j=1}^m, (x_j)_{j=1}^m)$ and $\mathcal{Q} = ((K_k)_{k=1}^n, (y_k)_{k=1}^n)$ be two partitions of the closed bounded interval I with* $\operatorname{step}(\mathcal{P}) < \delta_P$ *and* $\operatorname{step}(\mathcal{Q}) < \delta_Q$. *Let f be a continuous mapping $f : I \longrightarrow V$ where V is a complete normed space. Then*

$$\|S(\mathcal{P}, f) - S(\mathcal{Q}, f)\| \le \omega_f(\delta_P + \delta_Q) \operatorname{length}(I).$$

Proof. By Lemma 154, we have

$$\operatorname{length}(J_j) = \sum_{k=1}^{n} \operatorname{length}(J_j \cap K_k),$$

which allows us to write

$$S(\mathcal{P}, f) = \sum_{j=1}^{m} \sum_{k=1}^{n} f(x_j) \operatorname{length}(J_j \cap K_k).$$

Using a similar decomposition of K_k we have

$$S(\mathcal{Q}, f) = \sum_{j=1}^{m} \sum_{k=1}^{n} f(y_k) \operatorname{length}(J_j \cap K_k).$$

Subtracting off we get

$$S(\mathcal{P}, f) - S(\mathcal{Q}, f) = \sum_{j=1}^{m} \sum_{k=1}^{n} (f(x_j) - f(y_k)) \operatorname{length}(J_j \cap K_k). \tag{8.1}$$

The terms on the right are handled in one of two ways. If $J_j \cap K_k = \emptyset$ then $\operatorname{length}(J_j \cap K_k) = 0$. On the other hand, if $J_j \cap K_k \neq \emptyset$, then we can find $z_{jk} \in J_j \cap K_k$. Since $\operatorname{step}(\mathcal{P}) < \delta_P$, we have $|x_j - z_{jk}| < \delta_P$. Similarly, $|y_k - z_{jk}| < \delta_Q$ from which it follows that $|x_j - y_k| < \delta_P + \delta_Q$. Hence $\|f(x_j) - f(y_k)\| \le \omega_f(\delta_P + \delta_Q)$. Using these estimates in (8.1) we get

$$\begin{aligned} \|S(\mathcal{P}, f) - S(\mathcal{Q}, f)\| &\le \omega_f(\delta_P + \delta_Q) \sum_{j=1}^{m} \sum_{k=1}^{n} \operatorname{length}(J_j \cap K_k) \\ &= \omega_f(\delta_P + \delta_Q) \operatorname{length}(I), \end{aligned} \tag{8.2}$$

by a further application of Lemma 154. ■

We are now ready to establish the existence of the integral.

THEOREM 156 *Let I be a closed bounded interval and V a complete normed space. Let $f : I \longrightarrow V$ be a continuous mapping, and suppose that $(\mathcal{P}_k)$ is a sequence of Riemann partitions of I such that*

$$\operatorname{step}(\mathcal{P}_k) \longrightarrow 0$$

as $n \to \infty$. Then $(S(\mathcal{P}_k, f))$ is a convergent sequence in V. The limit is denoted

$$\int_I f(x)dx$$

and is independent of the sequence $(\mathcal{P}_k)$ of partitions used. Furthermore, for any partition $\mathcal{P}$ of step less than δ_P we have

$$\| \int_I f(x)dx - S(\mathcal{P}, f)\| \le \omega_f(\delta_P)\,\text{length}(I). \tag{8.3}$$

Proof. By Theorem 78, I is compact. By Theorem 85, f is uniformly continuous. It now follows from Lemma 155 that $(S(\mathcal{P}_k, f))$ is a Cauchy sequence in V. By the completeness of V we see that $(S(\mathcal{P}_k, f))$ converges.

Let $(\mathcal{Q}_k)$ be another sequence of partitions with the same property. Then we construct a further sequence $(\mathcal{R}_k)$ by intermingling the two.

$$\mathcal{R}_k = \begin{cases} \mathcal{P}_\ell & \text{if } k \text{ is odd and } k = 2\ell - 1, \\ \mathcal{Q}_\ell & \text{if } k \text{ is even and } k = 2\ell. \end{cases}$$

Repeating the above argument shows that $(S(\mathcal{R}_k, f))$ converges in V. It contains the sequences $(S(\mathcal{P}_k, f))$ and $(S(\mathcal{Q}_k, f))$ as subsequences. This establishes the uniqueness of the limit.

Finally let $(\mathcal{Q}_k)$ be any sequence with step$(\mathcal{Q}_k)$ tending to zero. Replacing $\mathcal{Q}$ with $\mathcal{Q}_k$ in (8.2) and letting k tend to ∞ we find (8.3). ∎

Next we establish elementary properties of the integral. In the remainder of this section I denotes a closed bounded interval of $\mathbb{R}$ and V, a complete normed vector space.

LEMMA 157 *The integral is a linear mapping from $C(I, V)$ to V.*

Proof. Let f_1 and f_2 be continuous functions from the closed bounded interval I to V. Let t_1 and t_2 be real numbers. We must show that

$$\int_I (t_1 f_1(x) + t_2 f_2(x))dx = t_1 \int_I f_1(x)dx + t_2 \int_I f_2(x)dx. \tag{8.4}$$

It is easy to check that for every Riemann partition $\mathcal{P}$ we have

$$S(\mathcal{P}, t_1 f_1 + t_2 f_2) = t_1 S(\mathcal{P}, f_1) + t_2 S(\mathcal{P}, f_2). \tag{8.5}$$

Now choose a sequence of partitions $(\mathcal{P}_k)$ with steps tending to zero and substitute $\mathcal{P} = \mathcal{P}_k$ in (8.5). The result of letting k tend to infinity in the resulting equality is precisely (8.4). ∎

LEMMA 158

- *Let* $f : I \longrightarrow \mathbb{R}^+$ *be continuous mapping. Then*

$$\int_I f(x)dx \geq 0.$$

- *Let* $f : I \longrightarrow \mathbb{R}$ *be continuous mapping. Then*

$$\inf_{x \in I} f(x) \operatorname{length}(I) \leq \int_I f(x)dx \leq \operatorname{length}(I) \sup_{x \in I} f(x).$$

- *Let* $f : I \longrightarrow V$ *be a continuous mapping. Then*

$$\| \int_I f(x)dx \| \leq \operatorname{length}(I) \sup_{x \in I} \|f(x)\|.$$

- *Let* $f : I \longrightarrow V$ *be the constant mapping given by* $f(x) = v$ *for all* $x \in I$*. Then*

$$\int_I f(x)dx = \operatorname{length}(I) v.$$

We leave the proof of the Lemma to the reader because it uses very similar arguments to those used above.

8.2 Properties of integrals

The following property of the integral is essential to the theory.

LEMMA 159 *Let* $a < b < c$ *be real numbers. Let* $f : [a, c] \longrightarrow V$ *be a continuous mapping into a complete normed vector space* V*. Then*

$$\int_{[a,c]} f(x)dx = \int_{[a,b]} f(x)dx + \int_{[b,c]} f(x)dx.$$

Proof. Form Riemann partitions $\mathcal{Q}_n$ of $[a, b]$ and $\mathcal{R}_n$ of $[b, c]$ by splitting these intervals into n equal subintervals. Let the representative points be the mid-points of the subintervals. The union of these collections of intervals and representative points forms a Riemann partition $\mathcal{P}_n$ of $[a, c]$ after a little fudging to ensure that the point b lies in only one of the subintervals. It follows easily that

$$S(\mathcal{P}_n, f) = S(\mathcal{Q}_n, f|_{[a,b]}) + S(\mathcal{R}_n, f|_{[b,c]}).$$

Passing to the limit as $n \to \infty$ yields the desired result. ■

The integrals that we have defined are integrals taken over sets — in our case, closed bounded intervals. This is the way that things are done in the Lebesgue Theory for example. From the point of view of analysis, it is actually the natural way to start. However calculus students are first introduced to integrals over directed intervals. These are easily defined by

$$\int_a^b f(x)dx = \begin{cases} \int_{[a,b]} f(x)dx & \text{if } a < b, \\ -\int_{[b,a]} f(x)dx & \text{if } a > b, \\ 0 & \text{if } a = b. \end{cases} \tag{8.6}$$

Interchanging the a and b on the left-hand side of (8.6) changes the sign of the integral.

This is the point of view that seems natural to differential geometers. It ultimately leads to an integration theory on oriented differentiable manifolds. The two theories may seem to be the same, but they lead in different directions. The ultimate result of reconciling the two theories is the Poincaré Duality Theorem which concerns the concept of topological degree. These matters are of course well beyond the scope of this text.

DEFINITION *Suppose that $a < b$ and that $c \in [a, b]$ is some basepoint. Let V be a complete normed vector space and let $f : [a, b] \to V$ be a continuous mapping. The* ***primitive*** *or* ***indefinite integral*** *of f is defined by*

$$g(x) = \int_c^x f(t)dt.$$

It is a function $g : [a, b] \to V$.

THEOREM 160 (FUNDAMENTAL THEOREM OF CALCULUS) *With the same notations and conditions as above, we have that $g'(x)$ exists for every point of $]a, b[$ and $g'(x) = f(x)$.*

Proof. Since f is continuous and by Theorem 85, we see that f is uniformly continuous. Now by Lemma 159, we have for $x, x+h \in [a,b]$,

$$g(x+h) - g(x) = \int_x^{x+h} f(t)dt.$$

Applying now (8.3) in the case where the Riemann partition consists of a single interval, and where the corresponding representative point is x we get

$$\|g(x+h) - g(x) - hf(x)\| \leq |h|\omega_f(|h|).$$

We now find that

$$\|\frac{g(x+h) - g(x)}{h} - f(x)\| \leq \omega_f(|h|) \longrightarrow 0,$$

as $h \longrightarrow 0$ by the uniform continuity of f. Thus the derivative $g'(x)$ exists and equals $f(x)$. ■

The Fundamental Theorem of Calculus gives us the ability to make substitutions in integrals.

THEOREM 161 (CHANGE OF VARIABLES THEOREM) *Let $\varphi :]a,b[\longrightarrow]\alpha,\beta[$ be a differentiable mapping with continuous derivative. Let $c \in]a,b[$ and $\gamma \in]\alpha,\beta[$ be basepoints such that $\varphi(c) = \gamma$. Let $f :]\alpha,\beta[\longrightarrow \mathbb{R}$ be a continuous mapping. Then for $u \in]a,b[$, we have*

$$\int_\gamma^{\varphi(u)} f(t)dt = \int_c^u f(\varphi(s))\varphi'(s)ds.$$

Proof. We define for $v \in]\alpha,\beta[$,

$$g(v) = \int_\gamma^v f(t)dt.$$

Then according to the Fundamental Theorem of Calculus, g is differentiable on $]\alpha,\beta[$ and

$$g'(v) = f(v).$$

Then, by the Chain Rule (page 149), stated for derivatives rather than differentials, for $u \in]a,b[$ we have

$$(g \circ \varphi)'(u) = (g' \circ \varphi)(u)\varphi'(u) = (f \circ \varphi)(u)\varphi'(u).$$

Since

$$u \longrightarrow (f \circ \varphi)(u)\varphi'(u)$$

is a continuous mapping, the Fundamental Theorem of Calculus can be applied again to show that if $h :]a, b[\longrightarrow \mathbb{R}$ is defined by

$$h(u) = \int_c^u f(\varphi(s))\varphi'(s)ds,$$

then $h'(u) = (f \circ \varphi)(u)\varphi'(u) = (g \circ \varphi)'(u)$. An application of Corollary 138 now shows that $h(u) - g(\varphi(u))$ is constant. Substituting $u = c$ shows that the constant is zero. Hence $h(u) = g(\varphi(u))$ for all $u \in]a, b[$. This is exactly what was to be proved. ■

The second objective of this section is to be able to differentiate under the integral sign.

THEOREM 162 *Let $\alpha < \beta$ and $a < b$. Let V be a complete normed vector space. Suppose that*

$$f, g : [a, b] \times [\alpha, \beta] \longrightarrow V$$

are continuous mappings such that $\dfrac{\partial g}{\partial t}(t, s)$ exists and equals $f(t, s)$ for all $(t, s) \in]a, b[\times [\alpha, \beta]$. Let us define a new function $G : [a, b] \to V$ by

$$G(t) = \int_{[\alpha,\beta]} g(t, s)ds \qquad (a \le t \le b).$$

Then $G'(t)$ exists for $a < t < b$ and

$$G'(t) = \int_{[\alpha,\beta]} f(t, s)ds \qquad (a < t < b).$$

Proof. For shortness of notation, let us define

$$F(t) = \int_{[\alpha,\beta]} f(t, s)ds \qquad (a < t < b).$$

Then, we have for $a < x < b$ and small enough h that

$$\begin{aligned} G(x+h) - G(x) - hF(x) &= \int_{[\alpha,\beta]} (g(x+h,s) - g(x,s) - hf(x,s))ds \\ &= \int_{[\alpha,\beta]} \left\{ \int_x^{x+h} (f(u,s) - f(x,s))du \right\} ds \end{aligned}$$

where we have used the Fundamental Theorem of Calculus in the last step. Sine the points (u,s) and (x,s) are separated by a distance of at most $|h|$, the inner integral satisfies

$$\left\| \int_x^{x+h} (f(u,s) - f(x,s))du \right\| \le |h|\omega_f(|h|).$$

It follows that

$$\|G(x+h) - G(x) - hF(x)\| \le (\beta - \alpha)|h|\omega_f(|h|).$$

Since f is a continuous function on the compact space $[a,b] \times [\alpha,\beta]$, it follows that f is uniformly continuous and this gives the desired conclusion. ■

EXAMPLE Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ be an infinitely differentiable mapping. Suppose that we now define a new mapping $\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$\theta(t,s) = \begin{cases} \dfrac{\varphi(t) - \varphi(s)}{t-s} & \text{if } t \ne s, \\ \varphi'(s) & \text{if } t = s. \end{cases}$$

We claim that θ is also infinitely differentiable. We approach this using Corollary 143, which allows us to assert that it is enough to show that the partial derivatives of θ of all orders exist and are continuous. This is clear except on the diagonal set where $t = s$.

The problem is solved very neatly using an integral representation of θ, namely

$$\theta(t,s) = \int_0^1 \varphi'((1-u)t + us)du,$$

agreeing with the previous definition by the Fundamental Theorem of Calculus and the Change of Variables Theorem. Repeated applications of Theorem 162 allow us to prove that

$$\frac{\partial^{\alpha+\beta}\theta}{\partial t^\alpha \partial s^\beta}(t,s) = \int_0^1 (1-u)^\alpha u^\beta \varphi^{(\alpha+\beta+1)}((1-u)t + us)du,$$

for all integers $\alpha, \beta \geq 0$. The partial derivatives of θ are easily seen to be continuous, using the fact that $\varphi^{(\alpha+\beta+1)}$ is uniformly continuous on the bounded intervals of $\mathbb{R}$. □

8.3 Taylor's Theorem

Before we can tackle Taylor's Theorem, we need to extend the Mean-Value Theorem.

THEOREM 163 (EXTENDED MEAN VALUE THEOREM) *Suppose that a and b are real numbers such that $a < b$. Let $g, h : [a, b] \longrightarrow \mathbb{R}$ be continuous maps. Suppose that g and h are differentiable at every point of $]a, b[$. Then there exists ξ such that $a < \xi < b$ and*

$$(g(b) - g(a))h'(\xi) = g'(\xi)(h(b) - h(a)).$$

Proof. Let us define

$$f(x) = g(x)(h(b) - h(a)) - (g(b) - g(a))h(x).$$

Then routine calculations show that

$$f(a) = g(a)h(b) - g(b)h(a) = f(b).$$

Since f is continuous on $[a, b]$ and differentiable on $]a, b[$, we can apply Rolle's Theorem (page 152) to establish the existence of $\xi \in]a, b[$ such that $f'(\xi) = 0$, a statement equivalent to the desired conclusion. ■

DEFINITION *Let f be a function $f :]a, b[\longrightarrow \mathbb{R}$ which is n times differentiable. Formally this means that the successive derivatives f', f'',...,$f^{(n)}$ exist on $]a, b[$. Let $c \in]a, b[$ be a basepoint. Then we can construct the **Taylor Polynomial** $T_{n,c}f$ of order n at c by*

$$T_{n,c}f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(c)(x - c)^k.$$

To make the notations clear we point out that $f^{(0)} = f$, that $0! = 1$ and that $(x - c)^0 = 1$. In fact even $0^0 = 1$ because it is viewed as an "empty product".

THEOREM 164 (TAYLOR'S THEOREM) *Let f be a function $f :]a,b[\longrightarrow \mathbb{R}$ which is $n+1$ times differentiable. Let $c \in]a,b[$ be a basepoint. Then there exists a point ξ between c and x, such that*

$$f(x) = T_{n,c}f(x) + \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-c)^{n+1}. \tag{8.7}$$

The statement ξ is between c and x means that

$$\begin{cases} c < \xi < x & \text{if } c < x, \\ c = \xi = x & \text{if } c = x, \\ x < \xi < c & \text{if } c > x. \end{cases}$$

The second term on the right of (8.7) is called the ***remainder term*** and in fact this specific form of the remainder is called the ***Lagrange remainder*** . It is the most common form. When we look at (8.7), we think of writing the function f as a polynomial plus an error term (the remainder). Of course, there is no guarantee that the remainder term is small.

All this presupposes that f is a function of x and indeed this is the obvious point of view when we are applying Taylor's Theorem. However for the proof, we take the other point of view and regard x as the constant and c as the variable.

Proof. First of all, if $x = c$ there is nothing to prove. We can therefore assume that $x \neq c$. We regard x as fixed and let c vary in $]a,b[$. We define

$$g(c) = f(x) - T_{n,c}f(x) \qquad \text{and} \qquad h(c) = (x-c)^{n+1}.$$

On differentiating g with respect to c we obtain a telescoping sum which yields

$$g'(\xi) = -\frac{1}{n!} f^{(n+1)}(\xi)(x-\xi)^n. \tag{8.8}$$

On the other hand we have, differentiating h with respect to c,

$$h'(\xi) = -(n+1)(x-\xi)^n.$$

Applying now the extended Mean-Value Theorem, we obtain

$$(g(c) - g(x))h'(\xi) = g'(\xi)(h(c) - h(x)),$$

where ξ is between c and x. Since both $g(x) = 0$ and $h(x) = 0$ (remember g and h are viewed as functions of c, so here we are substituting $c = x$), this is equivalent to

$$(f(x) - T_{n,c}f(x))(-(n+1)(x-\xi)^n) = (-\frac{1}{n!} f^{(n+1)}(\xi)(x-\xi)^n)(x-c)^{n+1},$$

Since $x \neq c$, we have that $x \neq \xi$ and we may divide by $(x - \xi)^n$ and obtain the desired conclusion. ■

In many situations, we can use estimates on the Lagrange remainder to establish the validity of power series expansion

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(c)(x - c)^k,$$

for x in some open interval around c. Such estimates are sometimes fraught with difficulties because all that one knows about ξ is that it lies between c and x. One has to face the fact that some information has been lost and allow for all such possible ξ. Usually, there are better ways of establishing the validity of power series expansions. There remains the question of obtaining estimates of the Taylor remainder in which no information is sacrificed.

THEOREM 165 (INTEGRAL REMAINDER THEOREM) *Let f be a function $f :]a, b[\longrightarrow \mathbb{R}$ which is $n+1$ times differentiable and such that $f^{(n+1)}$ is continuous. Let $c \in]a, b[$ be a basepoint. Then we have for $x \in]a, b[$*

$$f(x) = T_{n,c}f(x) + \frac{1}{n!} \int_{\xi=c}^{x} f^{(n+1)}(\xi)\,(x - \xi)^n d\xi, \tag{8.9}$$

or equivalently by change of variables

$$f(x) = T_{n,c}f(x) + \frac{1}{n!}(x - c)^{n+1} \int_{t=0}^{1} f^{(n+1)}((1 - t)c + tx)\,(1 - t)^n dt \tag{8.10}$$

This Theorem provides an explicit formula for the remainder term which involves an integral. Note that in order to define the integral it is supposed that f is slightly more regular than is the case with the Lagrange form of the remainder.

Proof. Again we tackle (8.9) by viewing x as the constant and c as the variable. Equation (8.9) follows immediately from the Fundamental Theorem of Calculus and (8.8). The second formulation (8.10) follows by the Change of Variables Theorem, using the substitution $\xi = (1 - t)c + tx$. ■

EXAMPLE Let $\alpha > 0$ and consider

$$f(x) = (1 - x)^{-\alpha}$$

for $-1 < x < 1$. The Taylor series of this function is

$$f(x) = 1 + \alpha x + \frac{\alpha(\alpha+1)}{2!}x^2 + \dots$$

actually valid for $-1 < x < 1$. If we try to obtain this result using the Lagrange form of the remainder

$$\frac{\alpha(\alpha+1)\dots(\alpha+n)}{(n+1)!}(1-\xi)^{-\alpha-n-1}x^{n+1}$$

we are able to show that the remainder tends to zero as n tends to infinity provided that

$$\sup \left| \frac{x}{1-\xi} \right| < 1,$$

where the sup is taken over all ξ between 0 and x. If $x > 0$ the worst case is when ξ is very close to x. Convergence of the Lagrange remainder to zero is guaranteed only if $0 < x < \frac{1}{2}$. On the other hand, if $x < 0$ then the worst location of ξ is $\xi = 0$. Convergence of the Lagrange remainder is then guaranteed for $-1 < x < 0$. Combining the two cases, we see that the Lagrange remainder can be controlled only for $-1 < x < \frac{1}{2}$.

For the same function, the integral form of the remainder is

$$\frac{\alpha(\alpha+1)\dots(\alpha+n)}{n!}\int_0^x (1-\xi)^{-\alpha-n-1}(x-\xi)^n d\xi.$$

For ξ between 0 and x we have

$$\left| \frac{x-\xi}{1-\xi} \right| \le |x|,$$

for $-1 < x < 1$. This estimate allows us to show that the remainder tends to zero over the full range $-1 < x < 1$. □

We can now settle some business postponed from the last chapter.

Proof of Proposition 153. In case $n = 1$, (8.9) gives

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \int_0^t (t-s)\varphi''(s)ds$$

assuming that φ is a twice continuously differentiable real-valued function defined on some neighbourhood of 0. Taking now

$$\varphi(t) = f(u + tw)$$

we find that

$$f(u + tw) = f(u) + tdf_u(w) + \int_0^t d^2 f_{u+sw}(w, w)(t - s)ds$$

which we can rewrite using the hypothesis $df_u = 0$ as

$$\begin{aligned} f(u + tw) &= f(u) + \tfrac{1}{2}t^2 d^2 f_u(w, w) + \int_0^t (d^2 f_{u+sw} - d^2 f_u)(w, w)(t - s)ds \\ &\geq f(u) + \tfrac{1}{2}t^2 \epsilon \|w\|^2 + \int_0^t (d^2 f_{u+sw} - d^2 f_u)(w, w)(t - s)ds. \end{aligned}$$

If $\|w\|$ is small enough, we have, using the continuity of $d^2 f$ that

$$|(d^2 f_{u+sw} - d^2 f_u)(w, w)| \leq \tfrac{1}{2}\epsilon \|w\|^2$$

and it follows that

$$f(u + tw) \geq f(u) + \tfrac{1}{4}t^2 \epsilon \|w\|^2$$

for such w. Hence u is a local minimum point of f. ■

8.4 Derivatives and Uniform Convergence

The results developed in this section are primarily intended for use in the theory of power series. The following Lemma is elementary.

LEMMA 166 *Let g_n and g be continuous functions on the bounded interval $[\alpha, \beta]$, taking values in a metric space. Suppose that $g_n \longrightarrow g$ uniformly on $[\alpha, \beta]$. Let $\xi, \xi_n \in [\alpha, \beta]$ be a sequence of points such that $\xi_n \longrightarrow \xi$. Then $g_n(\xi_n) \longrightarrow g(\xi)$.*

This Lemma will now be used to establish the key Theorem needed to start work on power series.

THEOREM 167 *Let V and W be normed spaces. Let $\Omega \subseteq V$ be an open set. Suppose that f_n are functions on Ω with values in W with continuous differentials df_n. Further suppose that f_n converges pointwise to a function f on Ω and that df_n converges uniformly on Ω to a function g on Ω with values in $\mathcal{CL}(V, W)$. Then f is continuously differentiable on Ω and $df = g$.*

Proof. Since the result is essentially local in nature, we can assume that Ω is a convex open set. Let $v \in \Omega$. Then for $u \in V$ of small enough norm, the line segment joining v to $v + u$ lies inside Ω. Applying Corollary 139 to the function

$$\varphi(u) = f_n(v + u) - d(f_n)_v(u)$$

we obtain

$$\|f_n(v + u) - f_n(v) - d(f_n)_v(u)\| \leq \|u\| \sup_{0 \leq t \leq 1} \|d(f_n)_{v+tu} - d(f_n)_v\|.$$

In particular, we can find $t_n \in [0, 1]$ such that

$$\|f_n(v + u) - f_n(v) - d(f_n)_v(u)\| \leq \|u\|(\|u\| + \|d(f_n)_{v+t_n u} - d(f_n)_v\|).$$

Some subsequence (t_{n_k}) of (t_n) converges to an element $t \in [0, 1]$. Of course, this subsequence and its limit t depends on u and v. Passing to the limit along this subsequence, using the hypotheses and Lemma 166, we find that

$$\|f(v + u) - f(v) - (g(v))(u)\| \leq \|u\|(\|u\| + \|g(v + tu) - g(v)\|).$$

Here, t depends on u and v. To be clear, we had better write

$$\|f(v + u) - f(v) - (g(v))(u)\| \leq \|u\|(\|u\| + \sup_{0 \leq t \leq 1} \|g(v + tu) - g(v)\|) \quad (8.11)$$

But g is continuous, since it is a uniform limit of continuous functions. It follows that the right hand side of (8.11) is little "o" of $\|u\|$. It follows that df_v exists and equals $g(v)$. ■

EXAMPLE We have already verified the Binomial Theorem for negative powers using Taylor's Theorem with the integral form of the remainder. Let us now approach the same question using power series. Let $\alpha > 0$ and define $c_0 = 1$, $c_1 = \alpha$, $c_2 = \frac{\alpha(\alpha+1)}{2!}$, and so forth. Then define polynomials f_n by

$$f_n(x) = \sum_{k=0}^{n} c_k x^k.$$

Let $0 < |x| < r < 1$. Then since

$$\sum_{k=0}^{\infty} c_k r^k < \infty \text{ and } \sum_{k=0}^{\infty} k c_k r^{k-1} < \infty$$

we see that (f_n) converges to a function f on the subinterval $]-r, r[$ of $]-1, 1[$ and that (f_n') converges uniformly to a function g on $]-r, r[$. It follows from Theorem 167 that the function f is differentiable on $]-r, r[$ and that its derivative is g. Now we check that

$$\begin{aligned}
(1-x)f_n'(x) - \alpha f_n(x) &= \sum_{k=0}^{n} k c_k x^{k-1} - \sum_{k=0}^{n} k c_k x^k - \alpha \sum_{k=0}^{n} c_k x^k, \\
&= \sum_{k=0}^{n-1} (k+1) c_{k+1} x^k - \sum_{k=0}^{n} k c_k x^k - \alpha \sum_{k=0}^{n} c_k x^k, \\
&= -(n+1) c_{n+1} x^n, \qquad (8.12)
\end{aligned}$$

since $(k+1)c_{k+1} = (\alpha + k)c_k$. Letting n tend to ∞ we obtain the identity

$$(1-x)f'(x) = \alpha f(x) \qquad \text{for } -r < x < r,$$

from the fact that the last member of (8.12) tends to zero. The next step is to let $r \longrightarrow 1$ so that

$$(1-x)f'(x) = \alpha f(x) \qquad \text{for } -1 < x < 1.$$

It is of course obvious that the limit functions from two different values of r agree where they are both defined.

Now let

$$h(x) = (1-x)^{\alpha} f(x) \qquad \text{for } -1 < x < 1.$$

Then an easy calculation gives that $h'(x) = 0$ and hence h is constant on $]-1, 1[$. It follows that

$$(1-x)^{-\alpha} = \sum_{k=0}^{\infty} c_k x^k.$$

for $-1 < x < 1$. □

This example is a scalar one and does not use the full force of Theorem 167. If however one wished to define the exponential of a square matrix X by say

$$\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k$$

and then wished to discuss the differential of exp, then the Corollary would be very useful.

9

The Implicit Function Theorem and its Cousins

In this chapter we present a number of theorems that are linked together by their method of proof. They are all established by invoking the Contraction Mapping Theorem (Theorem 52, page 80). The conclusions of all of these results are local in nature, a circumstance which tends to make both the statements and proofs look more complicated than they really are.

Before delving into the real subject matter of this chapter, it is first necessary to understand the concept of invertibility of continuous linear mappings.

LEMMA 168 *Let V be a complete normed vector space and let $T \in \mathcal{CL}(V, V)$ be a bijection of V onto V.*

- *Then T is invertible in the sense that its inverse mapping T^{-1} belongs to $\mathcal{CL}(V, V)$.*
- *Then, there exist a strictly positive real number ϵ such that $S \in \mathcal{CL}(V, V)$ and $\|S - T\|_{\mathrm{op}} \le \epsilon$ together imply that S is invertible.*
- *There is a neighbourhood of T on which the mapping $S \longrightarrow S^{-1}$ is continuous.*

Proof. For the first statement, it is routine to show that T^{-1} is linear. It is continuous by an easy application of the Open Mapping Theorem (on page 93). Let $\epsilon = \|T^{-1}\|_{\mathrm{op}}^{-1}$ and suppose that $\|S - T\|_{\mathrm{op}} \le \frac{1}{2}\epsilon$. It follows that

$$\|I - T^{-1}S\|_{\mathrm{op}} = \|T^{-1}(T - S)\|_{\mathrm{op}} \le \|T^{-1}\|_{\mathrm{op}}\|T - S\|_{\mathrm{op}} \le \tfrac{1}{2}.$$

We denote $X = I - T^{-1}S$. Then let

$$Y_n = I + X + X^2 + \cdots + X^n. \tag{9.1}$$

It is easy to check that Y_n is a Cauchy sequence in the complete space $\mathcal{CL}(V, V)$ and hence is convergent to a limit Y. An easy calculation involving telescoping sums yields $Y_n(I - X) = I - X^{n+1}$. Letting n tend to ∞ gives $Y(I - X) = I$, or equivalently $YT^{-1}S = I$ showing that YT^{-1} is a left inverse of S. Of course an entirely similar argument establishes the existence of $Z \in \mathcal{CL}(V, V)$ such that $T^{-1}Z$ is a right inverse of S. Then both inverses necessarily coincide and it follows that S is invertible.

Since $\|X\|_{\mathrm{op}} \le \frac{1}{2}$ it follows from (9.1) that $\|Y\|_{\mathrm{op}} \le 2$. Thus we find $\|S^{-1}\|_{\mathrm{op}} \le 2\|T^{-1}\|_{\mathrm{op}}$. Now let us suppose that we have two linear operators S_1 and S_2 such that

$$\|S_j - T\|_{\mathrm{op}} \le \tfrac{1}{2}\epsilon \quad \text{for } j = 1, 2.$$

Then clearly

$$\begin{aligned}
\|S_1^{-1} - S_2^{-1}\|_{\mathrm{op}} &= \|S_1^{-1}(S_2 - S_1)S_2^{-1}\|_{\mathrm{op}} \\
&\le \|S_1^{-1}\|_{\mathrm{op}}\|S_2^{-1}\|_{\mathrm{op}}\|S_1 - S_2\|_{\mathrm{op}} \\
&\le 4\|T^{-1}\|_{\mathrm{op}}^2\|S_1 - S_2\|_{\mathrm{op}}.
\end{aligned}$$

It follows that $S \longrightarrow S^{-1}$ is Lipschitz on the closed ball $B(T, \epsilon)$. ■

A more detailed examination will show that $S \longrightarrow S^{-1}$ is infinitely differentiable on $B(T, \epsilon)$.

9.1 Implicit Functions

Suppose that we have a system of n non-linear equations for n unknowns. Of course, if there are no additional restrictions, there may be no solutions, a unique solution or many solutions. There is a simple condition that forces a solution to be unique at least locally.

Let $\Omega \subseteq \mathbb{R}^n$ be an open set and let $x_0 \in \Omega$. Let $f : \Omega \longrightarrow \mathbb{R}^n$ be a continuously differentiable function. We can write our system of equations in the form

$$f(x) = \mathbf{0}. \tag{9.2}$$

Let us suppose that the point x_0 is a solution. In other words, let us suppose that $f(x_0) = \mathbf{0}$. We seek conditions that force the solution x_0 to be isolated. Of course we can examine the question in the more general context in which $\mathbb{R}^n$ is replaced by a complete normed real vector space.

LEMMA 169 *Let Ω be an open subset of a complete normed real vector space V. Let $f : \Omega \longrightarrow V$ be a continuously differentiable mapping. Let $v_0 \in \Omega$ be such that $f(v_0) = 0_V$. If the differential df_{v_0} is invertible then $v = v_0$ is an isolated solution of the equation $f(v) = 0_V$.*

Proof. By the differentiability of f, there exists a function $\varphi : \Omega \longrightarrow V$ in the class $\mathcal{E}_{\Omega,v_0}$ such that

$$f(v) = f(v_0) + df_{v_0}(v - v_0) + \varphi(v) \qquad \forall v \in \Omega.$$

If we assume that v and v_0 are both solutions of $f(v) = 0_V$, then we obtain

$$df_{v_0}(v - v_0) + \varphi(v) = 0_V \qquad \forall v \in \Omega,$$

or equivalently that

$$v - v_0 = -(df_{v_0})^{-1}\varphi(v) \qquad \forall v \in \Omega. \tag{9.3}$$

using the hypothesis that df_{v_0} is invertible. Now, choose ϵ so small that

$$\epsilon\|(df_{v_0})^{-1}\|_{\text{op}} < \tfrac{1}{2},$$

where $\| \ \|_{\text{op}}$ denotes the norm of $\mathcal{CL}(V, V)$. Next, we invoke the definition of $\mathcal{E}_{\Omega,v_0}$, page 145, to obtain the existence of $\delta > 0$ such that

$$v \in \Omega, \|v - v_0\| < \delta \Rightarrow \|(df_{v_0})^{-1}\varphi(v)\| \le \tfrac{1}{2}\|v - v_0\|.$$

Combining this with (9.3) now yields

$$v \in \Omega, \|v - v_0\| < \delta \quad \Rightarrow \quad \|v - v_0\| \le \tfrac{1}{2}\|v - v_0\| \quad \Rightarrow \quad v = v_0.$$

This is the desired conclusion. ■

Now, in the finite dimensional situation, suppose that we have a system of n non-linear equations for n unknowns which depend upon p parameters. This is a situation that occurs quite frequently in applications. Suppose that a solution is known for a certain setting of the parameters. We would like to be able to track the known solution as the parameters vary slightly.

More generally, we may replace $\mathbb{R}^n$ by a complete real normed vector space V and the parameter space $\mathbb{R}^p$ by a complete real normed space U.

Theorem 170 (Implicit Function Theorem) *Let U and V be complete normed real vector spaces. Then $U \oplus V$ is a complete normed real vector space which can be identified as a metric space with $U \times V$. Let $\Omega \subseteq V$ be an open set and let $v_0 \in \Omega$. Let $\Delta \subseteq U$ be an open set and let $t_0 \in \Delta$. Let $g : \Delta \times \Omega \longrightarrow V$ be a continuously differentiable function with the property that $g(t_0, v_0) = 0_V$. Suppose that the partial differential $dg^2_{(t_0,v_0)}$ of g in the V subspace is invertible as a continuous linear map from V to V. Then there exists an open set Δ_0 of U such that*

$$t_0 \in \Delta_0 \subseteq \Delta,$$

and a continuously differentiable mapping $h : \Delta_0 \longrightarrow \Omega$ such that

$$g(t, h(t)) = 0_V,$$

for all $t \in \Delta_0$ and $h(t_0) = v_0$.

Here the variable $t \in \Delta$ represents the parameter in U and for a given value of t the corresponding system of equations is given by

$$g(t, v) = 0_V.$$

The Implicit Function Theorem asserts the existence of a solution $v = h(t)$ which varies as a continuously differentiable function of t. Note that this function is in general not globally defined. It need only exist in a small neighbourhood of t_0.

Example Suppose that $\Omega = V = \mathbb{R}$ and $\Delta = U = \mathbb{R}$, the case $n = p = 1$ described in the preliminaries. Let $g(t, v) = t - v^2$, $v_0 = 1$ and $t_0 = 1$. We have

$$dg^2_{(t,v)} = \frac{\partial g}{\partial v}(t, v) = -2v$$

which is certainly invertible when $v = v_0 = 1$. The solution function h is the positive square root. We see that h can only be defined for $t > 0$ despite the fact that $g(t, v)$ is defined and infinitely differentiable for all real t and v. We can also see that the initial conditions $v_0 = -1$ and $t_0 = 1$ lead to a different solution function. □

The usual statement of the Implicit Function Theorem contains a uniqueness assertion which is essentially a restatement of Lemma 169.

Proof of the Implicit Function Theorem. We first of all replace the function g by $(dg^2_{(t_0,v_0)})^{-1} \circ g$. This does not affect any of the hypotheses or conclusions of the

Theorem and allows us to assume that $dg^2_{(t_0,v_0)} = I$, the identity endomorphism on V. Define now $\varphi : \Delta \times \Omega \longrightarrow V$ by

$$\varphi(t, v) = v - g(t, v).$$

Then φ is clearly a continuously differentiable function with the property

$$d\varphi^2_{(t_0,v_0)} = 0,$$

the zero endomorphism of V. By the continuity of $d\varphi^2$, there exist neighbourhoods Δ_1 and Ω_0 of t_0 and v_0 respectively such that

$$(t, v) \in \Delta_1 \times \Omega_0 \Rightarrow \|d\varphi^2_{(t,v)}\|_{\mathrm{op}} \leq \tfrac{1}{2}$$

We further suppose without loss of generality that Ω_0 is the closed ball centred at v_0 of some radius $r > 0$. Further, using $g(t_0, v_0) = 0_V$ and the continuity of g we find an neighbourhood $\Delta_0 \subseteq \Delta_1$ of t_0 such that

$$t \in \Delta_0 \Rightarrow \|g(t, v_0)\| \leq \tfrac{1}{2}r.$$

Suppose now that h_1 and h_2 are two continuous mappings from Δ_0 to Ω_0. Then we have

$$\|\varphi(t, h_1(t)) - \varphi(t, h_2(t))\| \leq \sup_{v\in\Omega_0} \|d\varphi^2_{(t,v)}\|_{\mathrm{op}}\|h_1(t) - h_2(t)\|,$$

by Theorem 7.18 (page 156). In order to apply this Theorem, we need to have Ω_0 convex and not merely a neighbourhood of v_0. For $t \in \Delta_0$ this leads to

$$\|\varphi(t, h_1(t)) - \varphi(t, h_2(t))\| \leq \tfrac{1}{2}\|h_1(t) - h_2(t)\|. \tag{9.4}$$

The key idea then is to define a mapping T by

$$Th(t) = \varphi(t, h(t)).$$

A consequence of (9.4) is that T is a contraction mapping for the uniform norm in the sense of Theorem 52, the contraction constant being $\frac{1}{2}$.

Filling in the details of the proof so that Theorem 52 can actually be applied requires a considerable amount of work. The space S on which the Contraction Mapping Theorem will be applied is the space of all continuous mappings h from Δ_0 to Ω_0 such that $h(t_0) = v_0$. We give S the uniform norm and observe that it is a complete metric space with this norm. It is in order to obtain completeness that we insist that Ω_0 is a closed set. To check the base point condition we observe that for $h \in S$, we have $h(t_0) = v_0$ whence

$$Th(t_0) = \varphi(t_0, h(t_0)) = h(t_0) - g(t_0, h(t_0)) = v_0 - g(t_0, v_0) = v_0.$$

To check that S is nonempty we simply consider the constant function

$$h_0(t) = v_0 \qquad \forall t \in \Delta_0.$$

Finally, we need to check that T actually maps S to S. Let $h \in S$ and apply (9.4) to h and h_0 to obtain

$$\|Th(t) - Th_0(t)\| \le \tfrac{1}{2}\|h(t) - h_0(t)\| \le \tfrac{1}{2}r,$$

for all $t \in \Delta_0$. But

$$\|Th_0(t) - v_0\| = \|\varphi(t, v_0) - v_0\| = \|g(t, v_0)\| \le \tfrac{1}{2}r.$$

It follows that

$$\|Th(t) - v_0\| \le \|Th(t) - Th_0(t)\| + \|Th_0(t) - v_0\| \le r,$$

for $t \in \Delta_0$ as required.

With the hypotheses of Theorem 52 complete, we are now guaranteed the existence of a continuous function $h : \Delta_0 \longrightarrow \Omega_0$ such that $Th(t) = h(t)$ for all $t \in \Delta_0$. Equivalently this means that $g(t, h(t)) = 0_V$ for all $t \in \Delta_0$. Also $h(t_0) = v_0$.

The remainder of the proof is to check that the function h is actually continuously differentiable and not merely continuous. This bootstrap kind of approach is quite standard in results of this type. Let t_1 be a fixed element of Δ_0. Then

$$\begin{aligned} 0_V &= g(t, h(t)) - g(t_1, h(t_1)) \\ &= dg^1_{(t_1, h(t_1))}(t - t_1) + dg^2_{(t_1, h(t_1))}(h(t) - h(t_1)) + \psi(t, h(t)) \end{aligned}$$

where $t \in \Delta_0$ and $\psi \in \mathcal{E}_{\Delta_0 \times \Omega_0, (t_1, h(t_1))}$. Now apply the linear transformation $(dg^2_{(t_1, h(t_1))})^{-1}$ to obtain

$$0_V = A(t - t_1) + (h(t) - h(t_1)) + \psi_1(t, h(t)) \tag{9.5}$$

for $t \in \Delta_0$ and where again $\psi_1 \in \mathcal{E}_{\Delta_0 \times \Omega_0, (t_1, h(t_1))}$. In the interests of brevity, we have denoted

$$A = (dg^2_{(t_1, h(t_1))})^{-1} \circ dg^1_{(t_1, h(t_1))},$$

a continuous linear mapping from $A : U \longrightarrow V$. Because of the continuity of h at t_1, the little "o" condition can be rewritten as the statement

$$\forall\epsilon > 0, \exists\delta > 0 \quad \text{such that}$$
$$\|t - t_1\| < \delta \Rightarrow \|\psi_1(t, h(t))\| \leq \epsilon(\|t - t_1\| + \|h(t) - h(t_1)\|). \qquad (9.6)$$

Now choose $\epsilon = \frac{1}{2}$. Then, combining (9.5) and (9.6) we see that

$$\tfrac{1}{2}\|h(t) - h(t_1)\| \leq \|A(t - t_1)\| + \tfrac{1}{2}\|t - t_1\|$$

for $\|t - t_1\| < \delta$, showing that

$$\|h(t) - h(t_1)\| \leq C\|t - t_1\|$$

for some suitable constant C. Substituting back in (9.6), this shows in turn that $\psi_1(t, h(t))$ is in fact little "o" of $\|t - t_1\|$. Finally rewriting (9.5) as

$$h(t) - h(t_1) = -A(t - t_1) - \psi_1(t, h(t))$$

we see that h is differentiable at t_1 and that

$$dh_{t_1} = -A = -(dg^2_{(t_1,h(t_1))})^{-1} \circ dg^1_{(t_1,h(t_1))}. \qquad (9.7)$$

Since operator inversion and multiplication are continuous operations, we see that the mapping $t \longrightarrow dh_t$ is itself continuous. ■

The uniqueness assertion can be viewed on an individual point basis. Let $t_1 \in \Delta_0$ and $v \in \Omega_0$ satisfy $g(t_1, v_1) = 0_V$. Then $v_1 = h(t_1)$. To establish this, we write

$$v_1 - h(t_1) = \varphi(t_1, v_1) - \varphi(t_1, h(t_1))$$

and it follows that

$$\|v_1 - h(t_1)\| \leq \tfrac{1}{2}\|v_1 - h(t_1)\|$$

as in (9.4). The desired conclusion, $v_1 = h(t_1)$ follows.

We note that if in the hypotheses of the Implicit Function Theorem we add that g is continuously differentiable up to order $k \geq 1$, then we may conclude that the solution function h is also continuously differentiable up to order k. We leave the rather technical proof to the reader. The key point is that this assertion is proved *a posteriori*, by simply repeatedly differentiating the known first differential (9.7). It follows from this that if g is infinitely differentiable then so is h.

There is also another possible approach in which the additional differentiability is built into the metric space S to which the Contraction Mapping Theorem is applied. This would be an *a priori* method, technically much more difficult.

The Implicit Function Theorem is obtained from the Contraction Mapping Theorem, but it is also possible to use the Implicit Function Theorem to build a further extension of Theorem 53. The key point to observe is that in the hybrid Theorem that results, the implicit function g defined by the fixed point is guaranteed to exist on the whole of S and not merely in some local sense about some initial point.

THEOREM 171 *Let X and Y be complete normed spaces, $k \in \mathbb{N}$ and $0 \leq \alpha < 1$. Let $c \in X$ and let $r > 0$. Let P be an open subset of Y and suppose that $f : P \times B(c, r) \longrightarrow X$ is a mapping such that*

- $d(f(p, x_1), f(p, x_2)) \leq \alpha d(x_1, x_2)$ *for* $x_1, x_2 \in B(c, r)$ *and each* $p \in P$.
- $d(c, f(p, c)) < (1 - \alpha)r$ *for all* $p \in P$.
- *The mapping* $f|_{P \times U(c,r)}$ *obtained by restricting* f *to* $P \times U(c, r)$ *is* k *times continuously differentiable from* $P \times U(c, r)$ *to* X.

Then there is a unique k times continuously differentiable mapping $g : P \longrightarrow U(c, r)$ such that $f(p, g(p)) = g(p)$ for all $p \in P$.

Proof. The hypotheses of Theorem 171 exceed those of Theorem 53, and therefore, we are guaranteed the existence of a continuous mapping $g : P \longrightarrow U(c, r)$ such that $x = g(p)$ is the unique solution of the fixed point equation $f(p, x) = x$. It remains to show that g is in fact k times continuously differentiable. But this is a local property, so let us use the Implicit Function Theorem, locally about $(p, g(p))$ to establish that g is continuously differentiable near $p \in P$. Let us denote $\varphi(p, x) = x - f(p, x)$. Then the key hypothesis that is needed in the Implicit Function Theorem is that $d\varphi^2_{(p,g(p))} = I - df^2_{(p,g(p))}$ is an invertible linear transformation on X. But, by Proposition 129 we see that $\|df^2_{(p,g(p))}\|_{\text{op}} \leq \alpha$. Since $\alpha < 1$, we see that

$$\varphi^2_{(p,g(p))} = \sum_{n=0}^{\infty} (df^2_{(p,g(p))})^n$$

is given by a convergent series, c.f. Lemma 168. This proves the result in the case $k = 1$. For larger values of k we proceed as explained above by directly computing the derivatives using (9.7). ∎

9.2 Inverse Functions

The next task is to establish what is in effect a corollary of the Implicit Function Theorem.

THEOREM 172 (INVERSE FUNCTION THEOREM) *Let V be a complete normed real vector space. Let $\Omega \subseteq V$ be an open set and let $v_0 \in \Omega$. Let $f : \Omega \longrightarrow V$ be a continuously differentiable function. We denote $t_0 = f(v_0)$. Suppose that the differential df_{v_0} is invertible, then there exists a neighbourhood Δ_0 of t_0 in V and a continuously differentiable function $h : \Delta_0 \longrightarrow \Omega$ such that $h(t_0) = v_0$ and*

$$f \circ h(t) = t \qquad \forall t \in \Delta_0.$$

Proof. We will apply the Implicit Function Theorem. Let $\Delta = U = V$ and set

$$g(t, v) = t - f(v) \qquad \forall t \in V, v \in \Omega.$$

The Implicit Function Theorem immediately gives the desired conclusion. ■

As with the Implicit Function Theorem itself, there are some corollaries. If f is k times continuously differentiable with $k \geq 1$, then we can assert that so is the inverse function h. If f is infinitely differentiable, then we can assert that h is infinitely differentiable.

9.3 Parametrization of Level Sets

In this section we give another application of the Implicit Function Theorem in the finite dimensional case.

THEOREM 173 (PARAMETRIZATION THEOREM) *Let $k \leq n$ and suppose that $\Omega \subseteq \mathbb{R}^n$ is an open set. Let $x_0 \in \Omega$. Let $\theta : \Omega \longrightarrow \mathbb{R}^k$ be a continuously differentiable function such that the derivative $d\theta_{x_0}$ has full rank k and $\theta(x_0) = \mathbf{0}$. Then there exists a neighbourhood U of $\mathbf{0}$ in $\mathbb{R}^{n-k}$ and a continuously differentiable function $\varphi : U \longrightarrow \Omega$ such that $\varphi(\mathbf{0}) = x_0$, $d\varphi_t$ has rank $n - k$ and $\theta \circ \varphi(t) = \mathbf{0}$ for all $t \in U$.*

Furthermore there is a neighbourhood V of x_0 with $V \subseteq \Omega$ such that for all $x \in V$ such that $\theta(x) = \mathbf{0}$ there exists $t \in U$ such that $x = \varphi(t)$.

Proof. Since $d\theta_{x_0}$ has rank k, it follows that its null space N has dimension $n-k$. Let π be a linear projection of $\mathbb{R}^n$ onto N (as in a direct sum $\mathbb{R}^n = N \oplus C$). Now we define a map

$$g : N \times \Omega \longrightarrow N \times \mathbb{R}^k$$

by $g(t, x) = (\pi(x) - t, \theta(x))$. We are going to apply the Implicit Function Theorem to g. The function g is continuously differentiable on $N \times \Omega$ because θ is continuously differentiable on Ω. A calculation shows that

$$dg^2_{(\pi(x_0),x_0)}(x) = (\pi(x), d\theta_{x_0}(x)). \tag{9.8}$$

If the right-hand side of (9.8) vanishes then we have $x \in N$ and $\pi(x) = \mathbf{0}$. But since $x \in N$, $x = \pi(x) = \mathbf{0}$. It follows that the partial differential $dg^2_{(\pi(x_0),x_0)}$ is an injective linear mapping. By dimensionality it is therefore invertible.

The consequence of the Implicit Function Theorem is the existence of a neighbourhood U of $\pi(x_0)$ in N and a continuously differentiable mapping $\varphi : U \longrightarrow \Omega$ such that $g(t, \varphi(t)) = \mathbf{0}$ for all $t \in U$. Thus we have $\theta \circ \varphi(t) = \mathbf{0}$ and $\pi(\varphi(t)) = t$ for all such t. Differentiating this last relation yields that $\pi \circ d\varphi_t$ is the identity endomorphism on N. It follows that $d\varphi_t$ has full rank $n - k$.

Finally, from the uniqueness part of the Implicit Function Theorem, there exist neighbourhoods U_0 of $\pi(x_0)$ and Ω_0 of x_0 such that $t_1 \in U_0$, $x_1 \in \Omega_0$ and $g(t_1, x_1) = \mathbf{0}$ together imply that

$$x_1 = \varphi(t_1). \tag{9.9}$$

Let $V = \Omega_0 \cap \pi^{-1}(U_0)$. Then for $x_1 \in V$ satisfying $\theta(x_1) = \mathbf{0}$ we define $t_1 = \pi(x_1) \in U_0$ and it follows that

$$g(t_1, x_1) = (\pi(x_1) - t_1, \theta(x_1)) = \mathbf{0}.$$

The desired conclusion (9.9) follows. ■

9.4 Existence of Solutions to Ordinary Differential Equations

The purpose of this section is to outline the use of the Contraction Mapping Theorem to establish the Picard Existence Theorem for ordinary differential equations. We present the vector-valued case.

THEOREM 174 (PICARD EXISTENCE THEOREM) *Let V be a complete normed real vector space and let Ω be an open subset of V with $v_0 \in \Omega$. Let J be an open interval of $\mathbb{R}$ with $t_0 \in J$. Let $F : J \times \Omega \longrightarrow V$ be a bounded continuous mapping. Suppose also that the partial differential dF^2 exists on $J \times \Omega$. Suppose further that dF^2 is continuous on $J \times \Omega$ and uniformly bounded in the sense*

$$\sup_{t \in J, v \in \Omega} \|dF^2_{(t,v)}\|_{\mathrm{op}} < \infty.$$

Then there is an open subinterval J_0 of J with $t_0 \in J_0$ and a continuously differentiable function $\varphi : J_0 \longrightarrow \Omega$ such that $\varphi(t_0) = v_0$ and

$$\varphi'(t) = F(t, \varphi(t)) \qquad \forall t \in J_0.$$

Since the function φ is a function mapping from a 1-dimensional space, we have used derivatives rather than differentials to describe the result.

EXAMPLE This example illustrates that the Picard Existence Theorem is necessarily local in nature. Let $V = \Omega = J = \mathbb{R}$, $t_0 = v_0 = 0$ and let $F(t, v) = 1 + v^2$. Then the only solution of the differential equation $v' = 1 + v^2$ satisfying the initial condition $v(0) = 0$ is the function $v = \tan t$, which has singularities at $\frac{\pi}{2}$ and $-\frac{\pi}{2}$. Thus the interval J_0 is, at its largest $]-\frac{\pi}{2}, \frac{\pi}{2}[$. □

Proof. The first step in the proof is to find closed bounded neighbourhoods Ω_1 of v_0 and J_1 of t_0 contained respectively in Ω and J. Let us choose Ω_1 specifically in the form $B(v_0, r)$ for some $r > 0$. We note that in the finite dimensional case, the Heine–Borel Theorem (page 103) and the continuity of F and dF^2 allow us to establish that the functions F and dF^2 are bounded on $J_1 \times \Omega_1$. In the general case, we must avail ourselves of the given boundedness hypotheses.

We now choose $\ell > 0$ so small that

$$\ell \sup_{t \in J, v \in \Omega} \|F(t, v)\|_V \leq r \tag{9.10}$$

$$\ell \sup_{t \in J, v \in \Omega} \|dF^2_{(t,v)}\|_{\mathrm{op}} \leq \tfrac{1}{2} \tag{9.11}$$

both hold. Finally let J_0 be an open interval containing the point t_0 contained in both the intervals $]t_0 - \ell, t_0 + \ell[$ and J_1.

Now we let S be the metric space of continuous functions φ from the open interval J_0 to the closed ball Ω_1 such that $\varphi(t_0) = v_0$. We use the uniform metric on S. In other words,

$$d_S(\varphi_1, \varphi_2) = \sup_{t \in J_0} \|\varphi_1(t) - \varphi_2(t)\|_V.$$

The contraction mapping T is defined by

$$T\varphi(t) = v_0 + \int_{t_0}^{t} F(s, \varphi(s))ds \tag{9.12}$$

It is clear from Theorem 160 (page 177) that $T\varphi$ is a continuously differentiable function, and in particular it is certainly continuous. We have

$$\|T\varphi(t) - v_0\| \leq |t - t_0| \|F\|_\infty \leq r$$

for $t \in J_0$ (by (9.10) and the third part of Lemma 158) so that $T\varphi$ actually takes values in Ω_1. Thus T actually maps S to S and it is also clear that S is complete since Ω_1 is closed and V is complete. The details of the proof are found on page 78.

To verify that T is a contraction mapping on S we use the identity

$$T\varphi_1(t) - T\varphi_2(t) = \int_{t_0}^{t} F(s, \varphi_1(s)) - F(s, \varphi_2(s))ds$$

Since Ω_1 is a convex set, we can estimate

$$\|F(s, \varphi_1(s)) - F(s, \varphi_2(s))\|_V \leq \left\{ \sup_{t \in J_0, v \in \Omega_1} \|dF^2_{(t,v)}\|_{\text{op}} \right\} \left\{ \sup_{t \in J_0} \|\varphi_1(t) - \varphi_2(t)\|_V \right\}$$

using Theorem 139 (page 156). This leads, by (9.11) and again using the third part of Lemma 158, to the estimate

$$\sup_{t \in J_0} \|T\varphi_1(t) - T\varphi_2(t)\|_V \leq \tfrac{1}{2} \left\{ \sup_{t \in J_0} \|\varphi_1(t) - \varphi_2(t)\|_V \right\}. \tag{9.13}$$

Finally, this gives

$$d_S(T\varphi_1, T\varphi_2) \leq \tfrac{1}{2} d_S(\varphi_1, \varphi_2).$$

Using the Contraction Mapping Theorem we find that T has a fixed point φ. In other words, the fixed point φ satisfies

$$\varphi(t) = v_0 + \int_{t_0}^{t} F(s, \varphi(s))ds$$

for all t in J_0. It follows from the Fundamental Theorem of Calculus (page 177) that φ is differentiable and satisfies the required differential equation

$$\varphi'(t) = F(t, \varphi(t)). \tag{9.14}$$

Since, φ and F are known to be continuous, we see that φ is continuously differentiable. Compare this "bootstrap" type argument to the one found in the proof of the Implicit Function Theorem. The initial condition $\varphi(t_0) = v_0$ holds since $\varphi \in S$. ∎

There are a number of interesting extensions of this result. First of all, if F is k times continuously differentiable with k an integer $k \geq 1$, then one can conclude that the solution φ is $k+1$ times continuously differentiable. This is established *a posteriori*, that is, by repeatedly differentiating the differential equation (9.14).

Secondly, the the solution of (9.14) is unique, given the initial condition. This can be seen from the uniqueness assertion of the Contraction Mapping Theorem and also directly. If φ_1 and φ_2 are two solutions of (9.14) satisfying the same initial condition and defined on an interval $K \subseteq J$ containing t_0, then (9.13) becomes

$$\sup_{t\in K_1} \|\varphi_1(t) - \varphi_2(t)\|_V \leq \tfrac{1}{2} \left\{ \sup_{t\in K_1} \|\varphi_1(t) - \varphi_2(t)\|_V \right\}.$$

where K_1 is the interval $K_1 =]t_0 - \ell, t_0 + \ell[\cap K$. This shows that φ_1 and φ_2 coincide on K_1. If $K \neq K_1$, we may then reapply the same argument where t_0 is replaced by a point of K_1 close to one or other of its endpoints. This allows the equality of φ_1 and φ_2 be be extended into new territory. Repeating this procedure shows that equality holds on the whole of K.

Thirdly, under stronger hypotheses, typically satisfied by linear equations, we have the existence of a global solution.

COROLLARY 175 *Let V be a complete normed real vector space and let v_0 be a point of V. Let J be an open interval of $\mathbb{R}$ with $t_0 \in J$. Let $F : J \times V \longrightarrow V$ be a continuously differentiable mapping. Suppose also that the partial differential dF^2 satisfies*

$$\sup_{t\in I, v\in V} \|dF^2_{(t,v)}\|_{\text{op}} < \infty.$$

on every compact subinterval I of J. Then there is a continuously differentiable function $\varphi : J \longrightarrow V$ such that $\varphi(t_0) = v_0$ and

$$\varphi'(t) = F(t, \varphi(t)) \qquad \forall t \in J.$$

Proof. The proof follows that of the Picard Existence Theorem. Notice that $\Omega = V$ so that we can take $r = \infty$ and equation (9.10) disappears. We work on I, a compact subinterval I of J. Since the choice of ℓ depends now only on the equation

$$\ell \sup_{t\in I, v\in V} \|dF^2_{(t,v)}\|_{\mathrm{op}} \leq \tfrac{1}{2},$$

we see that ℓ depends only on I. We obtain a solution φ_0 on the interval $J_0 = \operatorname{int} I \cap \,]t_0 - \ell, t_0 + \ell[$. Now choose a point $t_1 \in J_0$ close to an endpoint of J_0. Repeating the argument establishes the existence of a solution φ_1 on the interval $J_1 = \operatorname{int} I \cap \,]t_1 - \ell, t_1 + \ell[$, satisfying the initial condition $\varphi_1(t_1) = \varphi_0(t_1)$. The uniqueness argument outlined above shows that these two solutions agree on the overlap $J_0 \cap J_1$ and hence it is possible to glue them together into a solution on $J_0 \cup J_1$. Since ℓ depends only on I, by repeating this idea we eventually can construct a solution φ_I on the whole of $\operatorname{int} I$. Next, we write

$$J = \bigcup_{k=1}^{\infty} I_k$$

where the I_k are compact subintervals of J. We can assume without loss of generality that I_k increases with k. Again uniqueness tells us that

$$\varphi_{I_p}|_{I_q} = \varphi_{I_q}$$

for $q \leq p$. Thus the solutions φ_{I_p} can be glued together to provide a solution on the whole of J. ■

We now prepare to tackle the issue of the dependence of the solution of a differential equation on its initial conditions or even on perturbations of the equation itself. We need the following little lemma which while entirely trivial involves a huge conceptual leap.

LEMMA 176 *Let V and W be complete normed vector spaces and let Ω be an open subset of V. Let $G : \Omega \longrightarrow W$ be a k times continuously differentiable function. Now consider the space X (respectively Y) of bounded continuous functions for an open interval I in $\mathbb{R}$ to Ω (respectively W) with the uniform metric. Consider the map $\tilde{G}$*

$$\varphi \mapsto (t \mapsto G(\varphi(t))).$$

Then $\tilde{G}$ is a k times continuously differentiable map from X to Y.

Proof. For $0 \le \ell \le k$ we have

$$d^\ell \tilde{G}_\varphi(\psi_1, \dots, \psi_\ell) = (t \mapsto d^\ell G_{\varphi(t)}(\psi_1(t), \dots, \psi_\ell(t))).$$

Note that the function φ does not get differentiated here! ∎

THEOREM 177 *Let k be an integer with $k \ge 1$ and let P be an open subset of a complete normed real vector space and let $p_0 \in P$. Let V be a complete normed real vector space and let Ω be an open subset of V. Let $v_0 : P \longrightarrow \Omega$ be a k times continuously differentiable function. Let J be an open interval of $\mathbb{R}$ with $t_0 \in J$. Let $F : P \times J \times \Omega \longrightarrow V$ be a k times continuously differentiable mapping. Suppose also that the partial differential dF^3 exists on $P \times J \times \Omega$. Suppose further that dF^3 is uniformly bounded on $P \times J \times \Omega$ in the sense*

$$\sup_{p \in P, t \in J, v \in \Omega} \|dF^3_{(p,t,v)}\|_{\mathrm{op}} < \infty.$$

Then there is a neighbourhood P_0 of p_0 in P, an open subinterval J_0 of J with $t_0 \in J_0$ and a continuously differentiable function $\varphi : P_0 \times J_0 \longrightarrow \Omega$ such that $\varphi(p, t_0) = v_0(p)$ and

$$\frac{\partial \varphi}{\partial t}(p,t) = F(t, \varphi(p,t)) \qquad \forall p \in P_0, t \in J_0. \tag{9.15}$$

Proof. The proof will follow the proof of the Picard Existence Theorem, but we are going to use Theorem 171 instead of the Contraction Mapping Theorem. We will need to choose $r > 0$ as in the Picard Theorem and also to choose ℓ such that

$$\ell \sup_{p \in P, t \in J, v \in \Omega} \|F(p,t,v)\|_V \le r$$
$$\ell \sup_{p \in P, t \in J, v \in \Omega} \|dF^3_{(p,t,v)}\|_{\mathrm{op}} \le \tfrac{1}{2}$$

and furthermore we will choose the neighbouhood P_0 such that

$$\|v_0(p) - v_0(p_0)\|_V \le \tfrac{1}{5} r$$
$$\ell \|F(p,t,v) - F(p_0,t,v)\|_V \le \tfrac{1}{5} r$$

for all p in P_0. Since these conditions are the same as in the Picard Theorem. There is a solution φ_0 wich satisfies the unperturbed differential equation

$$\varphi_0'(t) = F(p_0, t, \varphi_0(t)) \qquad |t - t_0| < \ell, \tag{9.16}$$

$$\varphi_0(t_0) = v_0(p_0). \tag{9.17}$$

Now we define the mapping

$$T(p, \varphi)(t) = v_0(p) + \int_{t_0}^{t} F(p, s, \varphi(s))ds.$$

and check that it satisfies the hypotheses of Theorem 171. First we check the contraction condition. We find

$$\begin{aligned}
&\|F(p, s, \varphi_1(s)) - F(p, s, \varphi_2(s))\|_V \\
&\qquad \le \left\{ \sup_{t\in J_0, v\in\Omega_1} \|dF^3_{(p,t,v)}\|_{\text{op}} \right\} \left\{ \sup_{t\in J_0} \|\varphi_1(t) - \varphi_2(t)\|_V \right\}
\end{aligned}$$

leading to

$$\begin{aligned}
&\sup_{p,t} \|T(p, \varphi_1)(t) - T(p, \varphi_2)(t)\|_V \\
&\qquad \le \ell \left\{ \sup_{p\in P_0, t\in J_0, v\in\Omega_1} \|dF^3_{(p,t,v)}\|_{\text{op}} \right\} \left\{ \sup_{t\in J_0} \|\varphi_1(t) - \varphi_2(t)\|_V \right\} \\
&\qquad \le \tfrac{1}{2} \left\{ \sup_{t\in J_0} \|\varphi_1(t) - \varphi_2(t)\|_V \right\}
\end{aligned}$$

The second condition of Theorem 171 is also easy to check. We need to estimate $\|\varphi_0 - T(p, \varphi_0)\|$. Towards this and using the fact that φ_0 satisfies the unperturbed equation we get

$$\varphi_0(t) - T(p, \varphi_0)(t) = v_0(p_0) - v_0(p) + \int_{t_0}^{t} F(p_0, s, \varphi_0(s)) - F(p, s, \varphi_0(s))ds$$

and the estimate $\|\varphi_0 - T(p, \varphi_0)\| \le \frac{4}{5}r < r$ follows from the inequalities (9.16) and (9.17) used in the choice of P_0. Finally, we see that $(p, \varphi) \mapsto T(p, \varphi)$ is k times differentiable much as in Lemma 176. Thus, all the required hypotheses are verified. Hence there is a k times continuously differentiable mapping $g : P_0 \longrightarrow U(\varphi_0, r)$ such that $g(p_0) = \varphi_0$ and $T(p, g(p)) = g(p)$. Obviously, we will set

$\varphi(p,t) = (g(p))(t)$. The conditions $\varphi(p,t_0) = v_0(p)$ and (9.15) are satisfied. It is also clear that $p \mapsto \varphi(p,t)$ is k times differentiable for each fixed t. Now using (9.15) and the fact that F is k times differentiable, we see that $(p,t) \mapsto \dfrac{\partial \varphi}{\partial t}(p,t)$ is k times continuously differentiable in p. This captures the first t derivative. To capture second and higher order differentiability in t, we must differentiate (9.15) further. We leave the details to the reader. ■

Index