CHAPTER 3. METRIC SPACES

The reader will perhaps agree with the assessment that the very foundation upon which the edifice of real analysis is erected is the idea of a *limit*. In order to understand this concept (even intuitively), let alone define it formally, one relies heavily on the fact that there is a (readymade) notion of *distance* between any two real numbers (which, we recall, is the absolute value of their difference). It is this central idea of a distance that we seek to abstract and extend beyond the realm of real numbers. Doing so leads us to the study of metric spaces, which is the focus of this chapter.

§3.1. Preliminaries

Definition 3.1.1. Suppose that M is a nonempty set. A function $\rho : M \times M \to \mathbf{R}$ is called a *metric* on M if the following hold:

(M1) (*Nonnegativity*) $\rho(x, y) \geq 0$ for every $x, y \in M$.
(M2) $\rho(x, y) = 0$ if and only if $x = y$.
(M3) (*Symmetry*) $\rho(y, x) = \rho(x, y)$ for every $x, y \in M$.
(M4) (*Triangle Inequality*) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ for every $x, y, z \in M$.

The pair (M, ρ) is called a *metric space*.

The following is the simplest example of a metric space.

Example 3.1.2. Suppose that M is any nonempty set. It may be verified that the function $d : M \times M \to \mathbf{R}$ given by

$$d(x, y) = \begin{cases} 1, & \text{if } x \neq y; \\ 0, & \text{if } x = y, \end{cases}$$

is a metric on M. It is called the *discrete metric*.

A convenient way of obtaining metric spaces is to use the notion of a norm (defined below). The reader is advised to recall here the concept of a *vector space*, which must have given her great joy in a course on linear algebra.

Definition 3.1.3. Suppose that V is a vector space over $\mathbf{R}$. A function $\|\cdot\| : V \to \mathbf{R}$ is said to be a *norm* on V if the following hold:

(N1) (*Nonnegativity*) $\|x\| \geq 0$ for every $x \in V$.
(N2) $\|x\| = 0$ if and only if $x = 0$ (where 0 is the additive identity in V).
(N3) (*Homogeneity*) $\|\alpha x\| = |\alpha| \, \|x\|$, for every $x \in V$ and every $\alpha \in \mathbf{R}$.
(N4) (*Triangle inequality*) $\|x + y\| \leq \|x\| + \|y\|$ for every $x, y \in V$.

The pair $(V, \|\cdot\|)$ is called a *normed linear space*.

The following simple result shows that every normed linear space is automatically a metric space.

Proposition 3.1.4. *Suppose that $(V, \|\cdot\|)$ is a normed linear space. The function ρ, defined by $\rho(x, y) := \|x - y\|$, $x, y \in V$, is a metric on V (called the metric induced by the norm).*

Proof. EXERCISE. ∎

Example 3.1.5. (i) Let $|\cdot|$ denote the (usual) absolute value (of a real number). It may be verified that $(\mathbf{R}, |\cdot|)$ is a normed linear space. The corresponding metric (induced by the norm) is called the *usual metric* on $\mathbf{R}$.

(ii) Suppose that n is a fixed positive integer, and let $V = \mathbf{R}^n$, the set of all $n \times 1$ (column) matrices of real numbers. The reader will verify that the function

$$\|\mathbf{x}\|_1 := \sum_{k=1}^{n} |x_k|, \qquad \mathbf{x} = [\, x_1 \quad \cdots \quad x_n \,]^T ,$$

is a norm on $\mathbf{R}^n$. The normed linear space $(\mathbf{R}^n, \|\cdot\|_1)$ is denoted by ℓ_n^1.
(iii) It may also be verified that the function

$$\|\mathbf{x}\|_\infty := \max_{1 \le k \le n} |x_k|, \qquad \mathbf{x} = [\, x_1 \quad \cdots \quad x_n \,]^T ,$$

is a norm on $\mathbf{R}^n$. The normed linear space $(\mathbf{R}^n, \|\cdot\|_\infty)$ is denoted by ℓ_n^∞.
(iv) Let a and b be fixed real numbers with $a < b$. It may be verified that the set

$$C[a,b] := \{f : [a,b] \to \mathbf{R} \,:\, f \text{ is continuous on } [a,b]\}$$

is an infinite-dimensional vector space over $\mathbf{R}$. Moreover, the function

$$\|f\|_\infty := \max\{|f(x)| \,:\, a \le x \le b\}$$

is a norm on $C[a,b]$. It is usually called the *uniform norm* or the *supremum norm*.

Our next task is to define an entire family of norms on $\mathbf{R}^n$, one which will include Examples 3.1.5(i) and (ii) discussed above. We begin with some preliminaries.

Lemma 3.1.6. *If A and B are fixed positive numbers and $0 \le t \le 1$, then $A^t B^{1-t} \le tA + (1-t)B$.*

Proof. The function $f(x) = -\ln x$ is concave up on the interval $(0, \infty)$. Therefore

$$f(tA + (1-t)B) \le f(A) + \frac{f(B) - f(A)}{B - A}\,[tA + (1-t)B - A] = tf(A) + (1-t)f(B);$$

that is,

$$-\ln(tA + (1-t)B) \le t(-\ln A) + (1-t)(-\ln B) = -[\ln(A^t) + \ln(B^{1-t})] = -\ln(A^t B^{1-t}).$$

The required result follows upon exponentiation.

Aliter. The result being evident for $t = 0$ and $t = 1$, we shall suppose that $0 < t < 1$, and define

$$G(x) := tx + (1-t)B - x^t B^{1-t}, \qquad x > 0.$$

Elementary calculus shows that $G'(x) = t(1 - B^{1-t} x^{t-1})$, $x > 0$, and hence that

$$\min\{G(x) \,:\, x > 0\} = G(B) = 0.$$

It follows that $G(A) \ge 0$, which is a restatement of the stated inequality. ∎

The next pair of results is among the most basic in mathematics.

Theorem 3.1.7. *(Hölder's Inequality) Suppose that n is a fixed positive integer, and let $x_k, y_k \in \mathbf{R}$, $1 \le k \le n$. If $1 < p < \infty$ and $(1/p) + (1/q) = 1$, then*

$$\sum_{k=1}^{n} |x_k y_k| \le \left[\sum_{k=1}^{n} |x_k|^p\right]^{1/p} \left[\sum_{k=1}^{n} |y_k|^q\right]^{1/q}.$$

Proof. If $\sum_{k=1}^{n} |x_k|^p$ or $\sum_{k=1}^{n} |y_k|^q$ is zero, then the result is immediate. Assume now that neither quantity is zero and define

$$A_k := \frac{|x_k|^p}{\sum_{l=1}^{n} |x_l|^p} \quad \text{and} \quad B_k := \frac{|y_k|^q}{\sum_{l=1}^{n} |y_l|^q}, \qquad 1 \le k \le n.$$

If neither A_k nor B_k is zero, then Lemma 3.1.6 (with $t = 1/p$) provides the estimate

$$A_k^{1/p} B_k^{1/q} \le \frac{1}{p} A_k + \frac{1}{q} B_k. \tag{3.1.1}$$

On the other hand, the foregoing inequality is clearly valid if either A_k or B_k is zero. Thus (3.1.1) holds for every $1 \le k \le n$. In other words,

$$\frac{|x_k|}{\left(\sum_{l=1}^{n} |x_l|^p\right)^{1/p}} \frac{|y_k|}{\left(\sum_{l=1}^{n} |y_l|^q\right)^{1/q}} \le \frac{1}{p}\left(\frac{|x_k|^p}{\sum_{l=1}^{n} |x_l|^p}\right) + \frac{1}{q}\left(\frac{|y_k|^q}{\sum_{l=1}^{n} |y_l|^q}\right), \qquad 1 \le k \le n.$$

Summing from $k = 1$ to n we obtain

$$\frac{\sum_{k=1}^{n} |x_k y_k|}{\left(\sum_{l=1}^{n} |x_l|^p\right)^{1/p} \left(\sum_{l=1}^{n} |y_l|^q\right)^{1/q}} \le \frac{1}{p}\left(\frac{\sum_{k=1}^{n} |x_k|^p}{\sum_{l=1}^{n} |x_l|^p}\right) + \frac{1}{q}\left(\frac{\sum_{k=1}^{n} |y_k|^q}{\sum_{l=1}^{n} |y_l|^q}\right) = \frac{1}{p} + \frac{1}{q} = 1,$$

whence the result, via cross multiplication. ∎

Theorem 3.1.8. *(Minkowski's Inequality) Suppose that n is a fixed positive integer, and let $x_k, y_k \in \mathbf{R}$, $1 \le k \le n$. If $1 \le p < \infty$, then*

$$\left[\sum_{k=1}^{n} |x_k + y_k|^p\right]^{1/p} \le \left[\sum_{k=1}^{n} |x_k|^p\right]^{1/p} + \left[\sum_{k=1}^{n} |y_k|^p\right]^{1/p}.$$

Proof. The required result is obvious if $\sum_{k=1}^{n} |x_k + y_k|^p = 0$, so we may assume that $\sum_{k=1}^{n} |x_k + y_k|^p > 0$. Furthermore, the result for $p = 1$ is a direct consequence of the triangle inequality for real numbers; so let us assume that $p > 1$ and use the triangle inequality for real numbers once again:

$$\sum_{k=1}^{n} |x_k + y_k|^p = \sum_{k=1}^{n} |x_k + y_k|\, |x_k + y_k|^{p-1} \le \sum_{k=1}^{n} |x_k|\, |x_k + y_k|^{p-1} + \sum_{k=1}^{n} |y_k|\, |x_k + y_k|^{p-1}.$$

Let $1/q := 1 - (1/p)$. Applying Hölder's inequality to each of the summands on the right we obtain

$$\begin{aligned} \sum_{k=1}^{n} |x_k + y_k|^p &\le \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q}\right)^{1/q} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |x_k + y_k|^{(p-1)q}\right)^{1/q} \\ &= \left[\left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{1/p}\right] \left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{1/q}, \end{aligned}$$

the final step coming from the fact that $(1/p) + (1/q) = 1$. Remembering this relationship between p and q and dividing by the last term on the right yields the desired inequality. ∎

Example 3.1.9. Let n be a fixed positive integer and let $1 < p < \infty$. The reader should now be able to verify that the function

$$\|\mathbf{x}\|_p := \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}, \qquad \mathbf{x} = [\,x_1 \quad \cdots \quad x_n\,]^T,$$

is a norm on $\mathbf{R}^n$. The normed linear space $(\mathbf{R}^n, \|\cdot\|_p)$ is denoted by ℓ_n^p.

§3.2. Basic topology of metric spaces

The seeds of analysis on a metric space will be sown in this section.

Definition 3.2.1. Suppose that (M, ρ) is a metric space and $R > 0$. Let $x_0 \in M$. The *ρ-disc of radius R, centred at x_0*, is the set

$$U_\rho(x_0; R) := \{x \in M \,:\, \rho(x, x_0) < R\}.$$

Example 3.2.2. (i) Consider the real line $\mathbf{R}$ equipped with the usual metric, say ρ. If x_0 is any real number and $R > 0$, then $U_\rho(x_0; R)$ is the open interval $(x_0 - R, x_0 + R)$.
(ii) Let ρ denote the metric induced by the norm $\|\cdot\|_2$ on $\mathbf{R}^2$. Let $x_0 := [0 \quad 0]^T$ and $R > 0$. Then

$$U_\rho(x_0; R) = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \,:\, \sqrt{x_1^2 + x_2^2} < R \right\}.$$

(iii) Let ρ denote the metric induced by the norm $\|\cdot\|_1$ on $\mathbf{R}^2$. Let $x_0 := [0 \quad 0]^T$ and $R > 0$. Then

$$U_\rho(x_0; R) = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \,:\, |x_1| + |x_2| < R \right\}.$$

(iv) Let d denote the discrete metric on a nonempty set M. If $x_0 \in M$ and R is any positive number, then

$$U_d(x_0; R) = \begin{cases} \{x_0\}, & \text{if } R \le 1; \\ M, & \text{if } R > 1. \end{cases}$$

The following definition is of central importance in the theory of metric spaces.

Definition 3.2.3. (i) A nonempty subset G of a metric space (M, ρ) is said to be *open* (more precisely *ρ-open*) if for every $g \in G$, there is a positive number r_g (which may depend on g) such that $U_\rho(g; r_g) \subseteq G$. We note that the master space M is open (by definition), as is the empty set (vacuously).
(ii) A subset of a metric space (M, ρ) is said to be *closed* (more precisely *ρ-closed*) if its complement is ρ-open. Note that both the empty set and M are closed sets.

Example 3.2.4. Verification of the following facts is left to the reader.

(i) The interval $[0, 1]$ is a closed subset of the real line equipped with the usual metric. It is not an open set.
(ii) The set of natural numbers is a closed subset of the real line equipped with the usual metric. It is not an open set.
(iii) The interval $[0, 1)$ is neither closed nor open (with respect to the usual metric on $\mathbf{R}$).

A basic (and important) example of an open set is given below.

Theorem 3.2.5. *Suppose that (M, ρ) is a metric space and that R is a positive number. Let $x_0 \in M$. The disc $U_\rho(x_0; R)$ is a ρ-open subset of M.*

Proof. Let $g \in U_\rho(x_0; R)$, so that $\rho(x_0, g) < R$. Define $r_g := R - \rho(x_0, g) > 0$ and let $x \in U_\rho(g; r_g)$. Then $\rho(x, g) < r_g$ and the triangle inequality and symmetry (axioms (M3) and (M4) of Definition 3.1.1) imply the relations

$$\rho(x, x_0) \leq \rho(x, g) + \rho(g, x_0) < r_g + \rho(x_0, g) = R.$$

In other words $x \in U_\rho(x_0; R)$, hence $U_\rho(g; r_g) \subseteq U_\rho(x_0; R)$. As g was chosen arbitrarily, we conclude that $U_\rho(x_0; R)$ is an open set. ∎

We now turn our attention to the notion of convergent sequences in metric spaces.

Definition 3.2.6. Let $\{a_n\}$ be a sequence of elements in a metric space (M, ρ), *i.e.,* $a_n \in M$ for every positive integer n. Let $a \in M$. We shall say that $\{a_n\}$ *converges in (M, ρ) to a as n tends to infinity*, and write $\lim_{n\to\infty} a_n = a$ or $\{a_n\} \xrightarrow{n} a$, if $\lim_{n\to\infty} \rho(a_n, a) = 0$.

A quantitative description of convergence is given in the proposition below, which should be compared with Remark 2.2.1(3).

Proposition 3.2.7. *Let $\{a_n\}$ be a sequence in a metric space (M, ρ), and let $a \in M$. The following are equivalent:*
(i) $\lim_{n\to\infty} a_n = a$;
(ii) For every $\epsilon > 0$ there is a positive integer N (which may depend on ϵ) such that $\rho(a_n, a) < \epsilon$ for every $n \geq N$.

Proof. EXERCISE. ∎

Remark 3.2.8. The reader is invited to boost her self-confidence by attempting a proof of the following assertion: Let $\{a_n\}$ be a sequence in a metric space. If $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} a_n = b$, then $a = b$.

The next result provides a useful description of closed sets in terms of sequential convergence.

Theorem 3.2.9. *Suppose that F is a nonempty subset of a metric space (M, ρ). The following are equivalent:*
(i) F is ρ-closed.
(ii) If $\{a_n\}$ is a sequence in F and $\lim_{n\to\infty} a_n = a$, then $a \in F$.

Proof. Suppose that F is closed, and let $\{a_n\}$ be a sequence in F converging to an element a in M. We must show that a belongs to F. Now if $a \notin F$, then a belongs to F^c (the complement of F), an open set. So there is a positive number r_a such that $U_\rho(a; r_a) \subseteq F^c$. The convergence of $\{a_n\}$ to

a implies, via Proposition 3.2.7, the existence of a positive integer N such that $\rho(a_N, a) < r_a$, that is, $a_N \in U_\rho(a; r_a) \subseteq F^c$, contradicting the fact that every element of the sequence $\{a_n\}$ belongs to F.

Conversely, assume that condition (ii) holds. We shall show that F^c is open. If F^c is empty we are done, so let us suppose that it is nonempty and pick an arbitrary element g from it. We must demonstrate the existence of a positive number r_g such that $U_\rho(g; r_g) \subseteq F^c$. Now if no such number exists, then $U_\rho(g; R)$ must intersect F for every positive number R. In particular $U_\rho(g; 1/n) \cap F \neq \emptyset$ for every positive integer n. Choose $a_n \in U_\rho(g; 1/n) \cap F$, $n \in \mathbf{N}$. Then $\{a_n\}$ is a sequence in F and $\lim_{n\to\infty} a_n = g$ because $\rho(a_n, g) < 1/n$ for every n (implying that $\lim_{n\to\infty}\rho(a_n, g) = 0$). However g belongs to the complement of F and we have reached a contradiction to condition (ii). ∎

§3.3. Completeness

The following definition should be compared with Definition 2.2.12.

Definition 3.3.1. Let $\{x_n\}$ be a sequence in a metric space (M, ρ). We say that $\{x_n\}$ is *Cauchy* (more precisely *ρ-Cauchy*) if for every $\epsilon > 0$ there is a positive integer N (which may depend on ϵ) such that $\rho(x_m, x_n) < \epsilon$ for every $m, n \geq N$.

A sequence $\{x_n\}$ in a metric space (M, ρ) is said to be *convergent* if there is an $x \in M$ such that $\lim_{n\to\infty} x_n = x$. The upcoming result provides a basic connexion between convergent sequences and Cauchy sequences.

Theorem 3.3.2. *Let $\{x_n\}$ be a sequence in a metric space (M, ρ). If $\{x_n\}$ is convergent, then it is also Cauchy.*

Proof. EXERCISE (Mimic the proof of Theorem 2.2.13.) ∎

Definition 3.3.3. A metric space (M, ρ) is said to be *complete* if every Cauchy sequence in M converges to an element in M.

Example 3.3.4. (i) The reader will happily recall from Theorem 2.2.15 that the real line, equipped with the usual metric, is complete.
(ii) Let d denote the discrete metric on a nonempty set M. The reader should prove that (M, d) is complete.

Definition 3.3.5. A normed linear space $(V, \|\cdot\|)$ is called a *Banach space* if it is complete with respect to the metric induced by the norm.

Some very important examples of Banach spaces will be discussed below.

Example 3.3.6. The space ℓ_n^p is a Banach space for every positive integer n and every $1 \leq p \leq \infty$. We shall consider the case $1 \leq p < \infty$, leaving the case $p = \infty$ as an exercise for the reader. Let $\{\mathbf{x}^{(m)} : m \in \mathbf{N}\}$ be a Cauchy sequence in ℓ_n^p, $1 \leq p < \infty$. Let us denote $\mathbf{x}^{(m)} = \begin{bmatrix} x_1^{(m)} & \dots & x_n^{(m)} \end{bmatrix}^T$. Given $\epsilon > 0$, there is a positive integer $N = N_\epsilon$ such that $\|\mathbf{x}^{(r)} - \mathbf{x}^{(s)}\|_p < \epsilon$ for every $r, s \geq N$. That is,

$$\sum_{k=1}^{n} |x_k^{(r)} - x_k^{(s)}|^p < \epsilon^p, \qquad r, s \geq N, \tag{3.3.1}$$

whence $|x_k^{(r)} - x_k^{(s)}| < \epsilon$ for every $r, s \geq N$ and every $1 \leq k \leq n$. Thus for each (fixed) $1 \leq k \leq n$, the sequence $\{x_k^{(m)} : m \in \mathbf{N}\}$ is a Cauchy sequence of real numbers. So Theorem 2.2.15 provides a real number $\hat{x}_k$ such that $\lim_{m\to\infty} x_k^{(m)} = \hat{x}_k$. We shall show that, as m tends to infinity, $\{\mathbf{x}^{(m)} : m \in \mathbf{N}\}$ converges in ℓ_n^p to $\hat{\mathbf{x}} := [\,\hat{x}_1 \quad \ldots \quad \hat{x}_n\,]^T$. To that end let $\epsilon > 0$ be given and let $N = N_\epsilon$ be as above. Fixing $r \geq N$ and letting s go to infinity in (3.3.1) leads to the inequality

$$\sum_{k=1}^{n} |x_k^{(r)} - \hat{x}_k|^p \leq \epsilon^p, \qquad r \geq N;$$

equivalently,

$$\|\mathbf{x}^{(r)} - \hat{\mathbf{x}}\|_p \leq \epsilon, \qquad r \geq N,$$

and the proof is complete.

Let $\mathcal{S}$ denote the set of all real sequences, to wit:

$$\mathcal{S} := \{\, \{a_n\} \,:\, a_n \in \mathbf{R} \ \forall n \in \mathbf{N} \,\}.$$

Given a pair of sequences $\mathbf{a} = \{a_n\}$ and $\mathbf{b} = \{b_n\}$ in $\mathcal{S}$ and a real number α we define the operations of addition and scalar multiplication in $\mathcal{S}$ in the following natural manner:

$$\mathbf{a} + \mathbf{b} := \{a_n + b_n\} \quad \text{and} \quad \alpha\mathbf{a} := \{\alpha a_n\}.$$

It is a fairly routine exercise – which the reader is encouraged to attempt – to show that $\mathcal{S}$, endowed with the pair of operations given above, is an infinite-dimensional vector space over $\mathbf{R}$. The zero element in this vector space, namely the sequence all of whose elements are zero, will be denoted by $\mathbf{0}$. In what follows we shall be interested in some special (and classical) subspaces of $\mathcal{S}$.

Definition 3.3.7. (i) Let $1 \leq p < \infty$ be a fixed number. We define

$$\ell^p := \{\, \{a_n\} \in \mathcal{S} \,:\, \sum_{n=1}^{\infty} |a_n|^p < \infty \,\}.$$

(ii) The space of all bounded (real) sequences is denoted by ℓ^∞; precisely,

$$\ell^\infty := \{\, \{a_n\} \in \mathcal{S} \,:\, \sup_n |a_n| < \infty \,\}.$$

Proposition 3.3.8. *The sets ℓ^p, $1 \leq p \leq \infty$, are (vector) subspaces of $\mathcal{S}$. Moreover, the function*

$$\|\mathbf{a}\|_p := \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p}, \qquad \mathbf{a} = \{a_n\},$$

is a norm on ℓ^p for $1 \leq p < \infty$, and the function

$$\|\mathbf{a}\|_\infty := \sup\{|a_n| \,:\, n \in \mathbf{N}\}, \qquad \mathbf{a} = \{a_n\},$$

is a norm on ℓ^∞.

Proof. We shall consider the case $1 \le p < \infty$ and leave the case $p = \infty$ as an exercise for the reader. If $\mathbf{a} = \{a_n\} \in \ell^p$ and α is any real number, then

$$\lim_{N\to\infty} \sum_{n=1}^{N} |\alpha a_n|^p = \lim_{N\to\infty} |\alpha|^p \sum_{n=1}^{N} |a_n|^p = |\alpha|^p \sum_{n=1}^{\infty} |a_n|^p.$$

This demonstrates that ℓ^p is closed under scalar multiplication and also that $\|\alpha\mathbf{a}\|_p = |\alpha|\,\|\mathbf{a}\|_p$. Suppose now that $\mathbf{a} = \{a_n\}$ and $\mathbf{b} = \{b_n\}$ belong to ℓ^p. Minkowski's inequality asserts that, for any positive integer N,

$$\begin{aligned}\left(\sum_{n=1}^{N} |a_n + b_n|^p\right)^{1/p} &\le \left(\sum_{n=1}^{N} |a_n|^p\right)^{1/p} + \left(\sum_{n=1}^{N} |b_n|^p\right)^{1/p} \\ &\le \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |b_n|^p\right)^{1/p}.\end{aligned}$$

It follows that the infinite series $\sum_{n=1}^{\infty} |a_n + b_n|^p$ is convergent and that

$$\left(\sum_{n=1}^{\infty} |a_n + b_n|^p\right)^{1/p} \le \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |b_n|^p\right)^{1/p}.$$

In other words $\mathbf{a}+\mathbf{b} \in \ell^p$ and $\|\mathbf{a}+\mathbf{b}\|_p \le \|\mathbf{a}\|_p + \|\mathbf{b}\|_p$. The proof is completed upon observing the following: (i) the function $\|\cdot\|_p$ is patently nonnegative, (ii) $\|\mathbf{0}\|_p$ is evidently zero, and (iii) if the sum of an infinite series of nonnegative numbers is zero, then each summand must be zero. Ergo, if $\|\{a_n\}\|_p = 0$, then $\sum_{n=1}^{\infty} |a_n|^p = 0$, whence $|a_n|^p = 0$ for every n; that is $\mathbf{a} = \mathbf{0}$. ∎

Example 3.3.9. The space $(\ell^p, \|\cdot\|_p)$ is a Banach space for every $1 \le p \le \infty$. We shall take up the case $1 \le p < \infty$ and leave the remaining case, as usual, for the reader. We begin with a Cauchy sequence $\{\mathbf{a}^{(m)} : m \in \mathbf{N}\}$ in ℓ^p, with a view to showing that it converges, in $(\ell^p, \|\cdot\|_p)$, to an element in ℓ^p. Let $\mathbf{a}^{(m)} := \{a_n^{(m)} : n \in \mathbf{N}\}$, $m \in \mathbf{N}$. Given $\epsilon > 0$, there is a positive integer $N = N_\epsilon$ such that $\|\mathbf{a}^{(r)} - \mathbf{a}^{(s)}\|_p < \epsilon$ for every $r, s \ge N$. That is,

$$\sum_{n=1}^{\infty} |a_n^{(r)} - a_n^{(s)}|^p < \epsilon^p, \qquad r, s \ge N, \tag{3.3.2}$$

whence $|a_n^{(r)} - a_n^{(s)}| < \epsilon$ for every $r, s \ge N$ and every positive integer n. Thus for each (fixed) $n \in \mathbf{N}$, the sequence $\{a_n^{(m)} : m \in \mathbf{N}\}$ is a Cauchy sequence of real numbers. So Theorem 2.2.15 provides a real number $\hat{a}_n$ such that $\lim_{m\to\infty} a_n^{(m)} = \hat{a}_n$. We now show that the sequence $\hat{\mathbf{a}} := \{\hat{a}_n\}$ belongs to ℓ^p. The sequence $\{\mathbf{a}^{(m)} : m \in \mathbf{N}\}$ being Cauchy in ℓ^p, there is a positive integer K such that $\|\mathbf{a}^{(m)} - \mathbf{a}^{(K)}\|_p < 1$ for every $m \ge K$. Hence the triangle inequality guarantees that

$$\|\mathbf{a}^{(m)}\|_p = \|\mathbf{a}^{(m)} - \mathbf{a}^{(K)} + \mathbf{a}^{(K)}\|_p \le \|\mathbf{a}^{(m)} - \mathbf{a}^{(K)}\|_p + \|\mathbf{a}^{(K)}\|_p < 1 + \|\mathbf{a}^{(K)}\|_p =: \Delta, \quad m \ge K. \tag{3.3.3}$$

So for any positive integer S,

$$\sum_{n=1}^{S} |a_n^{(m)}|^p \le \|\mathbf{a}^{(m)}\|_p^p < \Delta^p, \qquad m \ge K. \tag{3.3.4}$$

As Δ is independent of m, we may let the latter quantity approach infinity in (3.3.4) to obtain

$$\sum_{n=1}^{S} |\hat{a}_n|^p \leq \Delta^p. \tag{3.3.5}$$

The choice of S was arbitrary and Δ does not depend on S, so (3.3.5) implies that the infinite series $\sum_{n=1}^{\infty} |\hat{a}_n|^p$ is convergent; that is, $\hat{\mathbf{a}}$ belongs to ℓ^p.

The argument will be concluded by showing that $\lim_{m\to\infty}\|\mathbf{a}^{(m)} - \hat{\mathbf{a}}\|_p = 0$ (the quantity $\|\mathbf{a}^{(m)} - \hat{\mathbf{a}}\|_p$ is well defined because both $\mathbf{a}^{(m)}$ and $\hat{\mathbf{a}}$ belong to ℓ^p, so the same is true of their difference). To that end let $\epsilon > 0$ be given and let $N = N_\epsilon$ be chosen in order to guarantee (3.3.2). Now (3.3.2) also implies that, for any positive integer S,

$$\sum_{n=1}^{S} |a_n^{(r)} - a_n^{(s)}|^p < \epsilon^p, \qquad r, s \geq N. \tag{3.3.6}$$

Keeping $r \geq N$ fixed and letting s approach infinity in (3.3.6), one obtains

$$\sum_{n=1}^{S} |a_n^{(r)} - \hat{a}_n|^p \leq \epsilon^p, \qquad r \geq N.$$

This being true for every S, we conclude that

$$\sum_{n=1}^{\infty} |a_n^{(r)} - \hat{a}_n|^p \leq \epsilon^p, \qquad r \geq N,$$

which is equivalent to the assertion that $\|\mathbf{a}^{(r)} - \hat{\mathbf{a}}\|_p \leq \epsilon$ whenever $r \geq N$. This completes the proof.

We shall now give a useful result connecting completeness and closed sets. It is worthwhile to bear in mind that, if (M, ρ) is a metric space and A is a subset of M, then (A, ρ) is automatically a metric space in its own right.

Proposition 3.3.10. *Suppose that (M, ρ) is a metric space and that A is a nonvoid subset of M. The following hold:*
(i) If (A, ρ) is complete then A is a ρ-closed subset of M.
(ii) If (M, ρ) is complete and A is a ρ-closed subset of M, then (A, ρ) is complete.

Proof. (i) We use Theorem 3.2.9. Let $\{a_n\}$ be a sequence in A, and assume that $\lim_{n\to\infty} a_n = a$. We must show that $a \in A$. As $\{a_n\}$ is a convergent sequence in (A, ρ), it is also Cauchy (Theorem 3.3.2). Now (A, ρ) is complete, so $\{a_n\}$ must converge to an element in A. By uniqueness of limit (Remark 3.2.8), a must belong to A.

(ii) Let $\{x_n\}$ be a Cauchy sequence in (A, ρ); so it is also a Cauchy sequence in (M, ρ). The completeness of (M, ρ) guarantees an $x \in M$ such that $\lim_{n\to\infty} x_n = x$. As A is a closed set, Theorem 3.2.9 asserts that $x \in A$. Thus (A, ρ) is complete. ∎

We wrap up the section with an elegant and well-known theorem. Should her mathematical excursions lead her into such territory, the reader may expect to encounter a classical application of this theorem in the basic theory of ordinary differential equations.

Theorem 3.3.11. *(Banach's Contraction Principle) Suppose that* (M, ρ) *is a complete metric space. Let* $T : M \to M$ *be a contraction, to wit, there is a number* α *such that* $0 < \alpha < 1$ *and* $\rho(T(x), T(y)) \leq \alpha\rho(x, y)$ *for every* x *and* y *in* M. *Then there is a unique element* x^* *in* M *such that* $T(x^*) = x^*$.

Proof. Let x_0 be an (arbitrary) element in M. Define a sequence $\{x_n\}$ inductively as follows: $x_n := T(x_{n-1})$, $n \in \mathbf{N}$. Using the hypothesis that T is a contraction along with simple induction, one arrives at the estimate $\rho(x_{n+1}, x_n) \leq \alpha^n \rho(x_1, x_0)$, for every positive integer n. Consequently, if r and s are positive integers and $s > r$, then (a repeated application of) the triangle inequality provides the following:

$$\rho(x_s, x_r) \leq \sum_{n=r}^{s-1} \rho(x_{n+1}, x_n) \leq \rho(x_1, x_0) \sum_{n=r}^{s-1} \alpha^n. \tag{3.3.7}$$

Remembering that the geometric series $\sum_{n=1}^{\infty} \alpha^n$ is convergent (α being a positive number smaller than 1), we find that, for every $\epsilon > 0$ there is a positive integer $N = N_\epsilon$ such that $\sum_{n=r}^{s-1} \alpha^n < \epsilon/(1+\rho(x_1, x_0))$ whenever $s > r \geq N$. This fact, coupled with (3.3.7), reveals that $\{x_n\}$ is a Cauchy sequence in M, and the completeness of the latter ensures an $x^* \in M$ such that $\lim_{n\to\infty} x_n = x^*$. Now

$$0 \leq \rho(T(x^*), x^*) \leq \rho(T(x^*), T(x_n)) + \rho(T(x_n), x^*) \leq \alpha\rho(x_n, x^*) + \rho(x_{n+1}, x^*), \quad n \in \mathbf{N}, \tag{3.3.8}$$

and each summand on the far right of (3.3.8) converges to zero as n tends to infinity (because $\lim_{n\to\infty} x_n = x^* = \lim_{n\to\infty} x_{n+1}$). It follows that $\rho(T(x^*), x^*) = 0$, and hence that $T(x^*) = x^*$. Finally, if $T(y^*) = y^*$ for $y^* \in M$, then

$$\rho(x^*, y^*) = \rho(T(x^*), T(y^*)) \leq \alpha\rho(x^*, y^*),$$

which, on account of the fact that $0 < \alpha < 1$, would lead to a contradiction unless $\rho(x^*, y^*) = 0$. The proof is complete. ∎

§3.4. Boundedness and total boundedness

Definition 3.4.1. Let (M, ρ) be a metric space and let A be a nonempty subset of M. The *diameter* of A is defined as follows:

$$\text{diam}(A) := \sup\{\rho(x, y) \, : \, x, y \in A\}.$$

We shall say that a subset of a metric space is *bounded* if it is empty or if it has finite diameter.

We now take up a refinement of the notion of boundedness. It will play an important role in our subsequent studies.

Definition 3.4.2. Suppose that (M, ρ) is a metric space and that A is a subset of M. We say that A is *totally bounded* if for every $\epsilon > 0$, there exist finitely many subsets, say $A_1, \ldots, A_{N(\epsilon)}$, such that the following hold:
(i) $\text{diam}(A_j) \leq \epsilon$ for every $1 \leq j \leq N(\epsilon)$.
(ii) $A = \cup_{j=1}^{N(\epsilon)} A_j$.

It is easy to prove the following:

Proposition 3.4.3. *Suppose that (M, ρ) is a metric space. The following hold:*
(i) Every finite subset of M is totally bounded.
(ii) If $B \subseteq A \subseteq M$ and A is totally bounded, then so is B.

Proof. EXERCISE. ∎

A basic (and perhaps expected) connexion between boundedness and total boundedness is brought out in the result below.

Proposition 3.4.4. *Suppose that (M, ρ) is a metric space and that $A \subseteq M$. If A is totally bounded, then it is also bounded.*

Proof. Total boundedness provides subsets $A_1, \ldots, A_m$ of M such that $\operatorname{diam}(A_j) \leq 1$ for every $1 \leq j \leq m$ and $A = \cup_{j=1}^{m} A_j$. Select (and fix) $a_j \in A_j$, $1 \leq j \leq m$, and define $\Delta := \max\{\rho(a_r, a_s) : 1 \leq r, s \leq m\}$. Suppose now that x and y are (any) two elements in A. Then $x \in A_k$ and $y \in A_l$ for some $1 \leq k, l \leq m$. The triangle inequality allows us to deduce that

$$\rho(x, y) \leq \rho(x, a_k) + \rho(a_k, a_l) + \rho(a_l, y) \leq \operatorname{diam}(A_k) + \Delta + \operatorname{diam}(A_l) \leq 2 + \Delta,$$

whereupon we may conclude that $\operatorname{diam}(A) \leq 2 + \Delta$. ∎

The purpose of the upcoming example is to show that total boundedness is a strictly stronger condition than boundedness.

Example 3.4.5. Let d denote the discrete metric on an infinite set M. The definition of d implies that the diameter of M is 1; in particular, (M, d) is bounded. We shall prove that (M, d) is not totally bounded by showing that M cannot be realized as a finite union of sets each of which has diameter at most $1/2$. Indeed, if B is a nonempty subset of M and $\operatorname{diam}(B) \leq 1/2$, then B must be a singleton (because the distance between any two distinct points in M is 1). So the finite union of such sets must be a finite set, whereas M is infinite.

The following theorem, which is of central importance, presents a very useful characterization of totally bounded sets.

Theorem 3.4.6. *Suppose that (M, ρ) is a metric space, and that A is a nonempty subset of M. The following are equivalent:*
(i) A is totally bounded.
(ii) Every sequence in A has a Cauchy subsequence.

Proof. Suppose that A is totally bounded, and let $\{a_n\}$ be a sequence in A. There exist sets $A_1^{(1)}, \ldots A_{N(1)}^{(1)}$ such that $\operatorname{diam}(A_j^{(1)}) \leq 1$ for every $1 \leq j \leq N(1)$ and $A = \cup_{j=1}^{N(1)} A_j^{(1)}$. The pigeonhole principle asserts that there must be some $A_k^{(1)}$, $1 \leq k \leq N(1)$, so that $a_n \in A_k^{(1)}$ for infinitely many values of n. Define $B_1 := A_k^{(1)}$. As B_1 is a subset of (the totally bounded set) A, B_1 must be totally bounded as well (Proposition 3.4.3). So there exist sets $A_j^{(2)}$, $1 \leq j \leq N(2)$, such that $\operatorname{diam}(A_j^{(2)}) \leq 1/2$ for every $1 \leq j \leq N(2)$ and $B_1 = \cup_{j=1}^{N(2)} A_j^{(2)}$. As before there is some set $A_l^{(2)}$ which contains a_n for infinitely many values of n. Let us call this set B_2. Proceeding inductively, we obtain a nested sequence of sets $B_1 \supseteq B_2 \supseteq \cdots$ such that $\operatorname{diam}(B_k) \leq 1/k$ for every $k \in \mathbf{N}$ and each B_k contains a_n for infinitely many values of n. Let $a_{n_1} \in B_1$. As B_2 contains a_n for infinitely many values of n, there is a positive integer $n_2 > n_1$ such that $a_{n_2} \in B_2$. Proceeding inductively we obtain a sequence of positive integers $n_1 < n_2 \cdots$ such that $a_{n_k} \in B_k$ for every $k \in \mathbf{N}$. Clearly

$\{a_{n_k} : k \in \mathbf{N}\}$ is a subsequence of $\{a_n\}$, and we assert that it is Cauchy. To that end let $\epsilon > 0$ be given, and choose a positive integer p such that $1/p < \epsilon$. If $n_r, n_s \geq n_p$, then both a_{n_r} and a_{n_s} belong to B_p. Hence $\rho(a_{n_r}, a_{n_s}) \leq \mathrm{diam}(B_p) \leq 1/p < \epsilon$.

Conversely, let us suppose that A is not totally bounded and exhibit a sequence in A which does not admit a Cauchy subsequence. The failure of total boundedness means that there is some $\epsilon_0 > 0$ such that A cannot be expressed as a finite union of sets of diameter at most ϵ_0. Equivalently, A cannot be realized as a subset of a finite union of sets whose diameters are at most ϵ_0.

Let $x_1 \in A$ and consider the disc $S_1 := U_\rho(x_1; \epsilon_0/2)$. The triangle inequality implies that, for $s, t \in S_1$, $\rho(s,t) \leq \rho(s, x_1) + \rho(x_1, t) < \epsilon_0$, whence $\mathrm{diam}(S_1) \leq \epsilon_0$. So by assumption there must be an element x_2 in A which does not belong to S_1. Let $S_2 := U_\rho(x_2; \epsilon_0/2)$, which, via the foregoing argument, is also a set of diameter at most ϵ_0. As A is not totally bounded there is some $x_3 \in A$ which does not belong to S_1 or S_2; in particular $\rho(x_3, x_j) \geq \epsilon_0/2$ for $j \in \{1, 2\}$. Continuing in this fashion we obtain a sequence $\{x_n\}$ in A and sets $S_j := U_\rho(x_j; \epsilon_0/2)$, $j \in \mathbf{N}$, such that x_n does not belong to $\cup_{j=1}^{n-1} S_j$ for every positive integer n. It follows that $\rho(x_r, x_s) \geq \epsilon_0/2$ for every pair of positive integers r and s, whence it is seen that no subsequence of $\{x_n\}$ can be Cauchy. This concludes the proof. ∎

The purport of Proposition 3.4.4 and Example 3.4.5 is that every totally bounded set is bounded but not vice versa. However, there is one very important case in which the two notions coincide, and this is what we propose to investigate next. The following preludial lemma will help set the stage.

Lemma 3.4.7. *Let A be a bounded subset of the real line $\mathbf{R}$, equipped with the usual metric. There is some positive number T such that $A \subseteq [-T, T]$.*

Proof. Select and fix an $a_0 \in A$. Then

$$|a| = |a - a_0 + a_0| \leq |a - a_0| + |a_0| \leq \mathrm{diam}(A) + |a_0|, \qquad a \in A.$$

Taking $T := \mathrm{diam}(A) + |a_0|$ yields the required result. ∎

Theorem 3.4.8. *Suppose that A is a subset of the real line $\mathbf{R}$, equipped with the usual metric. If A is bounded, then it is also totally bounded.*

Proof. Thanks to Theorem 3.4.6, it suffices to prove that every sequence in A has a Cauchy subsequence. We shall employ the classical method of bisection (or 'divide and conquer'). Let $\{a_n\}$ be a sequence in A, and let T be the number supplied by the preceding lemma. As $A \subseteq [-T, T] =: I_0$, one of the two intervals $[-T, 0]$ or $[0, T]$ must contain a_n for infinitely many values of n. Let I_1 denote this interval. Divide I_1 into two equal parts, so that one of the resultant subintervals, call it I_2, must contain a_n for infinitely many values of n. Proceeding in this fashion we obtain intervals $I_1 \supseteq I_2 \supseteq \cdots$ such that the length of I_k is half that of I_{k-1} for every positive integer k, and each I_k contains a_n for infinitely many values of n. The latter condition provides (consult the proof of Theorem 3.4.6) an increasing sequence of positive integers, say $n_1 < n_2 < \cdots$, such that $a_{n_k} \in I_k$ for every $k \in \mathbf{N}$. We show that the sequence $\{a_{n_k} : k \in \mathbf{N}\}$ – which is clearly a subsequence of $\{a_n\}$ – is Cauchy. To that end, let $\epsilon > 0$ be given and select a positive integer p such that $T/2^{p-1} < \epsilon$. If $n_r, n_s \geq n_p$, then both a_{n_r} and a_{n_s} belong to I_p; so $|a_{n_r} - a_{n_s}| \leq \mathrm{diam}(I_p) = T/2^{p-1} < \epsilon$. ∎

We close the section with a description of total boundedness which will be of use later.

Proposition 3.4.9. *Suppose that (M, ρ) is a metric space, and that A is a (nonempty) subset of M. The following are equivalent:*

(i) *A is totally bounded.*

(ii) *Given* $r > 0$*, there exist finitely many points* $x_1, \dots x_p$ *in* A *such that* $A \subseteq \bigcup_{k=1}^{p} U_\rho(x_k; r)$.

(iii) *Given* $\epsilon > 0$*, there exist sets* $B_1, \dots B_q$ *such that diam*$(B_k) \leq \epsilon$ *for every* $1 \leq k \leq q$*, and* $A \subseteq \bigcup_{k=1}^{q} B_k$.

Proof. Suppose that A is totally bounded, and let r be a given positive number. Total boundedness supplies sets $A_1, \dots A_p$ such that $\text{diam}(A_k) \leq r/2$ for every $1 \leq k \leq p$ and $A = \bigcup_{k=1}^{p} A_k$. Pick (and fix) points $x_k \in A_k$, for every $1 \leq k \leq p$; then each $x_k \in A$ because A_k is a subset of A. Moreover, if $k \in \{1, \dots, p\}$ and $x \in A_k$, then $\rho(x, x_k) \leq \text{diam}(A_k) \leq r/2 < r$, so $A_k \subseteq U_\rho(x_k; r)$. Consequently $A \subseteq \bigcup_{k=1}^{p} U_\rho(x_k; r)$, and this proves that statement (i) implies statement (ii).

Suppose now that (ii) holds, and let $\epsilon > 0$ be given. By assumption there exist points $x_1, \dots x_q$ such that $A \subseteq \bigcup_{k=1}^{q} U_\rho(x_k; \epsilon/2)$. Let $B_k := U_\rho(x_k; \epsilon/2)$ and observe that $\text{diam}(B_k) \leq \epsilon$ (if $x, y \in U_\rho(x_k; \epsilon/2)$, then $\rho(x, y) \leq \rho(x, x_k) + \rho(x_k, y) < \epsilon/2 + \epsilon/2 = \epsilon$). Thus (iii) is proven.

Assume that (iii) holds, and let $\epsilon > 0$ be given. There exist sets $B_1, \dots B_q$ such that $\text{diam}(B_k) \leq \epsilon$ for every $1 \leq k \leq q$, and $A \subseteq \bigcup_{k=1}^{q} B_k$. Put $A_k := B_k \cap A$ so that $A = \bigcup_{k=1}^{q} A_k$. Furthermore, as $A_k \subseteq B_k$, we find that $\text{diam}(A_k) \leq \text{diam}(B_k) \leq \epsilon$ for every k. It follows that A is totally bounded. ∎

§3.5. Compactness

In this section we introduce one of the most fundamental concepts in analysis.

Definition 3.5.1. Suppose that (M, ρ) is a metric space, and that A is a subset of M. We say that A is *compact* if (A, ρ) is complete and totally bounded.

The following famous theorem provides a pleasing description of compact subsets of the real line.

Theorem 3.5.2. *(Heine-Borel) Suppose that* A *is a subset of the real line* $\mathbf{R}$*, equipped with the usual metric. Then* A *is compact if and only if it is closed and bounded.*

Proof. If A is compact then it is complete, hence closed (Proposition 3.3.10). Moreover compactness implies total boundedness, hence boundedness (Proposition 3.4.4).

On the other hand, if A is closed and bounded, then it is totally bounded (Theorem 3.4.8). Furthermore, $\mathbf{R}$ is complete (Theorem 2.2.15) and A is a closed subset thereof, so A must be complete (Proposition 3.3.10). ∎

We now wish to give an extremely useful characterization of compactness. A preliminary lemma is in order.

Lemma 3.5.3. *Let* $\{a_n\}$ *be a Cauchy sequence in a metric space* (M, ρ)*. Assume that there is a subsequence* $\{a_{n_k} : k \in \mathbf{N}\}$ *of* $\{a_n\}$ *and an element* $a \in M$ *such that* $\lim_{k\to\infty} a_{n_k} = a$*. Then* $\lim_{n\to\infty} a_n = a$*.*

Proof. EXERCISE (Mimic the proof of Theorem 2.2.14). ∎

Theorem 3.5.4. *Suppose that (M, ρ) is a metric space, and that A is a subset of M. The following are equivalent:*
(i) A is compact.
(ii) If $\{a_n\}$ is a sequence in A, then there is a subsequence $\{a_{n_k} : k \in \mathbf{N}\}$ of $\{a_n\}$ and an element $a \in A$ such that $\lim_{k\to\infty} a_{n_k} = a$.

Proof. Assume that A is compact and let $\{a_n\}$ be a sequence in A. As A is compact, it is totally bounded, hence Theorem 3.4.6 ensures that $\{a_n\}$ admits a Cauchy subsequence, say $\{a_{n_k} : k \in \mathbf{N}\}$. The compactness of A also means that (A, ρ) is complete. So there is some $a \in A$ such that $\lim_{k\to\infty} a_{n_k} = a$.

Let us now suppose that condition (ii) is fulfilled. This condition, combined with the fact that every convergent sequence is Cauchy (Theorem 3.3.2), implies that every sequence in A admits a Cauchy subsequence. Therefore A is totally bounded (via Theorem 3.4.6), and it remains to show that A is complete. To that end let $\{a_n\}$ be a Cauchy sequence in (A, ρ). Our assumption provides a subsequence $\{a_{n_k} : k \in \mathbf{N}\}$ of $\{a_n\}$ as well as an element $a \in A$ such that $\lim_{k\to\infty} a_{n_k} = a$. So Lemma 3.5.3 asserts that $\lim_{n\to\infty} a_n = a$. ∎

Our final task in this section is to describe compactness from a different angle, specifically, in terms of open sets. The following definition sets the stage:

Definition 3.5.5. Suppose that (M, ρ) is a metric space and that A is a subset of M. We say that a collection $\{U_\alpha : \alpha \in \Lambda\}$ of (ρ-) open sets is an *open cover* for A if $A \subseteq \bigcup_{\alpha\in\Lambda} U_\alpha$.

Armed with the preceding definition, we are ready to formulate an alternate description of compactness.

Theorem 3.5.6. *Suppose that (M, ρ) is a metric space and that A is a (nonempty) subset of M. The following are equivalent:*
(i) A is compact.
(ii) Every open cover $\{U_\alpha : \alpha \in \Lambda\}$ of A admits a finite subcover, namely there is a finite subset Θ of Λ such that $A \subseteq \bigcup_{\alpha\in\Theta} U_\alpha$.

Proof. Assume that A is compact, and let $\{U_\alpha : \alpha \in \Lambda\}$ be an open cover for A. We claim that there exists an $r > 0$ such that, for every $x \in A$, there is an $\alpha(x) \in \Lambda$ so that the disc $U_\rho(x; r) \subseteq U_{\alpha(x)}$. If no such r exists, then for every positive integer n, we can find an $x_n \in A$ such that $U_\rho(x_n; 1/n)$ is not contained in U_α for any α. The compactness of A provides, via Theorem 3.5.4, a subsequence $\{x_{n_k} : k \in \mathbf{N}\}$ and a point $\hat{x} \in A$ such that $\lim_{k\to\infty} x_{n_k} = \hat{x}$. As $\{U_\alpha : \alpha \in \Lambda\}$ is a cover for A, there is some $\lambda \in \Lambda$ such that $\hat{x} \in U_\lambda$, and the fact that U_λ is open allows us to find a positive number δ such that $U_\rho(\hat{x}; \delta) \subseteq U_\lambda$. Choose a positive integer l so large that each of the numbers $1/n_l$ and $\rho(x_{n_l}, \hat{x})$ is less than $\delta/2$. So if $\rho(y, x_{n_l}) < 1/n_l$, then $\rho(y, \hat{x}) \leq \rho(y, x_{n_l}) + \rho(x_{n_l}, \hat{x}) < \delta$, whence $U_\rho(x_{n_l}; 1/n_l) \subseteq U_\rho(\hat{x}; \delta) \subseteq U_\lambda$. But this contradicts the very choice of the sequence $\{x_n\}$, so our initial claim is validated. To reiterate, we have established the existence of a positive number r such that for every $x \in A$, there is an $\alpha(x) \in \Lambda$ so that the disc $U_\rho(x; r) \subseteq U_{\alpha(x)}$. As A is compact, it is totally bounded, hence Proposition 3.4.9 guarantees finitely many points $x_1, \ldots, x_p$ in A so that $A \subseteq \bigcup_{k=1}^{p} U_\rho(x_k; r)$. Picking one (and exactly one) $\alpha(x_k)$ for each $1 \leq k \leq p$ and defining $\Theta := \{\alpha(x_k) : 1 \leq k \leq p\}$, we find that $A \subseteq \bigcup_{\alpha\in\Theta} U_\alpha$.

Assume now that condition (ii) holds, namely that every open cover for A admits a finite subcover. We show that every sequence in A admits a convergent subsequence, thus proving the compactness of A (via Theorem 3.5.4). Let $\{a_n\}$ be a sequence in A, and define $U_n := \{x \in M : \inf_{k>n} \rho(x, a_k) > 0\}$, $n \in \mathbf{N}$. We assert that each U_n is an open set. To that end let $y \in U_n$ and let $\inf_{k>n} \rho(y, a_k) =: \alpha > 0$. If $t \in U_\rho(y; \alpha/2)$ and $k > n$, then $\rho(t, a_k) \geq \rho(a_k, y) - \rho(y, t) > \alpha/2$, so $\inf_{k>n} \rho(t, a_k) \geq \alpha/2 > 0$. Thus $U_\rho(y; \alpha/2) \subseteq U_n$ and we have proved that U_n is open. As $U_n \subseteq U_{n+1}$ for every n (via Proposition 2.1.9(ii)), we see that, if Θ is any finite subset of the positive integers, then $\bigcup_{n\in\Theta} U_n = U_N$, where $N = \max\{n : n \in \Theta\}$. Combining this with the observation that $a_{N+1} \notin U_N$ for any N, we deduce that $A \not\subseteq \bigcup_{n\in\Theta} U_n$ for any finite subset Θ of $\mathbf{N}$, so our operative assumption on A translates to the statement that $\{U_n : n \in \mathbf{N}\}$ is not an open cover for A. Thus there is some $a \in A$ such that $\inf_{k>n} \rho(a, a_k) = 0$ for every positive integer n. Putting $n = 1$ in this last equation provides, via Proposition 2.1.8, a positive integer $n_1 > 1$ such that $\rho(a, a_{n_1}) < 1$. As $\inf_{k>n_1} \rho(a, a_k) = 0$, the aforesaid proposition supplies a positive integer $n_2 > n_1$ such that $\rho(a, a_{n_2}) < 1/2$. Proceeding inductively in this manner, we construct a subsequence $\{a_{n_k} : k \in \mathbf{N}\}$ of $\{a_n\}$ such that $\rho(a, a_{n_k}) < 1/k$ for every k. It follows that $\lim_{k\to\infty} a_{n_k} = a$ and the proof is complete. ∎

§3.6. Connectedness

In this section we undertake a brief study of connectedness in metric spaces. The discerning reader will observe that the flavour here is a bit less 'analytic' than what has been encountered hitherto.

Definition 3.6.1. Suppose that (M, ρ) is a metric space, and that A is a subset of M.
(i) We say that A is *disconnected* if there exist ρ-open sets G and H such that (i) $A \subseteq G \cup H$, (ii) both $A \cap G$ and $A \cap H$ are nonempty, and (iii) $A \cap G \cap H = \emptyset$. The sets G and H are said to provide a *disconnection* for A.
(ii) We say that A is *connected* if it is not disconnected.

Example 3.6.2. (i) Consider the real line $\mathbf{R}$ equipped with the usual metric. The set $\mathbf{Z}$ of integers is a disconnected subset of $\mathbf{R}$, for it may be verified that the sets $G = (-\infty, 1/2)$ and $H = (1/2, \infty)$ provide a disconnection for $\mathbf{Z}$.
(ii) Consider the normed linear space ℓ_2^2 and let $A := \{[\,x_1 \quad x_2\,]^T \in \mathbf{R}^2 : x_1x_2 = 1\}$. The reader is invited to check that the sets $G = \{[\,x_1 \quad x_2\,]^T \in \mathbf{R}^2 : x_1 > 0, x_2 > 0\}$ and $H = \{[\,x_1 \quad x_2\,]^T \in \mathbf{R}^2 : x_1 < 0, x_2 < 0\}$ provide a disconnection for A.

Remark 3.6.3. Let (M, ρ) be a metric space. Confirming the following simple fact is left as a task for the reader: M is disconnected if there exist nonempty open sets G and H such that $M = G \cup H$ and $G \cap H = \emptyset$. As before, G and H are said to provide a disconnection for M.

Here is an alternate description of disconnected metric spaces.

Theorem 3.6.4. *Suppose that (M, ρ) is a metric space. The following are equivalent:*
(i) M is disconnected.
(ii) There exists a nonempty, proper subset of M which is both ρ-open and ρ-closed.

Proof. Suppose that M is disconnected, and let G and H be as in Remark 3.6.3. Now G is an open set. Furthermore, it is a proper subset of M, because its complement, namely H, is nonempty. As H is also open, G must be closed as well.

Conversely, if S is a nonempty, proper subset of M which is both open and closed, then S and its complement provide a disconnection for M. ∎

We note in passing that Theorem 3.6.4 may be restated as follows: a metric space (M, ρ) is connected if and only if M does not contain any nonempty, proper subset which is both open and closed.

Example 3.6.5. Let d denote the discrete metric on a nonempty set M. The reader will verify that every subset of M is both d-open and d-closed. In particular (M, d) is disconnected.

The section concludes with an elegant description of the structure of connected subsets of the real line.

Theorem 3.6.6. *Suppose that E is a nonempty subset of the real line* $\mathbf{R}$*, equipped with the usual metric. The following are equivalent:*
(i) E is connected.
(ii) E has the following property: if $x, y \in E$ and $x < z < y$, then $z \in E$.

Proof. If E does not possess the property given in (ii), then there exist $x, y \in E$ and a $z_0 \notin E$ such that $x < z_0 < y$. Consider the open sets $G := (-\infty, z_0)$ and $H := (z_0, \infty)$. As $x \in G \cap E$ and $y \in H \cap E$, neither set is empty. Moreover, $G \cap H \cap E \subseteq G \cap H = \emptyset$, and the fact that $z_0 \notin E$ guarantees that $E \subseteq G \cup H$. Thus G and H provide a disconnection for E. This proves that condition (i) implies condition (ii).

The converse is trickier. We show that, if E satisfies condition (ii) and G and H are open sets such that neither $G \cap E$ nor $H \cap E$ is empty but $G \cap H \cap E = \emptyset$, then E cannot be contained in the union of G and H. This will show that condition (ii) implies condition (i). Let G and H be as above, and pick $x \in G \cap E$ and $y \in H \cap E$. We may assume without loss that $x < y$. Define $S := \{s \in G \cap E \,:\, s < y\}$ and note that it is a nonempty set because it contains x. Moreover, y is an upper bound of S, so $\alpha := \sup(S)$ is well defined. As $x \in S$ and α is the least upper bound of S, we find that $x \leq \alpha \leq y$, whence $\alpha \in E$ by assumption. We shall show that α cannot belong to G or H, thereby finishing the proof. Now if $\alpha \in H$, then the openness of H furnishes a positive number r such that $(\alpha - r, \alpha + r) \subseteq H$. Remembering Proposition 2.1.8 along with the definition of α yields an $s_0 \in S$ such that $\alpha - r < s_0 \leq \alpha$. In particular $s_0 \in G \cap E \cap H$, contradicting the fact that this set is empty. Thus $\alpha \notin H$. On the other hand, if $\alpha \in G$, then $\alpha < y$ because $\alpha \leq y$ and $y \notin G$. Combining this with the fact that G is open gives us a positive number η such that $(\alpha - \eta, \alpha + \eta) \subseteq G$ and $\alpha + \eta < y$. Now if t is a number such that $\alpha < t < \alpha + \eta$, then $x \leq \alpha < t < y$, and hence $t \in E$ by assumption. Therefore $t \in G \cap E$ and $t < y$, so $t \in S$, violating the fact that α is an upper bound of S. Thus $\alpha \notin G$ and the proof is complete. ∎

Remark 3.6.7. The reader should now be able to convince herself that the *only* connected subsets of the real line (equipped with the usual metric) are intervals.

§3.7. Continuous functions on metric spaces

Having concentrated thus far on various properties of metric spaces, we move now to the treatment of functions that are defined on metric spaces. We begin with an all important definition.

Definition 3.7.1. Suppose that (M, ρ) and (M', ρ') are metric spaces, and let $f : M \to M'$ be a function. We say that f is *continuous* at a point $x_0 \in M$ (or more precisely $\rho - \rho'$ *continuous* at x_0) if the following holds: for every sequence $\{x_n\}$ in M for which $\lim_{n\to\infty} x_n = x_0$, we have

$\lim_{n\to\infty} f(x_n) = f(x_0)$ in M'. We shall say that f is *continuous on M* if f is continuous at every point in M.

The following characterization of continuity (at a point) should be reminiscent of the well-known '$\epsilon - \delta$ definition' encountered in real analysis.

Theorem 3.7.2. *Suppose that (M, ρ) and (M', ρ') are metric spaces, and let $f : M \to M'$ be a function. The following are equivalent:*
(i) f is continuous at $x_0 \in M$.
(ii) For every $\epsilon > 0$ there is a $\delta > 0$ such that $\rho'(f(x), f(x_0)) < \epsilon$ whenever $\rho(x, x_0) < \delta$.

Proof. The failure of condition (ii) means that there is some $\epsilon_0 > 0$ such that, for every $\delta > 0$ there is some $x_\delta \in M$ for which $\rho(x_\delta, x_0) < \delta$ but $\rho'(f(x_\delta), f(x_0)) \geq \epsilon_0$. In particular, for every positive integer n one can find an $x_n \in M$ satisfying the pair of inequalities $\rho(x_n, x_0) < 1/n$ and $\rho'(f(x_n), f(x_0)) \geq \epsilon_0$. The first of these inequalities shows, via the 'Sandwich Principle', that $\lim_{n\to\infty} x_n = x_0$ in M, whereas the second guarantees that the sequence $\{f(x_n)\}$ does not converge to $f(x_0)$ in M' as n tends to infinity. Therefore f cannot be continuous at x_0, and we have proved that condition (i) implies condition (ii).

Assume now that condition (ii) holds. Let us also suppose that $\lim_{n\to\infty} x_n = x_0$ in M; we wish to prove that $\lim_{n\to\infty} f(x_n) = f(x_0)$ in M'. Let $\epsilon > 0$ be given. Our assumption provides a $\delta > 0$ such that $\rho'(f(x), f(x_0)) < \epsilon$ whenever $\rho(x, x_0) < \delta$. As $\lim_{n\to\infty} x_n = x_0$, there is some positive integer N (which may depend on δ and hence on ϵ) such that $\rho(x_n, x_0) < \delta$ for every $n \geq N$. It is immediate that $\rho'(f(x_n), f(x_0)) < \epsilon$ for every such n and the proof is complete. ∎

The following (general) definitions will be of considerable use.

Definition 3.7.3. Suppose that X and Y are nonempty sets, and that $f : X \to Y$ is a function. Let $A \subseteq X$ and $B \subseteq Y$. We define

$$f(A) := \{f(x) : x \in A\} \quad \text{and} \quad f^{-1}(B) := \{x \in X : f(x) \in B\}.$$

The first of these sets is called the *image of A* whilst the second is referred to as the *inverse image of B*.

Remark 3.7.4. Let X, Y, f, and B be as above. The reader is asked to verify that the inverse image of the complement of B is the complement of the inverse image of B.

The close connexion between continuity and open/closed sets is highlighted in this next result.

Theorem 3.7.5. *Suppose that (M, ρ) and (M', ρ') are metric spaces, and let $f : M \to M'$ be a function. The following are equivalent.*
(i) f is continuous on M.
(ii) $f^{-1}(G)$ is ρ-open in M whenever G is ρ'-open in M'.
(iii) $f^{-1}(F)$ is ρ-closed in M whenever F is ρ'-closed in M'.

Proof. Assume that f is continuous on M and let G be any open set in M'. If G is either the empty set or all of M', then its inverse image is either empty or all of M; in either case the inverse image of G is open. Suppose now that $f^{-1}(G)$ is nonempty and let x_0 be an arbitrary element in it; the latter condition implies that $f(x_0) \in G$. So the openness of G guarantees an $\epsilon > 0$ such that $U_{\rho'}(f(x_0); \epsilon) \subseteq G$. Now f is assumed to be continuous on M, so it must be continuous at

x_0. Therefore Theorem 3.7.2 supplies a $\delta > 0$ such that $\rho'(f(x), f(x_0)) < \epsilon$ whenever $\rho(x, x_0) < \delta$. In other words, $f(x) \in U_{\rho'}(f(x_0); \epsilon) \subseteq G$ whenever $x \in U_\rho(x_0; \delta)$, and this in turn implies that $U_\rho(x_0; \delta) \subseteq f^{-1}(G)$. Thus $f^{-1}(G)$ is open in M.

We now show that condition (iii) is implied by condition (ii). Let F be a closed set in M', so that F^c (the complement of F) is open in M'. Therefore $[f^{-1}(F)]^c = f^{-1}(F^c)$ (Remark 3.7.4) is open in M by assumption. Equivalently $f^{-1}(F)$ is closed in M.

Finally, we show that the failure of condition (i) implies the failure of condition (iii) (thereby proving that (iii) implies (i) and completing the cycle of implications). If f is not continuous on M then it is not continuous at some x_0 in M. So there is some sequence $\{x_n\}$ in M such that $\lim_{n\to\infty} x_n = x_0$ but $\{f(x_n)\}$ does not converge to $f(x_0)$. The latter condition implies the existence of an $\eta > 0$ and an increasing sequence of positive integers, say $n_1 < n_2 < \cdots$, such that $\rho'(f(x_{n_k}), f(x_0)) \geq \eta$ for every positive integer k. Hence $x_{n_k} \in f^{-1}([U_{\rho'}(f(x_0); \eta)]^c)$ for every $k \in \mathbf{N}$. As $\lim_{k\to\infty} x_{n_k} = x_0$ (every subsequence of a convergent sequence also converges to the limit of the original sequence) and x_0 is patently outside $f^{-1}([U_{\rho'}(f(x_0); \eta)]^c) = [f^{-1}(U_{\rho'}(f(x_0); \eta))]^c$ (because x_0 belongs to $f^{-1}(U_{\rho'}(f(x_0); \eta))$), the set $f^{-1}([U_{\rho'}(f(x_0); \eta)]^c)$ cannot be closed (thanks to Theorem 3.2.9). Thus we stand in violation of condition (iii), because $[U_{\rho'}(f(x_0); \eta)]^c$ is a closed set (via Theorem 3.2.5). ∎

The remainder of this section will be devoted to exploring the action of continuous functions on connected sets and compact sets. In particular, we shall discover that both properties are preserved under continuous maps, an altogether satisfactory state of affairs indeed.

Theorem 3.7.6. *Suppose that (M, ρ) and (M', ρ') are metric spaces, and let $f : M \to M'$ be a function. If f is continuous on M and A is a connected subset of M, then $f(A)$ is a connected subset of M'.*

Proof. If $f(A)$ is disconnected, then there exist ρ'-open sets G and H which provide a disconnection for $f(A)$ in M'. It is a fairly simple matter to check that the sets $f^{-1}(G)$ and $f^{-1}(H)$ – which are ρ-open because of Theorem 3.7.5 – provide a disconnection for A in M. It follows that if A is connected, then so is $f(A)$. ∎

As a direct consequence of the foregoing theorem one obtains a (modest) generalization of the familiar Intermediate Value Theorem.

Corollary 3.7.8. *(Generalized Intermediate Value Theorem) Suppose that (M, ρ) is a connected metric space, and that the real line $\mathbf{R}$ is equipped with the usual metric. Let $f : M \to \mathbf{R}$ be continuous on M. If $a, b \in M$ and γ is a real number between $f(a)$ and $f(b)$, then there is some $c \in M$ such that $f(c) = \gamma$.*

Proof. EXERCISE (Use Theorem 3.7.6 along with Remark 3.6.7). ∎

The reader will also convince herself that the standard Intermediate Value Theorem is, in fact, a consequence of Corollary 3.7.8.

Theorem 3.7.9. *Suppose that (M, ρ) and (M', ρ') are metric spaces, and let $f : M \to M'$ be a function. If f is continuous on M and A is a compact subset of M, then $f(A)$ is a compact subset of M'.*

Proof. On account of Theorem 3.5.4, it suffices to show the following: if $\{b_n\}$ is any sequence in $f(A)$ then there is a subsequence $\{b_{n_k} : k \in \mathbf{N}\}$ of $\{b_n\}$ as well as an element $b \in f(A)$ such that $\lim_{k\to\infty} b_{n_k} = b$. Let $b_n = f(a_n)$, $a_n \in A$, $n \in \mathbf{N}$. The compactness of A yields a subsequence

$\{a_{n_k} : k \in \mathbf{N}\}$ of $\{a_n\}$ as well as an element $a \in A$ such that $\lim_{k\to\infty} a_{n_k} = a$. As f is continuous at a and $\lim_{k\to\infty} a_{n_k} = a$, we must have $\lim_{k\to\infty} f(a_{n_k}) = f(a)$. Choosing $b = f(a)$ finishes the proof. ∎

Specializing to real-valued functions one obtains the following.

Theorem 3.7.10. *Suppose that (M, ρ) is a metric space, and that the real line $\mathbf{R}$ is equipped with the usual metric. Let $f : M \to \mathbf{R}$ be continuous on M, and let A be a compact subset of M. The following hold:*
(i) $f(A)$ is a closed and bounded subset of $\mathbf{R}$.
(ii) Let $\alpha^* := \sup\{f(a) : a \in A\}$ and $\alpha_* := \inf\{f(a) : a \in A\}$. There exist elements a^* and a_* in A such that $f(a^*) = \alpha^*$ and $f(a_*) = \alpha_*$.

Proof. The first assertion follows at once from Theorem 3.7.9 and the Heine-Borel Theorem.

The set $f(A)$ being bounded, it is contained in the interval $[-T, T]$ for some $T > 0$ (Lemma 3.4.7). In particular α^* and α_* are well defined. Let n be a(ny) positive integer. Proposition 2.1.8 provides an $a_n \in A$ such that $\alpha^* - (1/n) < f(a_n) \le \alpha^*$, whence

$$\lim_{n\to\infty} f(a_n) = \alpha^*. \tag{3.7.1}$$

Now the compactness of A yields a subsequence $\{a_{n_k} : k \in \mathbf{N}\}$ of $\{a_n\}$ as well as an element $a^* \in A$ such that $\lim_{k\to\infty} a_{n_k} = a^*$. As f is continuous at a^* and $\lim_{k\to\infty} a_{n_k} = a^*$, we must have

$$\lim_{k\to\infty} f(a_{n_k}) = f(a^*). \tag{3.7.2}$$

From (3.7.1) and (3.7.2) we find that $\alpha^* = f(a^*)$. The argument involving α_* is similar. ∎

We wish to close the section with an important example, but some groundwork must be laid first. Let M be a nonempty set; suppose that f and g are functions from M into the real line $\mathbf{R}$, and that α is a real number. We define the functions $f + g$, fg, αf, and $|f|$ in the expected way: for every $x \in M$,

$$(f+g)(x) := f(x) + g(x), \quad (fg)(x) := f(x)g(x), \quad (\alpha f)(x) := \alpha f(x), \quad \text{and} \quad |f|(x) := |f(x)|.$$

The following result is basic to further development.

Theorem 3.7.11. *Suppose that (M, ρ) is a metric space, and that the real line $\mathbf{R}$ is equipped with the usual metric. Let f and g be functions from M into $\mathbf{R}$, and let x_0 be a fixed element in M. The following hold:*
(i) If f and g are continuous at x_0, then so is $f + g$.
(ii) If f is continuous at x_0 and α is any fixed real number, then αf is also continuous at x_0.
(iii) If f is continuous at x_0, then so is $|f|$.
(iv) If f and g are continuous at x_0, then so is fg.

Proof. EXERCISE. ∎

The following example, which will bring this section to a close, invites an obvious comparison with Example 3.1.5(iv).

Example 3.7.12. Suppose that (K, ρ) is a compact metric space, and let the real line $\mathbf{R}$ be equipped with the usual metric. We define

$$C(K) := \{f : K \to \mathbf{R} \,:\, f \text{ is continuous on } K\}.$$

No doubt the reader will be delighted at being afforded the opportunity to verify the following facts: (i) $C(K)$, when endowed with the operations of addition and scalar multiplication discussed above, is a vector space over $\mathbf{R}$; (ii) $|f| \in C(K)$ whenever $f \in C(K)$, and (iii) if $f, g \in C(K)$, then fg belongs to $C(K)$ as well. The reader will further prolong her pleasure by also proving that the function

$$\|f\|_\infty := \sup\{|f(x)| \,:\, x \in K\}, \qquad f \in C(K),$$

is a norm on $C(K)$ (usually called the *supremum norm* or the *uniform norm*), and that

$$\|fg\|_\infty \le \|f\|_\infty \, \|g\|_\infty, \qquad f, g \in C(K).$$

§3.8. Uniform continuity

Let f be a function from a metric space (M, ρ) into a metric space (M', ρ'). Theorem 3.7.2 teaches us that f is continuous on M if for every $t \in M$ and every $\epsilon > 0$ there is a positive number δ, which may depend on t and ϵ, such that $\rho'(f(x), f(t)) < \epsilon$ whenever $\rho(x, t) < \delta$. We are now going to require something stronger, namely that the δ be independent of t. This leads to the notion of uniform continuity, a precise definition of which is given as under:

Definition 3.8.1. Suppose that (M, ρ) and (M', ρ') are metric spaces, and let $f : M \to M'$ be a function. We say that f is *uniformly continuous* on M if for every $\epsilon > 0$ there is a positive number δ (which may depend on ϵ) such that $\rho'(f(x), f(t)) < \epsilon$ whenever $x, t \in M$ and $\rho(x, t) < \delta$.

The following definition will help generate a fairly wide class of uniformly continuous functions.

Definition 3.8.2. Suppose that (M, ρ) and (M', ρ') are metric spaces. A function $f : M \to M'$ is said to be *Lipschitz* if there is a positive constant K such that $\rho'(f(x), f(t)) \le K\rho(x, t)$ for every pair of elements x and t in M. The constant K is often called the *Lipschitz constant* of f.

Example 3.8.3. Every Lipschitz function is uniformly continuous (choose $\delta := \epsilon/K$).

The Mean Value Theorem, a staple in univariate calculus, provides a convenient tool for manufacturing Lipschitz functions. Before taking this matter up in earnest, we formulate a technical definition. Given an interval I on the real line $\mathbf{R}$, the *interior* of I is defined as follows:

$$I^\circ := \{x \in I \,:\, \text{there is an } r_x > 0 \text{ such that } (x - r_x, x + r_x) \subseteq I\}.$$

As a quick example, the reader will have no difficulty in verifying that

$$(0, 1) = (0, 1)^\circ = [0, 1)^\circ = (0, 1]^\circ = [0, 1]^\circ.$$

The upcoming lemma will serve a useful purpose very soon.

Lemma 3.8.4. *Suppose that I is an interval on the real line $\mathbf{R}$, which is equipped with the usual metric. Let f be a real-valued function which is continuous on I and differentiable on the interior of I. Assume further that the quantity $K := \sup\{|f'(u)| : u \in I^\circ\}$ is finite. Then f is a Lipschitz function with Lipschitz constant K.*

Proof. Let $x, t \in I$. The Mean Value Theorem supplies a c between x and t such that $f(x)-f(t) = f'(c)(x-t)$. It follows at once that $|f(x) - f(t)| \leq K|x - t|$. ∎

Example 3.8.5. (i) The function $f(x) = x^2$ is Lipschitz, hence uniformly continuous, on $[-T, T]$ for every $T > 0$ (and hence uniformly continuous on $[a, b]$ for every $a, b \in \mathbf{R}$). Apply the preceding lemma with $I = [-T, T]$. We see that K may be chosen to be $2T$.
(ii) The function $g(x) = 1/x$ is Lipschitz, hence uniformly continuous, on any interval of the form $[a, \infty)$, $a > 0$; apply Lemma 3.8.4 with $I = [a, \infty)$ and $K = 1/a^2$.

We now wish to give examples of functions that fail to be uniformly continuous; as an (essential) preparation, the reader will find it highly instructive, if not downright enjoyable, to prove the following:

Remark 3.8.6. Suppose that (M, ρ) and (M', ρ') are metric spaces. A function $f : M \to M'$ is *not* uniformly continuous on M if and only if there is some $\epsilon_0 > 0$ and a pair of sequences $\{u_n\}$ and $\{v_n\}$ in M such that $\lim_{n\to\infty}\rho(u_n, v_n) = 0$ and $\rho'(f(u_n), f(v_n)) \geq \epsilon_0$ for every positive integer n.

Example 3.8.7. (i) The function $f(x) = x^2$ is not uniformly continuous on $\mathbf{R}$. This is seen by taking $\epsilon_0 = 2$, $u_n = n$, and $v_n = n + (1/n)$ in the foregoing remark.
(ii) The function $g(x) = 1/x$ is not uniformly continuous on $(0, \infty)$; take $\epsilon_0 = 1/2$, $u_n = 1/n$, and $v_n = 2/n$ in Remark 3.8.6.

It is our intention to conclude this section with a theorem that will demonstrate, in particular, that uniformly continuous functions do exist in profusion. We shall begin, however, with a preliminary result which is of interest in its own right.

Theorem 3.8.8. *Suppose that (M, ρ) is a metric space. If $\{a_n\}$ and $\{b_n\}$ are sequences in M and $\lim_{n\to\infty}a_n = a$ and $\lim_{n\to\infty}b_n = b$, then $\lim_{n\to\infty}\rho(a_n, b_n) = \rho(a, b)$.*

Proof. The triangle inequality implies the following pair of relations for every $n \in \mathbf{N}$:

$$\rho(a_n, b_n) - \rho(a, b) \leq \rho(a_n, a) + \rho(b, b_n)$$

and

$$\rho(a, b) - \rho(a_n, b_n) \leq \rho(a, a_n) + \rho(b_n, b).$$

These relations, taken in conjunction with the symmetry of the distance function ρ, yield the estimate

$$|\rho(a_n, b_n) - \rho(a, b)| \leq \rho(a_n, a) + \rho(b_n, b),$$

and our assumption guarantees that the right-hand side of the preceding inequality converges to zero as n tends to infinity. ∎

We are now in a position to prove the rather striking fact that every continuous function on a compact metric space is, in fact, uniformly continuous.

Theorem 3.8.9. *Suppose that (M,ρ) and (M',ρ') are metric spaces, and that $f : M \to M'$ is a function. If M is compact and f is continuous on M, then f is uniformly continuous on M.*

Proof. Assume contrariwise that f is not uniformly continuous. Remark 3.8.6 then furnishes a positive number ϵ_0 and a pair of sequences $\{u_n\}$ and $\{v_n\}$ in M such that

$$\lim_{n\to\infty}\rho(u_n,v_n)=0 \quad \text{and} \quad \rho'(f(u_n),f(v_n))\geq \epsilon_0 \ \forall\, n\in \mathbf{N}. \tag{3.8.1}$$

As M is compact, Theorem 3.5.4 guarantees a subsequence $\{u_{n_k} : k\in\mathbf{N}\}$ of $\{u_n\}$ and an element $u\in M$ such that $\lim_{k\to\infty}u_{n_k}=u$. Now the triangle inequality yields the estimate

$$\rho(v_{n_k},u)\leq \rho(v_{n_k},u_{n_k})+\rho(u_{n_k},u), \qquad k\in\mathbf{N}, \tag{3.8.2}$$

and the right-hand side of (3.8.2) converges to zero as k tends to infinity, via the first condition in (3.8.1) and the fact that $\lim_{k\to\infty}u_{n_k}=u$; hence $\lim_{k\to\infty}v_{n_k}=u$. Furthermore, f is continuous at u, so $\lim_{k\to\infty}f(u_{n_k})= f(u)= \lim_{k\to\infty}f(v_{n_k})$, whence $\lim_{k\to\infty}\rho'(f(u_{n_k}),f(v_{n_k}))=0$ on account of Theorem 3.8.8 and Property (M2) of a metric. But this clearly violates the second condition in (3.8.1), and the proof is complete. ∎

§3.9. Convergence of functions

This section is devoted to the study of convergence of sequences of functions defined on metric spaces.

Definition 3.9.1. Suppose that (M,ρ) and (M',ρ') are metric spaces. For every positive integer n, let f_n be a function from M to M'. We say that the sequence of functions $\{f_n\}$ *converges to (a function) $f : M \to M'$ pointwise on M* if $\lim_{n\to\infty}f_n(x)=f(x)$ for every $x\in M$.

Remark 3.9.2. The following is a quantitative definition of pointwise convergence given above: $\{f_n\}$ converges to f pointwise on M if for every $x\in M$ and every $\epsilon>0$, there is a positive integer N – which may depend on x and ϵ – such that $\rho'(f_n(x),f(x))<\epsilon$ for every $n\geq N$.

Example 3.9.3. (i) Let $f_n(x)=x^n$, $0\leq x\leq 1$, $n\in\mathbf{N}$. The sequence $\{f_n\}$ converges to the function

$$f(x):=\begin{cases}0, & \text{if } 0\leq x<1;\\ 1, & \text{if } x=1,\end{cases}$$

pointwise on $[0,1]$.
(ii) Consider the following sequence of functions defined on $[0,1]$:

$$f_n(x):=\begin{cases}2nx, & \text{if } 0\leq x\leq \frac{1}{2n};\\[2mm] -2n\left(x-\frac{1}{n}\right), & \text{if } \frac{1}{2n}\leq x\leq \frac{1}{n};\\[2mm] 0, & \text{if } \frac{1}{n}\leq x\leq 1.\end{cases}$$

(The graph of f_n comprises a straight-line segment from $(0,0)$ to $(1/2n,1)$, followed by a line segment from $(1/2n,1)$ to $(1/n,0)$, and a line segment from $(1/n,0)$ to $(1,0)$.) Then $\lim_{n\to\infty}f_n(x)=0$ for every $x\in[0,1]$. Thus the sequence $\{f_n\}$ converges to the zero function pointwise on $[0,1]$.

We now wish to strengthen the notion of pointwise convergence. The upcoming definition must be compared carefully with Remark 3.9.2.

Definition 3.9.4. Suppose that $\{f_n\}$ is a sequence of functions from a metric space (M, ρ) into a metric space (M', ρ'). We say that $\{f_n\}$ *converges to* f *uniformly on* M if for every $\epsilon > 0$ there is a positive integer N – which may depend on ϵ – such that $\rho'(f_n(x), f(x)) < \epsilon$ for every $n \geq N$ and every $x \in M$.

The next result exhibits a very desirable feature of uniform convergence, namely that the uniform limit of continuous functions is also continuous.

Theorem 3.9.5. *Suppose that* $\{f_n\}$ *is a sequence of functions from a metric space* (M, ρ) *into a metric space* (M', ρ'). *Assume that* f_n *is continuous on* M *for every* n. *If* $\{f_n\}$ *converges to* f *uniformly on* M, *then* f *is continuous on* M.

Proof. We shall use Theorem 3.7.2 to show that f is continuous at every $x_0 \in M$. Let $\epsilon > 0$ be given. Uniform convergence supplies a positive integer N such that

$$\rho'(f_N(x), f(x)) < \epsilon/3, \qquad \forall\, x \in M. \tag{3.9.1}$$

The function f_N being continuous at x_0, there is a $\delta > 0$ such that

$$\rho'(f_N(x), f_N(x_0)) < \epsilon/3 \qquad \text{whenever } \rho(x, x_0) < \delta. \tag{3.9.2}$$

So for every x in M such that $\rho(x, x_0) < \delta$, we obtain the following via the triangle inequality, (3.9.1), and (3.9.2):

$$\rho'(f(x), f(x_0)) \leq \rho'(f(x), f_N(x)) + \rho'(f_N(x), f_N(x_0)) + \rho'(f_N(x_0), f(x_0)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

The proof is complete. ∎

Example 3.9.6. An appeal to the preceding theorem shows that the sequence of functions considered in Example 3.9.3(i) is not uniformly convergent.

Remark 3.9.7. Suppose that (M, ρ) and (M', ρ') are metric spaces and $f_n : M \to M'$ for every positive integer n. Assume that the sequence $\{f_n\}$ converges to f pointwise on M. The reader will find it extremely instructive to provide a proof of the following assertion: $\{f_n\}$ does *not* converge to f uniformly on M if and only if there exists a positive number ϵ_0, a subsequence $\{n_k : k \in \mathbf{N}\}$ of the natural numbers, and a sequence $\{x_k\}$ in M such that $\rho'(f_{n_k}(x_k), f(x_k)) \geq \epsilon_0$ for every $k \in \mathbf{N}$.

Example 3.9.8. Consider the sequence $\{f_n\}$ defined in Example 3.9.3(ii), and recall that it converges to the zero function pointwise on $[0, 1]$. Choosing $\epsilon_0 = 1/2$, $n_k = k$, and $x_k = 1/2k$ in Remark 3.9.7, one finds that the convergence is not uniform. It is worth noting that this cannot be deduced via Theorem 3.9.5, for in this case, the limit function, along with every element in the sequence, is eminently continuous.

For the remainder of this section we shall focus exclusively on real-valued functions. We begin with an elegant theorem which provides a sufficient condition for uniform convergence. The alert reader will also find that it offers additional insight into the failure of uniform convergence in the foregoing example.

Theorem 3.9.9. *(Dini) Suppose that (M, ρ) is a compact metric space, and assume that the real line* **R** *is equipped with the usual metric. Let $\{g_n\}$ be sequence of functions from M to* **R** *satisfying the following conditions: (i) each g_n is continuous on M, (ii) for every $x \in M$, the sequence $\{g_n(x)\}$ is nonincreasing, and (iii) $\lim_{n\to\infty} g_n(x) = 0$ for each x in M. Then $\{g_n\}$ converges to the zero function uniformly on M.*

Proof. Suppose that $\{g_n\}$ does not converge uniformly. Remark 3.9.7 supplies an $\epsilon_0 > 0$, a subsequence $\{g_{n_k} : k \in \mathbf{N}\}$ of $\{g_n\}$, and a sequence $\{x_k\}$ in M such that $|g_{n_k}(x_k) - 0| = g_{n_k}(x_k) \geq \epsilon_0$ for every k. By Theorem 3.5.4, $\{x_k\}$ admits a subsequence $\{x_{k_p} : p \in \mathbf{N}\}$ which converges, as p tends to infinity, to an element $x^* \in M$. Let $y_p := x_{k_p}$ and $n_{k_p} := m(p)$, $p \in \mathbf{N}$, and note that $m(p) < m(p+1)$ for every p. Let us recall here that $\lim_{p\to\infty} y_p = x^*$ and

$$g_{m(p)}(y_p) \geq \epsilon_0 \quad \text{for every } p \in \mathbf{N}. \tag{3.9.3}$$

As $\{g_n(x^*) : n \in \mathbf{N}\}$ is a nonincreasing sequence converging to zero, there is a positive integer N such that

$$g_N(x^*) < \epsilon_0/4. \tag{3.9.4}$$

Furthermore, the continuity of g_N at x^* provides a positive number δ such that

$$|g_N(x) - g_N(x^*)| < \epsilon_0/4 \quad \text{whenever } \rho(x, x^*) < \delta. \tag{3.9.5}$$

Choose (and fix) a positive integer p such that $\rho(y_p, x^*) < \delta$ and $m(p) > N$. Remembering that the sequence $\{g_n(y_p) : n \in \mathbf{N}\}$ is nonincreasing, we find via (3.9.3)–(3.9.5) that

$$\begin{aligned}\epsilon_0 \leq g_{m(p)}(y_p) = g_{m(p)}(y_p) - g_N(x^*) + g_N(x^*) &\leq g_N(y_p) - g_N(x^*) + g_N(x^*)\\ &< \epsilon_0/4 + g_N(x^*) < \epsilon_0/4 + \epsilon_0/4 = \epsilon_0/2.\end{aligned}$$

This apparent contradiction finishes the proof. ∎

Corollary 3.9.10. *Suppose that (M, ρ) is a compact metric space, and assume that the real line* **R** *is equipped with the usual metric. Let $\{f_n\}$ be sequence of functions from M to* **R** *satisfying the following conditions: (i) each f_n is continuous on M, (ii) for every $x \in M$, the sequence $\{f_n(x)\}$ is nonincreasing, (iii) $\lim_{n\to\infty} f_n(x) = f(x)$ for each x in M, and (iv) f is continuous on M. Then $\{f_n\}$ converges to f uniformly on M.*

Proof. EXERCISE. ∎

The preceding result remains valid if the second condition there is replaced by the condition that the sequence $\{f_n(x)\}$ is nondecreasing for every fixed x in M. One simply applies the aforesaid corollary to $\{-f_n\}$.

Let (K, ρ) be a compact metric space, and recall the normed linear space $(C(K), \|\cdot\|_\infty)$ from Example 3.7.12. The following simple observation will prove quite useful.

Lemma 3.9.11. *Suppose that $\{f_n\}$ is a sequence in $C(K)$. The following are equivalent:*
(i) $\{f_n\}$ converges to f uniformly on K.
(ii) $\lim\limits_{n\to\infty} \|f_n - f\|_\infty = 0$.

Proof. If (i) holds, then f is continuous on K (Theorem 3.9.5). Let $\epsilon > 0$ be given. Uniform convergence supplies a positive integer N such that

$$|f_n(x) - f(x)| \leq \epsilon, \quad \forall\, n \geq N, \quad \forall\, x \in K, \tag{3.9.6}$$

whence

$$\|f_n - f\|_\infty = \sup\{|f_n(x) - f(x)| \,:\, x \in K\} \leq \epsilon, \quad \forall\, n \geq N, \tag{3.9.7}$$

and this proves (ii). The reverse implication follows at once upon observing that conditions (3.9.6) and (3.9.7) are, in fact, equivalent. ∎

The preceding lemma helps provide further information about the space $C(K)$.

Example 3.9.12. Let (K, ρ) be a compact metric space. We show that the normed linear space $(C(K), \|\cdot\|_\infty)$ is complete. Let $\{f_n\}$ be a Cauchy sequence in $(C(K), \|\cdot\|_\infty)$. Given $\epsilon > 0$ there is some positive integer P such that $\|f_m - f_n\|_\infty < \epsilon$ for every $m, n \geq P$. Let $x \in K$ be fixed. Then

$$|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty < \epsilon, \quad \forall\, m, n \geq P, \tag{3.9.8}$$

so $\{f_n(x)\}$ is a Cauchy sequence of real numbers, hence convergent. Let $f(x) := \lim_{n\to\infty} f_n(x)$, $x \in K$. Keeping m fixed in (3.9.8) and letting n there tend to infinity yields the estimate

$$|f_m(x) - f(x)| \leq \epsilon, \quad \forall\, m \geq P, \quad \forall\, x \in K.$$

In other words $\{f_n\}$ converges to f uniformly on K, and (the first part of) Lemma 3.9.11 ensures that $f \in C(K)$ and that $\lim_{n\to\infty}\|f_n - f\|_\infty = 0$. Thus every Cauchy sequence in $(C(K), \|\cdot\|_\infty)$ converges to an element in that space, *i.e.,* $(C(K), \|\cdot\|_\infty)$ is complete.

A fitting finale to the chapter will be provided by a celebrated theorem due to Karl Weierstraß. This seminal result paved the way towards the creation of a branch of analysis known as Approximation Theory, which has since blossomed into a fertile field of fruitful activity. The proof of Weierstraß's theorem we intend to present here is a delectable argument due to Sergei Bernstein. Justly famous in its own right, Bernstein's proof – which appeared about a quarter of a century after Weierstraß's original paper – also had a strong impact on the development of the theory of approximation.*

The proceedings will commence with a pivotal definition.

Definition 3.9.13. Suppose that f is a function defined on the interval $[0, 1]$, and let n be a nonnegative integer. The *n-th Bernstein polynomial associated to f* is defined as follows:

$$B_n(f; x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}, \qquad x \in \mathbf{R}.$$

Some simple, yet crucial properties of Bernstein polynomials are recorded in the result below.

Lemma 3.9.14. *Let $f_j(x) := x^j$, $j = 0, 1, 2$. The following hold for every positive integer n and every real number x:*
(i) $B_n(f_0; x) = f_0(x)$.
(ii) $B_n(f_1; x) = f_1(x)$.

---

* The original articles by Weierstraß and Bernstein, along with some other classic papers in the annals of approximation theory, may be found here:
`http://www.cs.wisc.edu/∼deboor/HAT/papers.html`

(iii) $B_n(f_2;x) = f_2(x) + \frac{x(1-x)}{n}$.

Proof. If x and y are real numbers, then the Binomial Theorem asserts that

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}. \tag{3.9.9}$$

Putting $y = 1-x$ in this equation gives (i). Differentiating both sides of (3.9.9) partially with respect to x, and multiplying the resulting equation by x/n, one obtains

$$x(x+y)^{n-1} = \sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} x^k y^{n-k}, \tag{3.9.10}$$

and substituting $y = 1-x$ in the foregoing equation yields

$$f_1(x) = x = \sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} x^k (1-x)^{n-k} = B_n(f_1;x).$$

Thus (ii) is proven. Now we differentiate (3.9.10) partially with respect to y and multiply the resulting equation by x/n to get

$$\frac{(n-1)x^2(x+y)^{n-2}}{n} + \frac{x(x+y)^{n-1}}{n} = \sum_{k=0}^{n} \binom{n}{k} \frac{k^2}{n^2} x^k y^{n-k}.$$

Setting $y = 1-x$ in the preceding equation gives the relation

$$\frac{(n-1)x^2 + x}{n} = \sum_{k=0}^{n} \binom{n}{k} \frac{k^2}{n^2} x^k (1-x)^{n-k} = B_n(f_2;x),$$

and a simplification of the left-hand side of the last equation results in (iii). ∎

We are now ready for the centrepiece of the current discussion.

Theorem 3.9.15. *(Weierstraß) Given (any) $f \in C([0,1])$, there is a sequence of polynomials which converges to f uniformly on $[0,1]$.*

Proof. (*Bernstein*) We show that $\lim_{n\to\infty} B_n(f;x) = f(x)$ for every $0 \le x \le 1$, and that the convergence is uniform on $[0,1]$. The function f being continuous on the compact set $[0,1]$, there is a positive number T such that $|f(x)| \le T$ for every $x \in [0,1]$ (via Theorem 3.7.10(i) and Lemma 3.4.7). Let $\epsilon > 0$ be given. Now Theorem 3.8.9 guarantees that f is uniformly continuous on $[0,1]$, so there is a $\delta > 0$ such that $|f(u) - f(v)| < \epsilon/2$ for every pair of points u and v in $[0,1]$ such that $|u - v| < \delta$.

Suppose now that $0 \le x \le 1$, let n be a positive integer, and let $p_{n,k}(x) := \binom{n}{k} x^k (1-x)^{n-k}$, $0 \le k \le n$. Using Lemma 3.9.14(i) (in the first step below) along with the triangle inequality and the patent nonnegativity of $p_{n,k}(x)$, we obtain

$$\begin{aligned} |B_n(f;x) - f(x)| &= \left| \sum_{k=0}^{n} f\left(\frac{k}{n}\right) p_{n,k}(x) - \sum_{k=0}^{n} f(x) p_{n,k}(x) \right| = \left| \sum_{k=0}^{n} \left[f\left(\frac{k}{n}\right) - f(x) \right] p_{n,k}(x) \right| \\ &\le \sum_{k=0}^{n} \left| f\left(\frac{k}{n}\right) - f(x) \right| p_{n,k}(x). \end{aligned} \tag{3.9.11}$$

Setting $I := \{k : 0 \le k \le n,\ |(k/n) - x| < \delta\}$ and $I' := \{k : 0 \le k \le n,\ |(k/n) - x| \ge \delta\}$, we may split the last sum in (3.9.11) into two:

$$\begin{aligned} |B_n(f;x) - f(x)| &\le \sum_{k \in I} \left| f\left(\frac{k}{n}\right) - f(x) \right| p_{n,k}(x) + \sum_{k \in I'} \left| f\left(\frac{k}{n}\right) - f(x) \right| p_{n,k}(x) \\ &=: S_1 + S_2. \end{aligned} \tag{3.9.12}$$

If $|(k/n) - x| < \delta$, then $|f(k/n) - f(x)| < \epsilon/2$, so the nonnegativity of the summands $p_{n,k}(x)$ lead to the estimates

$$S_1 \le \frac{\epsilon}{2} \sum_{k \in I} p_{n,k}(x) \le \frac{\epsilon}{2} \sum_{k=0}^{n} p_{n,k}(x) \le \frac{\epsilon}{2}, \tag{3.9.13}$$

the final inequality coming from Lemma 3.9.14(i).

The second sum S_2 will be handled using a lovely trick. Let $f_j(x) = x^j$, $j = 0, 1, 2$. Observing that $1 \le ((k/n) - x)^2/\delta^2$ for every $k \in I'$, we find, via the triangle inequality, the definition of the number T, and the nonnegativity of the functions $p_{n,k}$, that

$$\begin{aligned} S_2 &\le \sum_{k \in I'} \left[\left| f\left(\frac{k}{n}\right) \right| + |f(x)| \right] p_{n,k}(x) \le \frac{2T}{\delta^2} \sum_{k \in I'} \left(\frac{k}{n} - x \right)^2 p_{n,k}(x) \\ &\le \frac{2T}{\delta^2} \sum_{k=0}^{n} \left[\frac{k^2}{n^2} - 2x\frac{k}{n} + x^2 \right] p_{n,k}(x) \\ &= \frac{2T}{\delta^2} \left[B_n(f_2;x) - 2xB_n(f_1;x) + x^2 B_n(f_0;x) \right]. \end{aligned}$$

An appeal to Lemma 3.9.14 reveals that the foregoing estimate can be recast as follows:

$$S_2 \le \frac{2T}{\delta^2} \left[x^2 + \frac{x(1-x)}{n} - 2x^2 + x^2 \right] = \frac{2T}{\delta^2} \frac{x(1-x)}{n} \le \frac{2T}{n\delta^2}. \tag{3.9.14}$$

Choose a positive integer N so that $2T/(n\delta^2) < \epsilon/2$ for every $n \ge N$, and observe that such a choice depends only on f and ϵ. Combining (3.9.12)–(3.9.14) we conclude that

$$|B_n(f;x) - f(x)| < \epsilon, \quad \forall\, n \ge N, \quad \forall\, 0 \le x \le 1,$$

and this finishes the proof. ∎

A simple reformulation of Weierstraß's theorem is given below.

Corollary 3.9.16. *Given $f \in C([0,1])$ and $\epsilon > 0$, there is a polynomial p, which may depend on f and ϵ, such that $|f(x) - p(x)| < \epsilon$ for every $0 \le x \le 1$.*

Proof. Choose $p(x) = B_n(f;x)$ for a suitably large n. ∎

The specific choice of the interval $[0,1]$ in Theorem 3.9.15 and Corollary 3.9.16 is primarily one of convenience. In fact, the general form of Weierstraß's theorem may be deduced from the special case presented in the foregoing corollary. This is the content of the upcoming result, which will also serve to bring the curtain down on the present chapter.

Theorem 3.9.17. *(Weierstraß's Polynomial Approximation Theorem) Let a and b be real numbers with $a < b$. Given $F \in C([a, b])$ and $\epsilon > 0$, there is a polynomial P (which may depend on F and ϵ) such that $|F(t) - P(t)| < \epsilon$ for every $a \leq t \leq b$.*

Proof. Define $f(x) := F(a + x(b - a))$, $0 \leq x \leq 1$. Then $f \in C([0, 1])$, and Corollary 3.9.16 provides a polynomial p such that

$$|F(a + x(b - a)) - p(x)| = |f(x) - p(x)| < \epsilon \qquad \forall\, 0 \leq x \leq 1. \tag{3.9.15}$$

Putting $t := a + x(b - a)$ and noting that $P(t) := p((t - a)/(b - a))$ is also a polynomial (in the variable t), we find that (3.9.15) may be rewritten as follows:

$$|F(t) - P(t)| < \epsilon \qquad \forall\, a \leq t \leq b.$$

∎