Exercises

1. Give the details in the derivation of (6.7) and (6.8).
2. For each of the following points in $\mathbb{C}$, give the corresponding point of S: 0, $1+i$, $3+2i$.
3. Which subsets of S correspond to the real and imaginary axes in $\mathbb{C}$.
4. Let Λ be a circle lying in S. Then there is a unique plane P in $\mathbb{R}^3$ such that $P \cap S = \Lambda$. Recall from analytic geometry that

$$P = \{(x_1, x_2, x_3):\ x_1\beta_1 + x_2\beta_2 + x_3\beta_3 = l\}$$

where $(\beta_1, \beta_2, \beta_3)$ is a vector orthogonal to P and l is some real number. It can be assumed that $\beta_1^2+\beta_2^2+\beta_3^2 = 1$. Use this information to show that if Λ contains the point N then its projection on $\mathbb{C}$ is a straight line. Otherwise, Λ projects onto a circle in C.
5. Let Z and Z' be points on S corresponding to z and z' respectively. Let W be the point on S corresponding to $z+z'$. Find the coordinates of W in terms of the coordinates of Z and Z'.

Chapter II

Metric Spaces and the Topology of $\mathbb{C}$

§*1. Definition and examples of metric spaces*

A *metric space* is a pair (X, d) where X is a set and d is a function from $X \times X$ into $\mathbb{R}$, called a *distance function* or *metric*, which satisfies the following conditions for x, y, and z in X:

$$d(x, y) \geq 0$$

$$d(x, y) = 0 \text{ if and only if } x = y$$

$$d(x, y) = d(y, x) \text{ (symmetry)}$$

$$d(x, z) \leq d(x, y) + d(y, z) \text{ (triangle inequality)}$$

If x and $r > 0$ are fixed then define

$$B(x; r) = \{y \in X:\ d(x, y) < r\}$$

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

$B(x;\ r)$ and $\bar{B}(x;\ r)$ are called the *open* and *closed balls*, respectively, with center x and radius r.

Examples

1.1 Let $X = \mathbb{R}$ or $\mathbb{C}$ and define $d(z, w) = |z-w|$. This makes both $(\mathbb{R}, d)$ and $(\mathbb{C}, d)$ metric spaces. In fact, $(\mathbb{C}, d)$ will be the example of principal interest to us. If the reader has never encountered the concept of a metric space before this, he should continually keep $(\mathbb{C}, d)$ in mind during the study of this chapter.
1.2 Let (X, d) be a metric space and let $Y \subset X$; then (Y, d) is also a metric space.
1.3 Let $X = \mathbb{C}$ and define $d(x+iy, a+ib) = |x-a|+|y-b|$. Then $(\mathbb{C}, d)$ is a metric space.
1.4 Let $X = \mathbb{C}$ and define $d(x+iy, a+ib) = \max\{|x-a|, |y-b|\}$.
1.5 Let X be any set and define $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ if $x \neq y$. To show that the function d satisfies the triangle inequality one merely considers all possibilities of equality among x, y, and z. Notice here that $B(x;\ \epsilon)$ consists only of the point x if $\epsilon \leq 1$ and $B(x;\ \epsilon) = X$ if $\epsilon > 1$. This metric space does not appear in the study of analytic function theory.
1.6 Let $X = \mathbb{R}^n$ and for $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ define

$$d(x, y) = \left[\sum_{j=1}^{n} (x_j - y_j)^2\right]^{\frac{1}{2}}$$

1.7 Let S be any set and denote by $B(S)$ the set of all functions $f\colon S \to \mathbb{C}$ such that

$$\|f\|_\infty \equiv \sup\{|f(s)| : s \in S\} < \infty.$$

That is, $B(S)$ consists of all complex valued functions whose range is contained inside some disk of finite radius. For f and g in $B(S)$ define $d(f, g) = \|f-g\|_\infty$. We will show that d satisfies the triangle inequality. In fact if f, g, and h are in $B(S)$ and s is any point in S then $|f(s)-g(s)| = |f(s)-h(s)+h(s)-g(s)| \leq |f(s)-h(s)|+|h(s)-g(s)| \leq \|f-h\|_\infty + \|h-g\|_\infty$. Thus, when the supremum is taken over all s in S, $\|f-g\|_\infty \leq \|f-h\|_\infty + \|h-g\|_\infty$, which is the triangle inequality for d.

1.8 Definition. For a metric space (X, d) a set $G \subset X$ is *open* if for each x in G there is an $\epsilon > 0$ such that $B(x; \epsilon) \subset G$.

Thus, a set in $\mathbb{C}$ is open if it has no "edge." For example, $G = \{z \in \mathbb{C}: a < \operatorname{Re} z < b\}$ is open; but $\{z: \operatorname{Re} z < 0\} \cup \{0\}$ is not because $B(0; \epsilon)$ is not contained in this set no matter how small we choose ϵ.

We denote the *empty set*, the set consisting of no elements, by $\square$.

1.9 Proposition. *Let (X, d) be a metric space; then:*

(a) *The sets X and $\square$ are open;*

(b) *If $G_1, \ldots, G_n$ are open sets in X then so is $\bigcap_{k=1}^{n} G_k$;*

(c) *If $\{G_j: j \in J\}$ is a collection of open sets in X, J any indexing set, then $G = \cup\{G_j: j \in J\}$ is also open.*

Proof. The proof of (a) is a triviality. To prove (b) let $x \in G = \bigcap_{k=1}^{n} G_k$; then $x \in G_k$ for $k = 1, \ldots, n$. Thus, by the definition, for each k there is an $\epsilon_k > 0$ such that $B(x; \epsilon_k) \subset G_k$. But if $\epsilon = \min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$ then for $1 \leq k \leq n$ $B(x; \epsilon) \subset B(x; \epsilon_k) \subset G_k$. Thus $B(x; \epsilon) \subset G$ and G is open.

The proof of (c) is left as an exercise for the reader. ■

There is another class of subsets of a metric space which are distinguished. These are the sets which contain all their "edge"; alternately, the sets whose complements have no "edge."

1.10 Definition. A set $F \subset X$ is *closed* if its complement, $X-F$, is open.

The following proposition is the complement of Proposition 1.9. The proof, whose execution is left to the reader, is accomplished by applying de Morgan's laws to the preceding proposition.

1.11 Proposition. *Let (X, d) be a metric space. Then:*

(a) *The sets X and $\square$ are closed;*

(b) *If $F_1, \ldots, F_n$ are closed sets in X then so is $\bigcup_{k=1}^{n} F_k$;*

(c) *If $\{F_j: j \in J\}$ is any collection of closed sets in X, J any indexing set, then $F = \cap\{F_j: j \in J\}$ is also closed.*

The most common error made upon learning of open and closed sets is to interpret the definition of closed set to mean that if a set is not open it is closed. This, of course, is false as can be seen by looking at $\{z \in \mathbb{C}: \operatorname{Re} z > 0\} \cup \{0\}$; it is neither open nor closed.

1.12 Definition. Let A be a subset of X. Then the *interior* of A, int A, is the set $\bigcup\{G: G \text{ is open and } G \subset A\}$. The *closure* of A, A^-, is the set $\bigcap\{F: F \text{ is closed and } F \supset A\}$. Notice that int A may be empty and A^- may be X. If $A = \{a+bi: a \text{ and } b \text{ are rational numbers}\}$ then simultaneously $A^- = \mathbb{C}$ and int $A = \square$. By Propositions 1.9 and 1.11 we have that A^- is closed and int A is open. The *boundary* of A is denoted by ∂A and defined by $\partial A = A^- \cap (X-A)^-$.

1.13 Proposition. *Let A and B be subsets of a metric space (X, d). Then:*

(a) *A is open if and only if $A = \operatorname{int} A$;*

(b) *A is closed if and only if $A = A^-$;*

(c) $\operatorname{int} A = X-(X-A)^-$; $A^- = X-\operatorname{int}(X-A)$; $\partial A = A^- - \operatorname{int} A$;

(d) $(A \cup B)^- = A^- \cup B^-$;

(e) *$x_0 \in \operatorname{int} A$ if and only if there is an $\epsilon > 0$ such that $B(x_0; \epsilon) \subset A$;*

(f) *$x_0 \in A^-$ if and only if for every $\epsilon > 0$, $B(x_0; \epsilon) \cap A \neq \square$.*

Proof. The proofs of (a)–(e) are left to the reader. To prove (f) assume $x_0 \in A^- = X-\operatorname{int}(X-A)$; thus, $x_0 \notin \operatorname{int}(X-A)$. By part (e), for every $\epsilon > 0$ $B(x_0; \epsilon)$ is not contained in $X-A$. That is, there is a point $y \in B(x_0; \epsilon)$ which is not in $X-A$. Hence, $y \in B(x_0; \epsilon) \cap A$. Now suppose $x_0 \notin A^- = X-\operatorname{int}(X-A)$. Then $x_0 \in \operatorname{int}(X-A)$ and, by (e), there is an $\epsilon > 0$ such that $B(x_0; \epsilon) \subset X-A$. That is, $B(x_0; \epsilon) \cap A = \square$ so that x_0 does not satisfy the condition. ■

Finally, one last definition of a distinguished type of set.

1.14 Definition. A subset A of a metric space X is *dense* if $A^- = X$.

The set of rational numbers $\mathbb{Q}$ is dense in $\mathbb{R}$ and $\{x+iy: x, y \in \mathbb{Q}\}$ is dense in $\mathbb{C}$.

Exercises

1. Show that each of the examples of metric spaces given in (1.2)–(1.6) is, indeed, a metric space. Example (1.6) is the only one likely to give any difficulty. Also, describe $B(x; r)$ for each of these examples.
2. Which of the following subsets of $\mathbb{C}$ are open and which are closed: (a) $\{z: |z| < 1\}$; (b) the real axis; (c) $\{z: z^n = 1 \text{ for some integer } n \geq 1\}$; (d) $\{z \in \mathbb{C}: z \text{ is real and } 0 \leq z < 1\}$; (e) $\{z \in \mathbb{C}: z \text{ is real and } 0 \leq z \leq 1\}$?
3. If (X, d) is any metric space show that every open ball is, in fact, an open set. Also, show that every closed ball is a closed set.
4. Give the details of the proof of (1.9c).
5. Prove Proposition 1.11.
6. Prove that a set $G \subset X$ is open if and only if $X-G$ is closed.
7. Show that $(\mathbb{C}_\infty, d)$ where d is given by (I. 6.7) and (I. 6.8) is a metric space.
8. Let (X, d) be a metric space and $Y \subset X$. Suppose $G \subset X$ is open; show

that $G \cap Y$ is open in (Y, d). Conversely, show that if $G_1 \subset Y$ is open in (Y, d), there is an open set $G \subset X$ such that $G_1 = G \cap Y$.
9. Do Exercise 8 with "closed" in place of "open."
10. Prove Proposition 1.13.
11. Show that $\{\text{cis}\, k : k \geq 0\}$ is dense in $T = \{z \in \mathbb{C} : |z| = 1\}$. For which values of θ is $\{\text{cis}(k\theta) : k \geq 0\}$ dense in T?

§2. *Connectedness*

Let us start this section by giving an example. Let $X = \{z \in \mathbb{C}: |z| \leq 1\} \cup \{z: |z-3| < 1\}$ and give X the metric it inherits from $\mathbb{C}$. (Henceforward, whenever we consider subsets X of $\mathbb{R}$ or $\mathbb{C}$ as metric spaces we will assume, unless stated to the contrary, that X has the inherited metric $d(z, w) = |z-w|$.) Then the set $A = \{z: |z| \leq 1\}$ is simultaneously open and closed. It is closed because its complement in X, $B = X - A = \{z: |z-3| < 1\}$ is open; A is open because if $a \in A$ then $B(a; 1) \subset A$. (Notice that it may not happen that $\{z \in \mathbb{C}: |z-a| < 1\}$ is contained in A—for example, if $a = 1$. But the definition of $B(a; 1)$ is $\{z \in X: |z-a| < 1\}$ and this is contained in A.) Similarly B is also both open and closed in X.

This is an example of a non-connected space.

2.1 Definition. A metric space (X, d) is *connected* if the only subsets of X which are both open and closed are $\square$ and X. If $A \subset X$ then A is a *connected subset of* X if the metric space (A, d) is connected.

An equivalent formulation of connectedness is to say that X is not connected if there are disjoint open sets A and B in X, neither of which is empty, such that $X = A \cup B$. In fact, if this condition holds then $A = X - B$ is also closed.

2.2 Proposition. *A set $X \subset \mathbb{R}$ is connected iff X is an interval.*

Proof. Suppose $X = [a, b]$, a and b elements of $\mathbb{R}$. Let $A \subset X$ be an open subset of X such that $a \in A$, and $A \neq X$. We will show that A cannot also be closed—and hence, X must be connected. Since A is open and $a \in A$ there is an $\epsilon > 0$ such that $[a, a+\epsilon) \subset A$. Let

$$r = \sup \{\epsilon: [a, a+\epsilon) \subset A\}$$

Claim. $[a, a+r) \subset A$. In fact, if $a \leq x < a+r$ then, putting $h = a+r-x > 0$, the definition of supremum implies there is an ϵ with $r-h < \epsilon < r$ and $[a, a+\epsilon) \subset A$. But $a \leq x = a+(r-h) < a+\epsilon$ implies $x \in A$ and the claim is established.

However, $a+r \notin A$; for if, on the contrary, $a+r \in A$ then, by the openness of A, there is a $\delta > 0$ with $[a+r, a+r+\delta) \subset A$. But this gives $[a, a+r+\delta) \subset A$, contradicting the definition of r. Now if A were also closed then $a+r \in B = X-A$ which is open. Hence we could find a $\delta > 0$ such that $(a+r-\delta, a+r] \subset B$, contradicting the above claim.

The proof that other types of intervals are connected is similar and it will be left as an exercise.

The proof of the converse is Exercise 1. ■

If w and z are in $\mathbb{C}$ then we denote the straight line segment from z to w by

$$[z, w] = \{tw+(1-t)z: 0 \leq t \leq 1\}$$

A *polygon* from a to b is a set $P = \bigcup_{k=1}^{n} [z_k, w_k]$ where $z_1 = a$, $w_n = b$ and $w_k = z_{k+1}$ for $1 \leq k \leq n-1$; or, $P = [a, z_1 \ldots z_n, b]$.

2.3 Theorem. *An open set $G \subset \mathbb{C}$ is connected iff for any two points a, b in G there is a polygon from a to b lying entirely inside G.*

Proof. Suppose that G satisfies this condition and let us assume that G is not connected. We will obtain a contradiction. From the definition, $G = A \cup B$ where A and B are both open and closed, $A \cap B = \square$, and neither A nor B is empty. Let $a \in A$ and $b \in B$; by hypothesis there is a polygon P from a to b such that $P \subset G$. Now a moment's thought will show that one of the segments making up P will have one point in A and another in B. So we can assume that $P = [a, b]$. Define,

$$S = \{s \in [0, 1]: sb+(1-s)a \in A\}$$

$$T = \{t \in [0, 1]: tb+(1-t)a \in B\}$$

Then $S \cap T = \square$, $S \cup T = [0, 1]$, $0 \in S$ and $1 \in T$. However it can be shown that both S and T are open (Exercise 2), contradicting the connectedness of $[0, 1]$. Thus, G must be connected.

Now suppose that G is connected and fix a point a in G. To show how to construct a polygon (lying in G!) from a to a point b in G would be difficult. But we don't have to perform such a construction; we merely show that one exists. For a fixed a in G define

$$A = \{b \in G: \text{there is a polygon } P \subset G \text{ from } a \text{ to } b\}.$$

The plan is to show that A is simultaneously open and closed in G. Since $a \in A$ and G is connected this will give that $A = G$ and the theorem will be proved.

To show that A is open let $b \in A$ and let $P = [a, z_1, \ldots, z_n, b]$ be a polygon from a to b with $P \subset G$. Since G is open (this was not needed in the first half), there is an $\epsilon > 0$ such that $B(b; \epsilon) \subset G$. But if $z \in B(b; \epsilon)$ then $[b, z] \subset B(b; \epsilon) \subset G$. Hence the polygon $Q = P \cup [b, z]$ is inside G and goes from a to z. This shows that $B(b; \epsilon) \subset A$, and so A is open.

To show that A is closed suppose there is a point z in $G-A$ and let $\epsilon > 0$ be such that $B(z; \epsilon) \subset G$. If there is a point b in $A \cap B(z; \epsilon)$ then, as above, we can construct a polygon from a to z. Thus we must have that $B(z; \epsilon) \cap A = \square$, or $B(z; \epsilon) \subset G-A$. That is, $G-A$ is open so that A is closed. ■

2.4 Corollary. *If $G \subset \mathbb{C}$ is open and connected and a and b are points in G then there is a polygon P in G from a to b which is made up of line segments parallel to either the real or imaginary axis.*

Proof. There are two ways of proving this corollary. One could obtain a polygon in G from a to b and then modify each of its line segments so that a new polygon is obtained with the desired properties. However, this proof is more easily executed using compactness (see Exercise 5.7). Another proof can be obtained by modifying the proof of Theorem 2.3. Define the set A as in the proof of (2.3) but add the restriction that the polygon's segments are all parallel to one of the axes. The remainder of the proof will be valid with one exception. If $z \in B(b; \epsilon)$ then $[b, z]$ may not be parallel to an axis. But it is easy to see that if $z = x+iy$, $b = p+iq$ then the polygon $[b, p+iy] \cup [p+iy, z] \subset B(b; \epsilon)$ and has segments parallel to an axis. ■

It will now be shown that any set S in a metric space can be expressed, in a canonical way, as the union of connected pieces.

2.5 Definition. A subset D of a metric space X is a *component* of X if it is a maximal connected subset of X. That is, D is connected and there is no connected subset of X that properly contains D.

If the reader examines the example at the beginning of this section he will notice that both A and B are components and, furthermore, these are the only components of X. For another example let $X = \{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\}$. Then clearly every component of X is a point and each point is a component. Notice that while the components $\left\{\frac{1}{n}\right\}$ are all open in X, the component $\{0\}$ is not.

2.6 Lemma. *Let $x_0 \in X$ and let $\{D_j : j \in J\}$ be a collection of connected subsets of X such that $x_0 \in D_j$ for each j in J. Then $D = \bigcup \{D_j : j \in J\}$ is connected.*

Proof. Let A be a subset of the metric space (D, d) which is both open and closed and suppose that $A \neq \square$. Then $A \cap D_j$ is open in (D_j, d) for each j and it is also closed (Exercises 1.8 and 1.9). Since D_j is connected we get that either $A \cap D_j = \square$ or $A \cap D_j = D_j$. Since $A \neq \square$ there is at least one k such that $A \cap D_k \neq \square$; hence, $A \cap D_k = D_k$. In particular $x_0 \in A$ so that $x_0 \in A \cap D_j$ for every j. Thus $A \cap D_j = D_j$, or $D_j \subset A$, for each index j. This gives that $D = A$, so that D is connected. ■

2.7 Theorem. *Let (X, d) be a metric space. Then:*

(a) *Each x_0 in X is contained in a component of X.*

(b) *Distinct components of X are disjoint.*

Note that part (a) says that X is the union of its components.

Proof. (a) Let $\mathscr{D}$ be the collection of connected subsets of X which contain the given point x_0. Notice that $\{x_0\} \in \mathscr{D}$ so that $\mathscr{D} \neq \square$. Also notice that the hypotheses of the preceding lemma apply to the collection $\mathscr{D}$. Hence $C = \bigcup \{D : D \subset \mathscr{D}\}$ is connected and $x_0 \in C$. But C must be a component. In fact, if D is connected and $C \subset D$ then $x_0 \in D$ so that $D \in \mathscr{D}$; but then $D \subset C$, so that $C = D$. Thus C is maximal and part (a) is proved.

(b) Suppose C_1 and C_2 are components, $C_1 \neq C_2$, and suppose there is a point x_0 in $C_1 \cap C_2$. Again the lemma says that $C_1 \cup C_2$ is connected.

Since both C_1 and C_2 are components, this gives $C_1 = C_1 \cup C_2 = C_2$, a contradiction. ■

2.8 Proposition. (a) *If $A \subset X$ is connected and $A \subset B \subset A^-$, then B is connected.*

(b) *If C is a component of X then C is closed.*

The proof is left as an exercise.

2.9 Theorem. *Let G be open in $\mathbb{C}$; then the components of G are open and there are only a countable number of them.*

Proof. Let C be a component of G and let $x_0 \in C$. Since G is open there is an $\epsilon > 0$ with $B(x_0; \epsilon) \subset G$. By Lemma 2.6, $B(x_0; \epsilon) \cup C$ is connected and so must be C. That is $B(x_0; \epsilon) \subset C$ and C is, therefore, open.

To see that the number of components is countable let $S = \{a+ib$: a and b are rational and $a+bi \in G\}$. Then S is countable and each component of G contains a point of S, so that the number of components is countable. ■

Exercises

1. The purpose of this exercise is to show that a connected subset of $\mathbb{R}$ is an interval.
(a) Show that a set $A \subset R$ is an interval iff for any two points a and b in A with $a < b$, the interval $[a, b] \subset A$.
(b) Use part (a) to show that if a set $A \subset R$ is connected then it is an interval.
2. Show that the sets S and T in the proof of Theorem 2.3 are open.
3. Which of the following subsets X of $\mathbb{C}$ are connected; if X is not connected, what are its components: (a) $X = \{z: |z| \leq 1\} \cup \{z: |z-2| < 1\}$. (b) $X = [0, 1) \cup \left\{1+\frac{1}{n} : n \geq 1\right\}$. (c) $X = \mathbb{C} - (A \cup B)$ where $A = [0, \infty)$ and $B = \{z = r \text{ cis } \theta : r = \theta, 0 \leq \theta \leq \infty\}$?
4. Prove the following generalization of Lemma 2.6. If $\{D_j : j \in J\}$ is a collection of connected subsets of X and if for each j and k in J we have $D_j \cap D_k \neq \square$ then $D = \bigcup \{D_j : j \in J\}$ is connected.
5. Show that if $F \subset X$ is closed and connected then for every pair of points a, b in F and each $\epsilon > 0$ there are points $z_0, z_1, \ldots, z_n$ in F with $z_0 = a$, $z_n = b$ and $d(z_{k-1}, z_k) < \epsilon$ for $1 \leq k \leq n$. Is the hypothesis that F be closed needed? If F is a set which satisfies this property then F is not necessarily connected, even if F is closed. Give an example to illustrate this.

§3. *Sequences and completeness*

One of the most useful concepts in a metric space is that of a convergent sequence. Their central role in calculus is duplicated in the study of metric spaces and complex analysis.

3.1 Definition. If $\{x_1, x_2, \ldots\}$ is a sequence in a metric space (X, d) then

$\{x_n\}$ *converges* to x—in symbols $x = \lim x_n$ or $x_n \to x$—if for every $\epsilon > 0$ there is an integer N such that $d(x, x_n) < \epsilon$ whenever $n \geq N$.

Alternately, $x = \lim x_n$ if $0 = \lim d(x, x_n)$.

If $X = \mathbb{C}$ then $z = \lim z_n$ means that for each $\epsilon > 0$ there is an N such that $|z - z_n| < \epsilon$ when $n \geq N$.

Many concepts in the theory of metric spaces can be phrased in terms of sequences. The following is an example.

3.2 Proposition. *A set $F \subset X$ is closed iff for each sequence $\{x_n\}$ in F with $x = \lim x_n$ we have $x \in F$.*

Proof. Suppose F is closed and $x = \lim x_n$ where each x_n is in F. So for every $\epsilon > 0$, there is a point x_n in $B(x; \epsilon)$; that is $B(x; \epsilon) \cap F \neq \square$, so that $x \in F^- = F$ by Proposition 2.8.

Now suppose F is not closed; so there is a point x_0 in F^- which is not in F. By Proposition 1.13(f), for every $\epsilon > 0$ we have $B(x_0; \epsilon) \cap F \neq \square$. In particular for every integer n there is a point x_n in $B\left(x_0; \frac{1}{n}\right) \cap F$. Thus, $d(x_0, x_n) < \frac{1}{n}$ which implies that $x_n \to x_0$. Since $x_0 \notin F$, this says the condition fails. ■

3.3 Definition. If $A \subset X$ then a point x in X is a *limit point* of A if there is a sequence $\{x_n\}$ of distinct points in A such that $x = \lim x_n$.

The reason for the word "distinct" in this definition can be illustrated by the following example. Let $X = \mathbb{C}$ and let $A = [0, 1] \cup \{i\}$; each point in $[0, 1]$ is a limit point of A but i is not. We do not wish to call a point such as i a limit point; but if "distinct" were dropped from the definition we could taken $x_n = i$ for each i and have $i = \lim x_n$.

3.4 Proposition. (a) *A set is closed iff it contains all its limit points.*

(b) *If $A \subset X$ then $A^- = A \cup \{x\colon x$ is a limit point of $A\}$.*

The proof is left as an exercise.

From real analysis we know that a basic property of $\mathbb{R}$ is that any sequence whose terms get closer together as n gets large, must be convergent. Such sequences are called Cauchy sequences. One of their attributes is that you know the limit will exist even though you can't produce it.

3.5 Definition. A sequence $\{x_n\}$ is called a *Cauchy sequence* if for every $\epsilon > 0$ there is an integer N such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq N$.

If (X, d) has the property that each Cauchy sequence has a limit in X then (X, d) is *complete.*

3.6 Proposition. $\mathbb{C}$ *is complete.*

Proof. If $\{x_n + iy_n\}$ is a Cauchy sequence in $\mathbb{C}$ then $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in $\mathbb{R}$. Since $\mathbb{R}$ is complete, $x_n \to x$ and $y_n \to y$ for points x, y in $\mathbb{R}$. It follows that $x + iy = \lim (x_n + iy_n)$, and so $\mathbb{C}$ is complete. ■

Consider $\mathbb{C}_\infty$ with its metric d (I. 6.7 and I. 6.8). Let z_n, z be points in $\mathbb{C}$;

it can be shown that $d(z_n, z) \to 0$ if and only if $|z_n - z| \to 0$. In spite of this, any sequence $\{z_n\}$ with $\lim |z_n| = \infty$ is Cauchy in $\mathbb{C}_\infty$, but, of course, is not Cauchy in $\mathbb{C}$.

If $A \subset X$ we define the *diameter* of A by $\operatorname{diam} A = \sup \{d(x, y)\colon x$ and y are in $A\}$.

3.7 Cantor's Theorem. *A metric space (X, d) is complete iff for any sequence $\{F_n\}$ of non-empty closed sets with $F_1 \supset F_2 \supset \ldots$ and* $\operatorname{diam} F_n \to 0$, $\bigcap_{n=1}^{\infty} F_n$ *consists of a single point.*

Proof. Suppose (X, d) is complete and let $\{F_n\}$ be a sequence of closed sets having the properties: (i) $F_1 \supset F_2 \supset \ldots$ and (ii) $\lim \operatorname{diam} F_n = 0$. For each n, let x_n be an arbitrary point in F_n; if $n, m \geq N$ then x_n, x_m are in F_N so that, by definition, $d(x_n, x_m) \leq \operatorname{diam} F_N$. By the hypothesis N can be chosen sufficiently large that $\operatorname{diam} F_N < \epsilon$; this shows that $\{x_n\}$ is a Cauchy sequence. Since X is complete, $x_0 = \lim x_n$ exists. Also, x_n is in F_N for all $n \geq N$ since $F_n \subset F_N$; hence, x_0 is in F_N for every N and this gives $x_0 \in \bigcap_{n=1}^{\infty} F_n = F$. So F contains at least one point; if, also, y is in F then both x_0 and y are in F_n for each n and this gives $d(x_0, y) \leq \operatorname{diam} F_n \to 0$. Therefore $d(x_0, y) = 0$, or $x_0 = y$.

Now let us show that X is complete if it satisfies the stated condition. Let $\{x_n\}$ be a Cauchy sequence in X and put $F_n = \{x_n, x_{n+1}, \ldots\}^-$; then $F_1 \supset F_2 \supset \ldots$. If $\epsilon > 0$, choose N such that $d(x_n, x_m) < \epsilon$ for each n, $m \geq N$; this gives that $\operatorname{diam} \{x_n, x_{n+1}, \ldots\} \leq \epsilon$ for $n \geq N$ and so $\operatorname{diam} F_n \leq \epsilon$ for $n \geq N$ (Exercise 3). Thus $\operatorname{diam} F_n \to 0$ and, by hypothesis, there is a point x_0 in X with $\{x_0\} = F_1 \cap F_2 \cap \ldots$. In particular x_0 is in F_n, and so $d(x_0, x_n) \leq \operatorname{diam} F_n \to 0$. Therefore, $x_0 = \lim x_n$. ■

There is a standard exercise associated with this theorem. It is to find a sequence of sets $\{F_n\}$ in $\mathbb{R}$ which satisfies two of the conditions:

(a) each F_n is closed,

(b) $F_1 \supset F_2 \supset \ldots$,

(c) $\operatorname{diam} F_n \to 0$;

but which has $F = F_1 \cap F_2 \cap \ldots$ either empty or consisting of more than one point. Everyone should get examples satisfying the possible combinations.

3.8 Proposition. *Let (X, d) be a complete metric space and let $Y \subset X$. Then (Y, d) is a complete metric space iff Y is closed in X.*

Proof. It is left as an exercise to show that (Y, d) is complete whenever Y is a closed subset. Now assume (Y, d) to be complete; let x_0 be a limit point of Y. Then there is a sequence $\{y_n\}$ of points in Y such that $x_0 = \lim y_n$. Hence $\{y_n\}$ is a Cauchy sequence (Exercise 5) and must converge to a point y_0 in Y, since (Y, d) is complete. It follows that $y_0 = x_0$ and so Y contains all its limit points. Hence Y is closed by Proposition 3.4. ■

Exercises

1. Prove Proposition 3.4.
2. Furnish the details of the proof of Proposition 3.8.
3. Show that diam A = diam A^-.
4. Let z_n, z be points in $\mathbb{C}$ and let d be the metric on $\mathbb{C}_\infty$. Show that $|z_n - z| \to 0$ if and only if $d(z_n, z) \to 0$. Also show that if $|z_n| \to \infty$ then $\{z_n\}$ is Cauchy in $\mathbb{C}_\infty$. (Must $\{z_n\}$ converge in $\mathbb{C}_\infty$?)
5. Show that every convergent sequence in (X, d) is a Cauchy sequence.
6. Give three examples of non complete metric spaces.
7. Put a metric d on $\mathbb{R}$ such that $|x_n - x| \to 0$ if and only if $d(x_n, x) \to 0$, but that $\{x_n\}$ is a Cauchy sequence in $(\mathbb{R}, d)$ when $|x_n| \to \infty$. (Hint: Take inspiration from $\mathbb{C}_\infty$.)
8. Suppose $\{x_n\}$ is a Cauchy sequence and $\{x_{n_k}\}$ is a subsequence that is convergent. Show that $\{x_n\}$ must be convergent.

§4. *Compactness*

The concept of compactness is an extension of the benefits of finiteness to infinite sets. Most properties of compact sets are analogues of properties of finite sets which are quite trivial. For example, every sequence in a finite set has a convergent subsequence. This is quite trivial since there must be at least one point which is repeated an infinite number of times. However the same statement remains true if "finite" is replaced by "compact."

4.1 Definition. A subset K of a metric space X is *compact* if for every collection $\mathscr{G}$ of open sets in X with the property

4.2 $$K \subset \bigcup \{G : G \in \mathscr{G}\},$$

there is a finite number of sets $G_1, \ldots, G_n$ in $\mathscr{G}$ such that $K \subset G_1 \cup G_2 \cup \ldots \cup G_n$. A collection of sets $\mathscr{G}$ satisfying (4.2) is called a *cover* of K; if each member of $\mathscr{G}$ is an open set it is called an *open cover* of K.

Clearly the empty set and all finite sets are compact. An example of a non compact set is $D = \{z \in \mathbb{C} : |z| < 1\}$. If $G_n = \left\{z : |z| < 1 - \frac{1}{n}\right\}$ for $n = 2, 3, \ldots$, then $\{G_2, G_3, \ldots\}$ is an open cover of D for which there is no finite subcover.

4.3 Proposition. *Let K be a compact subset of X; then:*
(a) *K is closed;*
(b) *If F is closed and $F \subset K$ then F is compact.*

Proof. To prove part (a) we will show that $K = K^-$. Let $x_0 \in K^-$; by Proposition 1.13(f), $B(x_0; \epsilon) \cap K \neq \square$ for each $\epsilon > 0$. Let $G_n = X - \bar{B}\left(x_0; \frac{1}{n}\right)$ and suppose that $x_0 \notin K$. Then each G_n is open and $K \subset \bigcup_{n=1}^{\infty} G_n$ (because $\bigcap_{n=1}^{\infty}$

$\bar{B}\left(x_0; \frac{1}{n}\right) = \{x_0\}$). Since K is compact there is an integer m such that $K \subset \bigcup_{n=1}^{m} G_n$. But $G_1 \subset G_2 \subset \ldots$ so that $K \subset G_m = X - \bar{B}\left(x_0; \frac{1}{m}\right)$. But this gives that $B\left(x_0; \frac{1}{m}\right) \cap K = \square$, a contradiction. Thus $K = K^-$.

To prove part (b) let $\mathscr{G}$ be an open cover of F. Then, since F is closed, $\mathscr{G} \cup \{X - F\}$ is an open cover of K. Let $G_1, \ldots, G_n$ be sets in $\mathscr{G}$ such that $K \subset G_1 \cup \ldots \cup G_n \cup (X - F)$. Clearly, $F \subset G_1 \cup \ldots \cup G_n$ and so F is compact. ■

If $\mathscr{F}$ is a collection of subsets of X we say that $\mathscr{F}$ has the *finite intersection property* (f.i.p.) if whenever $\{F_1, F_2, \ldots, F_n\} \subset \mathscr{F}$, $F_1 \cap F_2 \cap \ldots \cap F_n \neq \square$. An example of such a collection is $\{D - G_2, D - G_3, \ldots\}$ where the sets G_n are as in the example preceding Proposition 4.3.

4.4 Proposition. *A set $K \subset X$ is compact iff every collection $\mathscr{F}$ of closed subsets of K with the f.i.p. has $\bigcap \{F : F \in \mathscr{F}\} \neq \square$.*

Proof. Suppose K is compact and $\mathscr{F}$ is a collection of closed subsets of K having the f.i.p. Assume that $\bigcap \{F : F \in \mathscr{F}\} = \square$ and let $\mathscr{G} = \{X - F : F \in \mathscr{F}\}$. Then, $\bigcup \{X - F : F \in \mathscr{F}\} = X - \bigcap \{F : F \in \mathscr{F}\} = X$ by the assumption; in particular, $\mathscr{G}$ is an open cover of K. Thus, there are $F_1, \ldots, F_n \in \mathscr{F}$ such that $K \subset \bigcup_{k=1}^{n} (X - F_k) = X - \bigcap_{k=1}^{n} F_k$. But this gives that $\bigcap_{k=1}^{n} F_k \subset X - K$, and since each F_k is a subset of K it must be that $\bigcap_{k=1}^{n} F_k = \square$. This contradicts the f.i.p.

The proof of the converse is left as an exercise. ■

4.5 Corollary. *Every compact metric space is complete.*

Proof. This follows easily by applying the above proposition and Theorem 3.7. ■

4.6. Corollary. *If X is compact then every infinite set has a limit point in X.*

Proof. Let S be an infinite subset of X and suppose S has no limit points. Let $\{a_1, a_2, \ldots\}$ be a sequence of distinct points in S; then $F_n = \{a_n, a_{n+1}, \ldots\}$ also has no limit points. But if a set has no limit points it contains all its limits points and must be closed! Thus, each F_n is closed and $\{F_n : n \geq 1\}$ has the f.i.p. However, since the points $a_1, a_2, \ldots$ are distinct, $\bigcap_{n=1}^{\infty} F_n = \square$, contradicting the above proposition. ■

4.7 Definition. A metric space (X, d) is *sequentially compact* if every sequence in X has a convergent subsequence.

It will be shown that compact and sequentially compact metric spaces are the same. To do this the following is needed.

4.8 Lebesgue's Covering Lemma. *If (X, d) is sequentially compact and*

$\mathscr{G}$ is an open cover of X then there is an $\epsilon > 0$ such that if x is in X, there is a set G in $\mathscr{G}$ with $B(x; \epsilon) \subset G$.

Proof. The proof is by contradiction; suppose that $\mathscr{G}$ is an open cover of X and no such $\epsilon > 0$ can be found. In particular, for every integer n there is a point x_n in X such that $B\left(x_n; \frac{1}{n}\right)$ is not contained in any set G in $\mathscr{G}$. Since X is sequentially compact there is a point x_0 in X and a subsequence $\{x_{n_k}\}$ such that $x_0 = \lim x_{n_k}$. Let $G_0 \in \mathscr{G}$ such that $x_0 \in G_0$ and choose $\epsilon > 0$ such that $B(x_0; \epsilon) \subset G_0$. Now let N be such that $d(x_0, x_{n_k}) < \epsilon/2$ for all $n_k \geq N$. Let n_k be any integer larger than both N and $2/\epsilon$, and let $y \in B(x_{n_k}; 1/n_k)$. Then $d(x_0, y) \leq d(x_0, x_{n_k}) + d(x_{n_k}, y) < \epsilon/2 + 1/n_k < \epsilon$. That is, $B(x_{n_k}; 1/n_k) \subset B(x_0; \epsilon) \subset G_0$, contradicting the choice of x_{n_k}. ■

There are two common misinterpretations of Lebesgue's Covering Lemma; one implies that it says nothing and the other that it says too much. Since $\mathscr{G}$ is an open covering of X it follows that each x in X is contained in some G in $\mathscr{G}$. Thus there is an $\epsilon > 0$ such that $B(x; \epsilon) \subset G$ since G is open. The lemma, however, gives one $\epsilon > 0$ such that for any x, $B(x; \epsilon)$ is contained in some member of $\mathscr{G}$. The other misinterpretation is to believe that for the $\epsilon > 0$ obtained in the lemma, $B(x; \epsilon)$ is contained in each G in $\mathscr{G}$ such that $x \in G$.

4.9. Theorem. *Let (X, d) be a metric space; then the following are equivalent statements:*

(a) *X is compact;*
(b) *Every infinite set in X has a limit point;*
(c) *X is sequentially compact;*
(d) *X is complete and for every $\epsilon > 0$ there are a finite number of points $x_1, \ldots, x_n$ in X such that*

$$X = \bigcup_{k=1}^{n} B(x_k; \epsilon).$$

(The property mentioned in (d) is called *total boundedness.*)

Proof. That (a) implies (b) is the statement of Corollary 4.6.

(b) implies (c): Let $\{x_n\}$ be a sequence in X and suppose, without loss of generality, that the points $x_1, x_2, \ldots$ are all distinct. By (b), the set $\{x_1, x_2, \ldots\}$ has a limit point x_0. Thus there is a point $x_{n_1} \in B(x_0; 1)$; similarly, there is an integer $n_2 > n_1$ with $x_{n_2} \in B(x_0; 1/2)$. Continuing we get integers $n_1 < n_2 < \ldots$, with $x_{n_k} \in B(x_0; 1/k)$. Thus, $x_0 = \lim x_{n_k}$ and X is sequentially compact.

(c) implies (d): To see that X is complete let $\{x_n\}$ be a Cauchy sequence, apply the definition of sequential compactness, and appeal to Exercise 3.8.

Now let $\epsilon > 0$ and fix $x_1 \in X$. If $X = B(x_1; \epsilon)$ then we are done; otherwise choose $x_2 \in X - B(x_1; \epsilon)$. Again, if $X = B(x_1; \epsilon) \cup B(x_2; \epsilon)$ we are done;

if not, let $x_3 \in X - [B(x_1; \epsilon) \cup B(x_2; \epsilon)]$. If this process never stops we find a sequence $\{x_n\}$ such that

$$x_{n+1} \in X - \bigcup_{k=1}^{n} B(x_k; \epsilon).$$

But this implies that for $n \neq m$, $d(x_n, x_m) \geq \epsilon > 0$. Thus $\{x_n\}$ can have no convergent subsequence, contradicting (c).

(d) implies (c): This part of the proof will use a variation of the "pigeon hole principle." This principle states that if you have more objects than you have receptacles then at least one receptacle must hold more than one object. Moreover, if you have an infinite number of points contained in a finite number of balls then one ball contains infinitely many points. So part (d) says that for every $\epsilon > 0$ and any infinite set in X, there is a point $y \in X$ such that $B(y; \epsilon)$ contains infinitely many points of this set. Let $\{x_n\}$ be a sequence of distinct points. There is a point y_1 in X and a subsequence $\{x_n^{(1)}\}$ of $\{x_n\}$ such that $\{x_n^{(1)}\} \subset B(y_1; 1)$. Also, there is a point y_2 in X and a subsequence $\{x_n^{(2)}\}$ of $\{x_n^{(1)}\}$ such that $\{x_n^{(2)}\} \subset B(y_2; \frac{1}{2})$. Continuing, for each integer $k \geq 2$ there is a point y_k in X and a subsequence $\{x_n^{(k)}\}$ of $\{x_n^{(k-1)}\}$ such that $\{x_n^{(k)}\} \subset B(y_k; 1/k)$. Let $F_k = \{x_n^{(k)}\}^-$; then diam $F_k \leq 2/k$ and $F_1 \supset F_2 \supset \ldots$. By Theorem $\cancel{3.6}$ 3.7, $\bigcap_{k=1}^{\infty} F_k = \{x_0\}$. We claim that $x_k^{(k)} \to x_0$ (and $\{x_k^{(k)}\}$ is a subsequence of $\{x_n\}$). In fact, $x_0 \in F_k$ so that $d(x_0, x_k^{(k)}) \leq$ diam $F_k \leq 2/k$, and $x_0 = \lim x_k^{(k)}$.

(c) implies (a): Let $\mathscr{G}$ be an open cover of X. The preceding lemma gives an $\epsilon > 0$ such that for every $x \in X$ there is a G in $\mathscr{G}$ with $B(x; \epsilon) \subset G$. Now (c) also implies (d); hence there are points $x_1, \ldots, x_n$ in X such that $X = \bigcup_{k=1}^{n} B(x_k; \epsilon)$. Now for $1 \leq k \leq n$ there is a set $G_k \in \mathscr{G}$ with $B(x_k; \epsilon) \subset G_k$. Hence $X = \bigcup_{k=1}^{n} G_k$; that is, $\{G_1, \ldots, G_n\}$ is a finite subcover of $\mathscr{G}$. ■

4.10 Heine-Borel Theorem. *A subset K of $\mathbb{R}^n$ $(n \geq 1)$ is compact iff K is closed and bounded.*

Proof. If K is compact then K is totally bounded by part (d) of the preceding theorem. It follows that K must be closed (Proposition 4.3); also, it is easy to show that a totally bounded set is also bounded.

Now suppose that K is closed and bounded. Hence there are real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ such that $K \subset F = [a_1, b_1] \times \ldots \times [a_n, b_n]$. If it can be shown that F is compact then, because K is closed, it follows that K is compact (Proposition 4.3(b)). Since $\mathbb{R}^n$ is complete and F is closed it follows that F is complete. Hence, again using part (d) of the preceding theorem we need only show that F is totally bounded. This is easy although somewhat "messy" to write down. Let $\epsilon > 0$; we now will write F as the union of n-dimensional rectangles each of diameter less than ϵ. After doing this we will have $F = \bigcup_{k=1}^{m} B(x_k; \epsilon)$ where each x_k belongs to one of the aforementioned rectangles. The execution of the details of this strategy is left to the reader (Exercise 3). ■

Exercises

1. Finish the proof of Proposition 4.4.
2. Let $p = (p_1, \ldots, p_n)$ and $q = (q_1, \ldots, q_n)$ be points in $\mathbb{R}^n$ with $p_k < q_k$ for each k. Let $R = [p_1, q_1]\times\ldots\times[p_n, q_n]$ and show that

$$\operatorname{diam} R = d(p, q) = \left[\sum_{k=1}^{n} (q_k - p_k)^2\right]^{\frac{1}{2}}.$$

3. Let $F = [a_1, b_1]\times\ldots\times[a_n, b_n] \subset \mathbb{R}^n$ and let $\epsilon > 0$; use Exercise 2 to show that there are rectangles $R_1, \ldots, R_m$ such that $F = \bigcup_{k=1}^{m} R_k$ and diam $R_k < \epsilon$ for each k. If $x_k \in R_k$ then it follows that $R_k \subset B(x_k; \epsilon)$.
4. Show that the union of a finite number of compact sets is compact.
5. Let X be the set of all bounded sequences of complex numbers. That is, $\{x_n\} \in X$ iff $\sup\{|x_n|: n \ge 1\} < \infty$. If $x = \{x_n\}$ and $y = \{y_n\}$, define $d(x, y) = \sup\{|x_n - y_n|: n \ge 1\}$. Show that for each x in X and $\epsilon > 0$, $\bar{B}(x; \epsilon)$ is not totally bounded although it is complete. (Hint: you might have an easier time of it if you first show that you can assume $x = (0, 0, \ldots)$.)
6. Show that the closure of a totally bounded set is totally bounded.

§5. *Continuity*

One of the most elementary properties of a function is continuity. The presence of continuity guarantees a certain degree of regularity and smoothness without which it is difficult to obtain any theory of functions on a metric space. Since the main subject of this book is the theory of functions of a complex variable which possess derivatives (and so are continuous), the study of continuity is basic.

5.1 Definition. Let (X, d) and (Ω, ρ) be metric spaces and let $f: X \to \Omega$ be a function. If $a \in X$ and $\omega \in \Omega$, then $\lim_{x\to a} f(x) = \omega$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $\rho(f(x), \omega) < \epsilon$ whenever $0 < d(x, a) < \delta$. The function f is *continuous at the point* a if $\lim_{x\to a} f(x) = f(a)$. If f is continuous at each point of X then f *is a continuous function from* X to Ω.

5.2 Proposition. *Let $f: (X, d) \to (\Omega, \rho)$ be a function and $a \in X$, $\alpha = f(a)$. The following are equivalent statements:*

(a) *f is continuous at a;*
(b) *For every $\epsilon > 0$, $f^{-1}(B(\alpha; \epsilon))$ contains a ball with center at a;*
(c) *$\alpha = \lim f(x_n)$ whenever $a = \lim x_n$.*

The proof will be left as an exercise for the reader.

That was the last proposition concerning continuity of a function at a

point. From now on we will concern ourselves only with functions continuous on all of X.

5.3 Proposition. *Let $f: (X, d) \to (\Omega, \rho)$ be a function. The following are equivalent statements:*

(a) *f is continuous;*
(b) *If Δ is open in Ω then $f^{-1}(\Delta)$ is open in X;*
(c) *If Γ is closed in Ω then $f^{-1}(\Gamma)$ is closed in X.*

Proof. (a) implies (b): Let Δ be open in Ω and let $x \in f^{-1}(\Delta)$. If $\omega = f(x)$ then ω is in Δ; by definition, there is an $\epsilon > 0$ with $B(\omega; \epsilon) \subset \Delta$. Since f is continuous, part (b) of the preceding proposition gives a $\delta > 0$ with $B(x; \delta) \subset f^{-1}(B(\omega; \epsilon)) \subset f^{-1}(\Delta)$. Hence, $f^{-1}(\Delta)$ is open.

(b) implies (c): If $\Gamma \subset \Omega$ is closed then let $\Delta = \Omega - \Gamma$. By (b), $f^{-1}(\Delta) = X - f^{-1}(\Gamma)$ is open, so that $f^{-1}(\Gamma)$ is closed.

(c) implies (a): Suppose there is a point x in X at which f is not continuous. Then there is an $\epsilon > 0$ and a sequence $\{x_n\}$ such that $\rho(f(x_n), f(x)) \ge \epsilon$ for every n while $x = \lim x_n$. Let $\Gamma = \Omega - B(f(x); \epsilon)$; then Γ is closed and each x_n is in $f^{-1}(\Gamma)$. Since (by (c)) $f^{-1}(\Gamma)$ is closed we have $x \in f^{-1}(\Gamma)$. But this implies $\rho(f(x), f(x)) \ge \epsilon > 0$, a contradiction. ■

The following type of result is probably well understood by the reader and so the proof is left as an exercise.

5.4 Proposition. *Let f and g be continuous functions from X into $\mathbb{C}$ and let $\alpha, \beta \in \mathbb{C}$. Then $\alpha f + \beta g$ and fg are both continuous. Also, f/g is continuous provided $g(x) \ne 0$ for every x in X.*

5.5 Proposition. *Let $f: X \to Y$ and $g: Y \to Z$ be continuous functions. Then $g \circ f$ (where $g \circ f(x) = g(f(x))$) is a continuous function from X into Z.*

Proof. If U is open in Z then $g^{-1}(U)$ is open in Y; hence, $f^{-1}(g^{-1}(U)) = (g \circ f)^{-1}(U)$ is open in X. ■

5.6 Definition. A function $f: (X, d) \to (\Omega, \rho)$ is *uniformly continuous* if for every $\epsilon > 0$ there is a $\delta > 0$ (depending only on ϵ) such that $\rho(f(x), f(y)) < \epsilon$ whenever $d(x, y) < \delta$. We say that f is a *Lipschitz function* if there is a constant $M > 0$ such that $\rho(f(x), f(y)) \le M d(x, y)$ for all x and y in X.

It is easy to see that every Lipschitz function is uniformly continuous. In fact, if ϵ is given, take $\delta = \epsilon/M$. It is even easier to see that every uniformly continuous function is continuous. What are some examples of such functions? If $X = \Omega = \mathbb{R}$ then $f(x) = x^2$ is continuous but not uniformly continuous. If $X = \Omega = [0, 1]$ then $f(x) = x^{\frac{1}{2}}$ is uniformly continuous but is not a Lipschitz function. The following provides a wealthy supply of Lipschitz functions.

Let $A \subset X$ and $x \in X$; define the *distance from x to the set A, $d(x, A)$, by*

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

5.7 Proposition. *Let $A \subset X$; then:*

(a) $d(x, A) = d(x, A^-)$;

(b) $d(x, A) = 0$ *iff* $x \in A^-$;

(c) $|d(x, A) - d(y, A)| \leq d(x, y)$ *for all* x, y in X.

Proof. (a) If $A \subset B$ then it is clear from the definition that $d(x, B) \leq d(x, A)$. Hence, $d(x, A^-) \leq d(x, A)$. On the other hand, if $\epsilon > 0$ there is a point y in A^- such that $d(x, A^-) \geq d(x, y) - \epsilon/2$. Also, there is a point a in A with $d(y, a) < \epsilon/2$. But $|d(x, y) - d(x, a)| \leq d(y, a) < \epsilon/2$ by the triangle inequality. In particular, $d(x, y) > d(x, a) - \epsilon/2$. This gives, $d(x, A^-) \geq d(x, a) - \epsilon \geq d(x, A) - \epsilon$. Since ϵ was arbitrary $d(x, A^-) \geq d(x, A)$, so that (a) is proved.

(b) If $x \in A^-$ then $0 = d(x, A^-) = d(x, A)$. Now for any x in X there is a minimizing sequence $\{a_n\}$ in A such that $d(x, A) = \lim d(x, a_n)$. So if $d(x, A) = 0$, $\lim d(x, a_n) = 0$; that is, $x = \lim a_n$ and so $x \in A^-$.

(c) For a in A $d(x, a) \leq d(x, y) + d(y, a)$. Hence, $d(x, A) = \inf \{d(x, a): a \in A\} \leq \inf \{d(x, y) + d(y, a): a \in A\} = d(x, y) + d(y, A)$. This gives $d(x, A) - d(y, A) \leq d(x, y)$. Similarly $d(y, A) - d(x, A) \leq d(x, y)$ so the desired inequality follows. ∎

Notice that part (c) of the proposition says that $f: X \to \mathbb{R}$ defined by $f(x) = d(x, A)$ is a Lipschitz function. If we vary the set A we get a large supply of these functions.

It is not true that the product of two uniformly continuous (Lipschitz) functions is again uniformly continuous (Lipschitz). For example, $f(x) = x$ is Lipschitz but $f \cdot f$ is not even uniformly continuous. However if both f and g are bounded then the conclusion is valid (see Exercise 3).

Two of the most important properties of continuous functions are contained in the following result.

5.8 Theorem. *Let* $f: (X, d) \to (\Omega, \rho)$ *be a continuous function.*

(a) *If* X *is compact then* $f(X)$ *is a compact subset of* Ω.

(b) *If* X *is connected then* $f(X)$ *is a connected subset of* Ω.

Proof. To prove (a) and (b) it may be supposed, without loss of generality, that $f(X) = \Omega$. (a) Let $\{\omega_n\}$ be a sequence in Ω; then there is, for each $n \geq 1$, a point x_n in X with $\omega_n = f(x_n)$. Since X is compact there is a point x in X and a subsequence $\{x_{n_k}\}$ such that $x = \lim x_{n_k}$. But if $\omega = f(x)$, then the continuity of f gives that $\omega = \lim \omega_{n_k}$; hence Ω is compact by Theorem 4.9. (b) Suppose $\Sigma \subset \Omega$ is both open and closed in Ω and that $\Sigma \neq \square$. Then, because $f(X) = \Omega$, $\square \neq f^{-1}(\Sigma)$; also, $f^{-1}(\Sigma)$ is both open and closed because f is continuous. By connectivity, $f^{-1}(\Sigma) = X$ and this gives $\Omega = \Sigma$. Thus, Ω is connected. ∎

5.9 Corollary. *If* $f: X \to \Omega$ *is continuous and* $K \subset X$ *is either compact or connected in* X *then* $f(K)$ *is compact or connected, respectively, in* Ω.

5.10 Corollary. *If* $f: X \to R$ *is continuous and* X *is connected then* $f(X)$ *is an interval.*

This follows from the characterization of connected subsets of $\mathbb{R}$ as intervals.

5.11 Intermediate Value Theorem. *If* $f: [a, b] \to \mathbb{R}$ *is continuous and* $f(a) \leq \xi \leq f(b)$ *then there is a point* x, $a \leq x \leq b$, *with* $f(x) = \xi$.

5.12 Corollary. *If* $f: X \to \mathbb{R}$ *is continuous and* $K \subset X$ *is compact then there are points* x_0 *and* y_0 *in* K *with* $f(x_0) = \sup \{f(x): x \in K\}$ *and* $f(y_0) = \inf \{f(x): x \in K\}$.

Proof. If $\alpha = \sup \{f(x): x \in K\}$ then α is in $f(K)$ because $f(K)$ is closed and bounded in $\mathbb{R}$. Similarly $\beta = \inf \{f(x): x \in K\}$ is in $f(K)$. ∎

5.13 Corollary. *If* $K \subset X$ *is compact and* $f: X \to \mathbb{C}$ *is continuous then there are points* x_0 *and* y_0 *in* K *with*

$$|f(x_0)| = \sup \{|f(x)|: x \in K\} \text{ and } |f(y_0)| = \inf \{|f(x)|: x \in K\}.$$

Proof. This corollary follows from the preceding one because $g(x) = |f(x)|$ defines a continuous function from X into $\mathbb{R}$.

5.14 Corollary. *If* K *is a compact subset of* X *and* x *is in* X *then there is a point* y *in* K *with* $d(x, y) = d(x, K)$.

Proof. Define $f: X \to \mathbb{R}$ by $f(y) = d(x, y)$. Then f is continuous and, by Corollary 5.12, assumes a minimum value on K. That is, there is a point y in K with $f(y) \leq f(z)$ for every $z \in K$. This gives $d(x, y) = d(x, K)$. ∎

The next two theorems are extremely important and will be used repeatedly throughout this book with no specific reference to the theorem numbers.

5.15. Theorem. *Suppose* $f: X \to \Omega$ *is continuous and* X *is compact; then* f *is uniformly continuous.*

Proof. Let $\epsilon > 0$; we wish to find a $\delta > 0$ such that $d(x, y) < \delta$ implies $\rho(f(x), f(y)) < \epsilon$. Suppose there is no such δ; in particular, each $\delta = 1/n$ will fail to work. Then for every $n \geq 1$ there are points x_n and y_n in X with $d(x_n, y_n) < 1/n$ but $\rho(f(x_n), f(y_n)) \geq \epsilon$. Since X is compact there is a subsequence $\{x_{n_k}\}$ and a point $x \in X$ with $x = \lim x_{n_k}$.

Claim. $x = \lim y_{n_k}$. In fact, $d(x, y_{n_k}) \leq d(x, x_{n_k}) + 1/n_k$ and this tends to zero as k goes to ∞.

But if $\omega = f(x)$, $\omega = \lim f(x_{n_k}) = \lim f(y_{n_k})$ so that

$$\begin{aligned} \epsilon &\leq \rho(f(x_{n_k}), f(y_{n_k})) \\ &\leq \rho(f(x_{n_k}), \omega) + \rho(\omega, f(y_{n_k})) \end{aligned}$$

and the right hand side of this inequality goes to zero. This is a contradiction and completes the proof. ∎

5.16. Definition. If A and B are subsets of X then define the *distance from* A *to* B, $d(A, B)$, by

$$d(A, B) = \inf \{d(a, b): a \in A, b \in B\}.$$

Notice that if B is the single-point set $\{x\}$ then $d(A, \{x\}) = d(x, A)$. If

$A = \{y\}$ and $B = \{x\}$ then $d(\{x\}, \{y\}) = d(x, y)$. Also, if $A \cap B \neq \square$ then $d(A, B) = 0$, but we can have $d(A, B) = 0$ with A and B disjoint. The most popular type of example is to take $A = \{(x, 0)\colon x \in \mathbb{R}\} \subset \mathbb{R}^2$ and $B = \{(x, e^x)\colon x \in \mathbb{R}\}$. Notice that A and B are both closed and disjoint and still $d(A, B) = 0$.

5.17 Theorem. *If A and B are disjoint sets in X with B closed and A compact then $d(A, B) > 0$.*

Proof. Define $f\colon X \to \mathbb{R}$ by $f(x) = d(x, B)$. Since $A \cap B = \square$ and B is closed, $f(a) > 0$ for each a in A. But since A is compact there is a point a in A such that $0 < f(a) = \inf\{f(x)\colon x \in A\} = d(A, B)$. ■

Exercises

1. Prove Proposition 5.2.
2. Show that if f and g are uniformly continuous (Lipschitz) functions from X into $\mathbb{C}$ then so is $f+g$.
3. We say that $f\colon X \to \mathbb{C}$ is bounded if there is a constant $M > 0$ with $|f(x)| \leq M$ for all x in X. Show that if f and g are bounded uniformly continuous (Lipschitz) functions from X into $\mathbb{C}$ then so is fg.
4. Is the composition of two uniformly continuous (Lipschitz) functions again uniformly continuous (Lipschitz)?
5. Suppose $f\colon X \to \Omega$ is uniformly continuous; show that if $\{x_n\}$ is a Cauchy sequence in X then $\{f(x_n)\}$ is a Cauchy sequence in Ω. Is this still true if we only assume that f is continuous? (Prove or give a counterexample.)
6. Recall the definition of a dense set (1.14). Suppose that Ω is a complete metric space and that $f\colon (D, d) \to (\Omega; \rho)$ is uniformly continuous, where D is dense in (X, d). Use Exercise 5 to show that there is a uniformly continuous function $g\colon X \to \Omega$ with $g(x) = f(x)$ for every x in D.
7. Let G be an open subset of $\mathbb{C}$ and let P be a polygon in G from a to b. Use Theorems 5.15 and 5.17 to show that there is a polygon $Q \subset G$ from a to b which is composed of line segments which are parallel to either the real or imaginary axes.
8. Use Lebesgue's Covering Lemma (4.8) to give another proof of Theorem 5.15.
9. Prove the following converse to Exercise 2.5. Suppose (X, d) is a compact metric space having the property that for every $\epsilon > 0$ and for any points a, b in X, there are points $z_0, z_1, \ldots, z_n$ in X with $z_0 = a$, $z_n = b$, and $d(z_{k-1}, z_k) < \epsilon$ for $1 \leq k \leq n$. Then (X, d) is connected. (Hint: Use Theorem 5.17.)
10. Let f and g be continuous functions from (X, d) to (Ω, p) and let D be a dense subset of X. Prove that if $f(x) = g(x)$ for x in D then $f = g$. Use this to show that the function g obtained in Exercise 6 is unique.

§6. *Uniform convergence*

Let X be a set and (Ω, ρ) a metric space and suppose $f, f_1, f_2, \ldots$ are functions from X into Ω. The sequence $\{f_n\}$ *converges uniformly* to f — written $f = u-\lim f_n$—if for every $\epsilon > 0$ there is an integer N (depending on ϵ alone) such that $\rho(f(x), f_n(x)) < \epsilon$ for all x in X, whenever $n \geq N$. Hence,

$$\sup\{\rho(f(x), f_n(x))\colon x \in X\} \leq \epsilon$$

whenever $n \geq N$.

The first problem is this: If X is not just a set but a metric space and each f_n is continuous does it follow that f is continuous? The answer is yes.

6.1 Theorem. *Suppose $f_n\colon (X, d) \to (\Omega, \rho)$ is continuous for each n and that $f = u-\lim f_n$; then f is continuous.*

Proof. Fix x_0 in X and $\epsilon > 0$; we wish to find a $\delta > 0$ such that $\rho(f(x_0), f(x)) < \epsilon$ when $d(x_0, x) < \delta$. Since $f = u-\lim f_n$, there is a function f_n with $\rho(f(x), f_n(x)) < \epsilon/3$ for all x in X. Since f_n is continuous there is a $\delta > 0$ such that $\rho(f_n(x_0), f_n(x)) < \epsilon/3$ when $d(x_0, x) < \delta$. Therefore, if $d(x_0, x) < \delta$, $\rho(f(x_0), f(x)) \leq \rho(f(x_0), f_n(x_0)) + \rho(f_n(x_0), f_n(x)) + \rho(f_n(x), f(x)) < \epsilon$. ■

Let us consider the special case where $\Omega = \mathbb{C}$. If $u_n\colon X \to \mathbb{C}$, let $f_n(x) = u_1(x) + \ldots + u_n(x)$. We say $f(x) = \sum_{n=1}^{\infty} u_n(x)$ iff $f(x) = \lim f_n(x)$ for each x in X. The *series* $\sum_1^{\infty} u_n$ is *uniformly* convergent to f iff $f = u-\lim f_n$.

6.2 Weierstrass M-Test. *Let $u_n\colon X \to \mathbb{C}$ be a function such that $|u_n(x)| \leq M_n$ for every x in X and suppose the constants satisfy $\sum_{n=1}^{\infty} M_n < \infty$. Then $\sum_1^{\infty} u_n$ is uniformly convergent.*

Proof. Let $f_n(x) = u_1(x) + \ldots + u_n(x)$. Then for $n > m$, $|f_n(x) - f_m(x)| = |u_{m+1}(x) + \ldots + u_n(x)| \leq \sum_{k=m+1}^{n} M_k$ for each x. Since $\sum_1^{\infty} M_k$ converges, $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{C}$. Thus there is a number $\xi \in \mathbb{C}$ with $\xi = \lim f_n(x)$. Define $f(x) = \xi$; this gives a function $f\colon X \to \mathbb{C}$. Now

$$|f(x) - f_n(x)| = \left|\sum_{k=n+1}^{\infty} u_n(x)\right| \leq \sum_{k=n+1}^{\infty} |u_n(x)| \leq \sum_{k=n+1}^{\infty} M_k;$$

since $\sum_1^{\infty} M_k$ is convergent, for any $\epsilon > 0$ there is an integer N such that $\sum_{k=n+1}^{\infty} M_k < \epsilon$ whenever $n \geq N$. This gives $|f(x) - f_n(x)| < \epsilon$ for all x in X when $n \geq N$. ■

Exercise

1. Let $\{f_n\}$ in a sequence of uniformly continuous functions from (X, d) into (Ω, p) and suppose that $f = u-\lim f_n$ exists. Prove that f is uniformly continuous. If each f_n is a Lipschitz function with constant M_n and $\sup M_n < \infty$, show that f is a Lipschitz function. If $\sup M_n = \infty$, show that f may fail to be Lipschitz.