4

CONTINUITY

The function concept and some of the related terminology were introduced in Definitions 2.1 and 2.2. Although we shall (in later chapters) be mainly interested in real and complex functions (i.e., in functions whose values are real or complex numbers) we shall also discuss vector-valued functions (i.e., functions with values in R^k) and functions with values in an arbitrary metric space. The theorems we shall discuss in this general setting would not become any easier if we restricted ourselves to real functions, for instance, and it actually simplifies and clarifies the picture to discard unnecessary hypotheses and to state and prove theorems in an appropriately general context.

The domains of definition of our functions will also be metric spaces, suitably specialized in various instances.

LIMITS OF FUNCTIONS

4.1 Definition Let X and Y be metric spaces; suppose $E \subset X$, f maps E into Y, and p is a limit point of E. We write $f(x) \to q$ as $x \to p$, or

(1) $$\lim_{x \to p} f(x) = q$$

if there is a point $q \in Y$ with the following property: For every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), q) < \varepsilon \tag{2}$$

for all points $x \in E$ for which

$$0 < d_X(x, p) < \delta. \tag{3}$$

The symbols d_X and d_Y refer to the distances in X and Y, respectively.

If X and/or Y are replaced by the real line, the complex plane, or by some euclidean space R^k, the distances d_X, d_Y are of course replaced by absolute values, or by norms of differences (see Sec. 2.16).

It should be noted that $p \in X$, but that p need not be a point of E in the above definition. Moreover, even if $p \in E$, we may very well have $f(p) \neq \lim_{x \to p} f(x)$.

We can recast this definition in terms of limits of sequences:

4.2 Theorem Let X, Y, E, f, *and* p *be as in Definition* 4.1. *Then*

$$\lim_{x \to p} f(x) = q \tag{4}$$

if and only if

$$\lim_{n \to \infty} f(p_n) = q \tag{5}$$

for every sequence $\{p_n\}$ *in* E *such that*

$$p_n \neq p, \qquad \lim_{n \to \infty} p_n = p. \tag{6}$$

Proof Suppose (4) holds. Choose $\{p_n\}$ in E satisfying (6). Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that $d_Y(f(x), q) < \varepsilon$ if $x \in E$ and $0 < d_X(x, p) < \delta$. Also, there exists N such that $n > N$ implies $0 < d_X(p_n, p) < \delta$. Thus, for $n > N$, we have $d_Y(f(p_n), q) < \varepsilon$, which shows that (5) holds.

Conversely, suppose (4) is false. Then there exists some $\varepsilon > 0$ such that for every $\delta > 0$ there exists a point $x \in E$ (depending on δ), for which $d_Y(f(x), q) \geq \varepsilon$ but $0 < d_X(x, p) < \delta$. Taking $\delta_n = 1/n$ $(n = 1, 2, 3, \ldots)$, we thus find a sequence in E satisfying (6) for which (5) is false.

Corollary *If f has a limit at p, this limit is unique.*

This follows from Theorems 3.2(*b*) and 4.2.

4.3 Definition Suppose we have two complex functions, f and g, both defined on E. By $f+g$ we mean the function which assigns to each point x of E the number $f(x)+g(x)$. Similarly we define the difference $f-g$, the product fg, and the quotient f/g of the two functions, with the understanding that the quotient is defined only at those points x of E at which $g(x) \neq 0$. If f assigns to each point x of E the same number c, then f is said to be a constant function, or simply a constant, and we write $f = c$. If f and g are real functions, and if $f(x) \geq g(x)$ for every $x \in E$, we shall sometimes write $f \geq g$, for brevity.

Similarly, if $\mathbf{f}$ and $\mathbf{g}$ map E into R^k, we define $\mathbf{f}+\mathbf{g}$ and $\mathbf{f}\cdot\mathbf{g}$ by

$$(\mathbf{f}+\mathbf{g})(x) = \mathbf{f}(x)+\mathbf{g}(x), \qquad (\mathbf{f}\cdot\mathbf{g})(x) = \mathbf{f}(x)\cdot\mathbf{g}(x);$$

and if λ is a real number, $(\lambda\mathbf{f})(x) = \lambda\mathbf{f}(x)$.

4.4 Theorem *Suppose $E \subset X$, a metric space, p is a limit point of E, f and g are complex functions on E, and*

$$\lim_{x\to p} f(x) = A, \qquad \lim_{x\to p} g(x) = B.$$

Then (*a*) $\lim_{x\to p} (f+g)(x) = A+B$;

(*b*) $\lim_{x\to p} (fg)(x) = AB$;

(*c*) $\lim_{x\to p} \left(\frac{f}{g}\right)(x) = \frac{A}{B}$, *if* $B \neq 0$.

Proof In view of Theorem 4.2, these assertions follow immediately from the analogous properties of sequences (Theorem 3.3).

Remark If $\mathbf{f}$ and $\mathbf{g}$ map E into R^k, then (*a*) remains true, and (*b*) becomes

(*b'*) $\lim_{x\to p} (\mathbf{f}\cdot\mathbf{g})(x) = \mathbf{A}\cdot\mathbf{B}$.

(Compare Theorem 3.4.)

CONTINUOUS FUNCTIONS

4.5 Definition Suppose X and Y are metric spaces, $E \subset X$, $p \in E$, and f maps E into Y. Then f is said to be *continuous at* p if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \varepsilon$$

for all points $x \in E$ for which $d_X(x, p) < \delta$.

If f is continuous at every point of E, then f is said to be *continuous on* E.

It should be noted that f has to be defined at the point p in order to be continuous at p. (Compare this with the remark following Definition 4.1.)

If p is an isolated point of E, then our definition implies that every function f which has E as its domain of definition is continuous at p. For, no matter which $\varepsilon > 0$ we choose, we can pick $\delta > 0$ so that the only point $x \in E$ for which $d_X(x, p) < \delta$ is $x = p$; then

$$d_Y(f(x), f(p)) = 0 < \varepsilon.$$

4.6 Theorem *In the situation given in Definition 4.5, assume also that p is a limit point of E. Then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.*

Proof This is clear if we compare Definitions 4.1 and 4.5.

We now turn to compositions of functions. A brief statement of the following theorem is that a continuous function of a continuous function is continuous.

4.7 Theorem *Suppose X, Y, Z are metric spaces, $E \subset X$, f maps E into Y, g maps the range of f, $f(E)$, into Z, and h is the mapping of E into Z defined by*

$$h(x) = g(f(x)) \qquad (x \in E).$$

If f is continuous at a point $p \in E$ and if g is continuous at the point $f(p)$, then h is continuous at p.

This function h is called the *composition* or the *composite* of f and g. The notation

$$h = g \circ f$$

is frequently used in this context.

Proof Let $\varepsilon > 0$ be given. Since g is continuous at $f(p)$, there exists $\eta > 0$ such that

$$d_Z(g(y), g(f(p))) < \varepsilon \text{ if } d_Y(y, f(p)) < \eta \text{ and } y \in f(E).$$

Since f is continuous at p, there exists $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \eta \text{ if } d_X(x, p) < \delta \text{ and } x \in E.$$

It follows that

$$d_Z(h(x), h(p)) = d_Z(g(f(x)), g(f(p))) < \varepsilon$$

if $d_X(x, p) < \delta$ and $x \in E$. Thus h is continuous at p.

4.8 Theorem *A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y.*

(Inverse images are defined in Definition 2.2.) This is a very useful characterization of continuity.

Proof Suppose f is continuous on X and V is an open set in Y. We have to show that every point of $f^{-1}(V)$ is an interior point of $f^{-1}(V)$. So, suppose $p \in X$ and $f(p) \in V$. Since V is open, there exists $\varepsilon > 0$ such that $y \in V$ if $d_Y(f(p), y) < \varepsilon$; and since f is continuous at p, there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \varepsilon$ if $d_X(x, p) < \delta$. Thus $x \in f^{-1}(V)$ as soon as $d_X(x, p) < \delta$.

Conversely, suppose $f^{-1}(V)$ is open in X for every open set V in Y. Fix $p \in X$ and $\varepsilon > 0$, let V be the set of all $y \in Y$ such that $d_Y(y, f(p)) < \varepsilon$. Then V is open; hence $f^{-1}(V)$ is open; hence there exists $\delta > 0$ such that $x \in f^{-1}(V)$as soon as $d_X(p, x) < \delta$. But if $x \in f^{-1}(V)$, then $f(x) \in V$, so that $d_Y(f(x), f(p)) < \varepsilon$.

This completes the proof.

Corollary *A mapping f of a metric space X into a metric space Y is continuous if and only if $f^{-1}(C)$ is closed in X for every closed set C in Y.*

This follows from the theorem, since a set is closed if and only if its complement is open, and since $f^{-1}(E^c) = [f^{-1}(E)]^c$ for every $E \subset Y$.

We now turn to complex-valued and vector-valued functions, and to functions defined on subsets of R^k.

4.9 Theorem *Let f and g be complex continuous functions on a metric space X. Then $f + g$, fg, and f/g are continuous on X.*

In the last case, we must of course assume that $g(x) \neq 0$, for all $x \in X$.

Proof At isolated points of X there is nothing to prove. At limit points, the statement follows from Theorems 4.4 and 4.6.

4.10 Theorem

(*a*) *Let $f_1, \ldots, f_k$ be real functions on a metric space X, and let $\mathbf{f}$ be the mapping of X into R^k defined by*

$$\mathbf{f}(x) = (f_1(x), \ldots, f_k(x)) \qquad (x \in X); \tag{7}$$

then $\mathbf{f}$ is continuous if and only if each of the functions $f_1, \ldots, f_k$ is continuous.
(*b*) *If $\mathbf{f}$ and $\mathbf{g}$ are continuous mappings of X into R^k, then $\mathbf{f} + \mathbf{g}$ and $\mathbf{f} \cdot \mathbf{g}$ are continuous on X.*

The functions $f_1, \ldots, f_k$ are called the *components* of $\mathbf{f}$. Note that $\mathbf{f} + \mathbf{g}$ is a mapping into R^k, whereas $\mathbf{f} \cdot \mathbf{g}$ is a *real* function on X.

Proof Part (*a*) follows from the inequalities

$$|f_j(x) - f_j(y)| \le |\mathbf{f}(x) - \mathbf{f}(y)| = \left\{ \sum_{i=1}^{k} |f_i(x) - f_i(y)|^2 \right\}^{\frac{1}{2}},$$

for $j = 1, \ldots, k$. Part (*b*) follows from (*a*) and Theorem 4.9.

4.11 Examples If $x_1, \ldots, x_k$ are the coordinates of the point $\mathbf{x} \in R^k$, the functions ϕ_i defined by

$$\phi_i(\mathbf{x}) = x_i \qquad (\mathbf{x} \in R^k) \tag{8}$$

are continuous on R^k, since the inequality

$$|\phi_i(\mathbf{x}) - \phi_i(\mathbf{y})| \le |\mathbf{x} - \mathbf{y}|$$

shows that we may take $\delta = \varepsilon$ in Definition 4.5. The functions ϕ_i are sometimes called the *coordinate functions.*

Repeated application of Theorem 4.9 then shows that every monomial

$$x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k} \tag{9}$$

where $n_1, \ldots, n_k$ are nonnegative integers, is continuous on R^k. The same is true of constant multiples of (9), since constants are evidently continuous. It follows that every polynomial P, given by

$$P(\mathbf{x}) = \Sigma c_{n_1 \cdots n_k} x_1^{n_1} \ldots x_k^{n_k} \qquad (\mathbf{x} \in R^k), \tag{10}$$

is continuous on R^k. Here the coefficients $c_{n_1 \cdots n_k}$ are complex numbers, $n_1, \ldots, n_k$ are nonnegative integers, and the sum in (10) has finitely many terms.

Furthermore, every rational function in $x_1, \ldots, x_k$, that is, every quotient of two polynomials of the form (10), is continuous on R^k wherever the denominator is different from zero.

From the triangle inequality one sees easily that

$$\big|\,|\mathbf{x}| - |\mathbf{y}|\,\big| \le |\mathbf{x} - \mathbf{y}| \quad (\mathbf{x}, \mathbf{y} \in R^k). \tag{11}$$

Hence the mapping $\mathbf{x} \to |\mathbf{x}|$ is a continuous real function on R^k.

If now $\mathbf{f}$ is a continuous mapping from a metric space X into R^k, and if ϕ is defined on X by setting $\phi(p) = |\mathbf{f}(p)|$, it follows, by Theorem 4.7, that ϕ is a continuous real function on X.

4.12 Remark We defined the notion of continuity for functions defined on a *subset* E of a metric space X. However, the complement of E in X plays no role whatever in this definition (note that the situation was somewhat different for limits of functions). Accordingly, we lose nothing of interest by discarding the complement of the domain of f. This means that we may just as well talk only about continuous mappings of one metric space into another, rather than of mappings of subsets. This simplifies statements and proofs of some theorems. We have already made use of this principle in Theorems 4.8 to 4.10, and will continue to do so in the following section on compactness.

CONTINUITY AND COMPACTNESS

4.13 Definition A mapping $\mathbf{f}$ of a set E into R^k is said to be *bounded* if there is a real number M such that $|\mathbf{f}(x)| \leq M$ for all $x \in E$.

4.14 Theorem *Suppose f is a continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact.*

Proof Let $\{V_\alpha\}$ be an open cover of $f(X)$. Since f is continuous, Theorem 4.8 shows that each of the sets $f^{-1}(V_\alpha)$ is open. Since X is compact, there are finitely many indices, say $\alpha_1, \ldots, \alpha_n$, such that

$$X \subset f^{-1}(V_{\alpha_1}) \cup \cdots \cup f^{-1}(V_{\alpha_n}). \tag{12}$$

Since $f(f^{-1}(E)) \subset E$ for every $E \subset Y$, (12) implies that

$$f(X) \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}. \tag{13}$$

This completes the proof.

Note: We have used the relation $f(f^{-1}(E)) \subset E$, valid for $E \subset Y$. If $E \subset X$, then $f^{-1}(f(E)) \supset E$; equality need not hold in either case.

We shall now deduce some consequences of Theorem 4.14.

4.15 Theorem *If $\mathbf{f}$ is a continuous mapping of a compact metric space X into R^k, then $\mathbf{f}(X)$ is closed and bounded. Thus, $\mathbf{f}$ is bounded.*

This follows from Theorem 2.41. The result is particularly important when f is real:

4.16 Theorem *Suppose f is a continuous real function on a compact metric space X, and*

$$M = \sup_{p \in X} f(p), \qquad m = \inf_{p \in X} f(p). \tag{14}$$

Then there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

The notation in (14) means that M is the least upper bound of the set of all numbers $f(p)$, where p ranges over X, and that m is the greatest lower bound of this set of numbers.

The conclusion may also be stated as follows: *There exist points p and q in X such that $f(q) \le f(x) \le f(p)$ for all $x \in X$*; that is, f attains its maximum (at p) and its minimum (at q).

Proof By Theorem 4.15, $f(X)$ is a closed and bounded set of real numbers; hence $f(X)$ contains

$$M = \sup f(X) \qquad \text{and} \qquad m = \inf f(X),$$

by Theorem 2.28.

4.17 Theorem *Suppose f is a continuous 1-1 mapping of a compact metric space X onto a metric space Y. Then the inverse mapping f^{-1} defined on Y by*

$$f^{-1}(f(x)) = x \qquad (x \in X)$$

is a continuous mapping of Y onto X.

Proof Applying Theorem 4.8 to f^{-1} in place of f, we see that it suffices to prove that $f(V)$ is an open set in Y for every open set V in X. Fix such a set V.

The complement V^c of V is closed in X, hence compact (Theorem 2.35); hence $f(V^c)$ is a compact subset of Y (Theorem 4.14) and so is closed in Y (Theorem 2.34). Since f is one-to-one and onto, $f(V)$ is the complement of $f(V^c)$. Hence $f(V)$ is open.

4.18 Definition Let f be a mapping of a metric space X into a metric space Y. We say that f is *uniformly continuous* on X if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(p), f(q)) < \varepsilon \tag{15}$$

for all p and q in X for which $d_X(p, q) < \delta$.

Let us consider the differences between the concepts of continuity and of uniform continuity. First, uniform continuity is a property of a function on a set, whereas continuity can be defined at a single point. To ask whether a given function is uniformly continuous at a certain point is meaningless. Second, if f is continuous on X, then it is possible to find, for each $\varepsilon > 0$ and for each point p of X, a number $\delta > 0$ having the property specified in Definition 4.5. This δ depends on ε *and* on p. If f is, however, uniformly continuous on X, then it is possible, for each $\varepsilon > 0$, to find *one* number $\delta > 0$ which will do for *all* points p of X.

Evidently, every uniformly continuous function is continuous. That the two concepts are equivalent on compact sets follows from the next theorem.

4.19 Theorem *Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is uniformly continuous on X.*

Proof Let $\varepsilon > 0$ be given. Since f is continuous, we can associate to each point $p \in X$ a positive number $\phi(p)$ such that

$$q \in X,\ d_X(p, q) < \phi(p) \quad \text{implies} \quad d_Y(f(p), f(q)) < \frac{\varepsilon}{2}. \tag{16}$$

Let $J(p)$ be the set of all $q \in X$ for which

$$d_X(p, q) < \tfrac{1}{2}\phi(p). \tag{17}$$

Since $p \in J(p)$, the collection of all sets $J(p)$ is an open cover of X; and since X is compact, there is a finite set of points $p_1, \ldots, p_n$ in X, such that

$$X \subset J(p_1) \cup \cdots \cup J(p_n). \tag{18}$$

We put

$$\delta = \tfrac{1}{2} \min\, [\phi(p_1), \ldots, \phi(p_n)]. \tag{19}$$

Then $\delta > 0$. (This is one point where the finiteness of the covering, inherent in the definition of compactness, is essential. The minimum of a finite set of positive numbers is positive, whereas the inf of an infinite set of positive numbers may very well be 0.)

Now let q and p be points of X, such that $d_X(p, q) < \delta$. By (18), there is an integer m, $1 \le m \le n$, such that $p \in J(p_m)$; hence

$$d_X(p, p_m) < \tfrac{1}{2}\phi(p_m), \tag{20}$$

and we also have

$$d_X(q, p_m) \le d_X(p, q) + d_X(p, p_m) < \delta + \tfrac{1}{2}\phi(p_m) \le \phi(p_m).$$

Finally, (16) shows that therefore

$$d_Y(f(p), f(q)) \le d_Y(f(p), f(p_m)) + d_Y(f(q), f(p_m)) < \varepsilon.$$

This completes the proof.

An alternative proof is sketched in Exercise 10.

We now proceed to show that compactness is essential in the hypotheses of Theorems 4.14, 4.15, 4.16, and 4.19.

4.20 Theorem *Let E be a noncompact set in R^1. Then*

(*a*) *there exists a continuous function on E which is not bounded;*

(*b*) *there exists a continuous and bounded function on E which has no maximum.*

If, in addition, E is bounded, then

(*c*) *there exists a continuous function on* E *which is not uniformly continuous.*

Proof Suppose first that E is bounded, so that there exists a limit point x_0 of E which is not a point of E. Consider

$$f(x) = \frac{1}{x - x_0} \qquad (x \in E). \tag{21}$$

This is continuous on E (Theorem 4.9), but evidently unbounded. To see that (21) is not uniformly continuous, let $\varepsilon > 0$ and $\delta > 0$ be arbitrary, and choose a point $x \in E$ such that $|x - x_0| < \delta$. Taking t close enough to x_0, we can then make the difference $|f(t) - f(x)|$ greater than ε, although $|t - x| < \delta$. Since this is true for every $\delta > 0$, f is not uniformly continuous on E.

The function g given by

$$g(x) = \frac{1}{1 + (x - x_0)^2} \qquad (x \in E) \tag{22}$$

is continuous on E, and is bounded, since $0 < g(x) < 1$. It is clear that

$$\sup_{x \in E} g(x) = 1,$$

whereas $g(x) < 1$ for all $x \in E$. Thus g has no maximum on E.

Having proved the theorem for bounded sets E, let us now suppose that E is unbounded. Then $f(x) = x$ establishes (*a*), whereas

$$h(x) = \frac{x^2}{1 + x^2} \qquad (x \in E) \tag{23}$$

establishes (*b*), since

$$\sup_{x \in E} h(x) = 1$$

and $h(x) < 1$ for all $x \in E$.

Assertion (*c*) would be false if boundedness were omitted from the hypotheses. For, let E be the set of all integers. Then every function defined on E is uniformly continuous on E. To see this, we need merely take $\delta < 1$ in Definition 4.18.

We conclude this section by showing that compactness is also essential in Theorem 4.17.

4.21 Example Let X be the half-open interval $[0, 2\pi)$ on the real line, and let $\mathbf{f}$ be the mapping of X onto the circle Y consisting of all points whose distance from the origin is 1, given by

$$\mathbf{f}(t) = (\cos t, \sin t) \qquad (0 \le t < 2\pi). \tag{24}$$

The continuity of the trigonometric functions cosine and sine, as well as their periodicity properties, will be established in Chap. 8. These results show that $\mathbf{f}$ is a continuous 1-1 mapping of X onto Y.

However, the inverse mapping (which exists, since $\mathbf{f}$ is one-to-one and onto) fails to be continuous at the point $(1, 0) = \mathbf{f}(0)$. Of course, X is not compact in this example. (It may be of interest to observe that $\mathbf{f}^{-1}$ fails to be continuous in spite of the fact that Y *is* compact!)

CONTINUITY AND CONNECTEDNESS

4.22 Theorem *If f is a continuous mapping of a metric space X into a metric space Y, and if E is a connected subset of X, then $f(E)$ is connected.*

Proof Assume, on the contrary, that $f(E) = A \cup B$, where A and B are nonempty separated subsets of Y. Put $G = E \cap f^{-1}(A)$, $H = E \cap f^{-1}(B)$.

Then $E = G \cup H$, and neither G nor H is empty.

Since $A \subset \bar{A}$ (the closure of A), we have $G \subset f^{-1}(\bar{A})$; the latter set is closed, since f is continuous; hence $\bar{G} \subset f^{-1}(\bar{A})$. It follows that $f(\bar{G}) \subset \bar{A}$. Since $f(H) = B$ and $\bar{A} \cap B$ is empty, we conclude that $\bar{G} \cap H$ is empty.

The same argument shows that $G \cap \bar{H}$ is empty. Thus G and H are separated. This is impossible if E is connected.

4.23 Theorem *Let f be a continuous real function on the interval $[a, b]$. If $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.*

A similar result holds, of course, if $f(a) > f(b)$. Roughly speaking, the theorem says that a continuous real function assumes all intermediate values on an interval.

Proof By Theorem 2.47, $[a, b]$ is connected; hence Theorem 4.22 shows that $f([a, b])$ is a connected subset of R^1, and the assertion follows if we appeal once more to Theorem 2.47.

4.24 Remark At first glance, it might seem that Theorem 4.23 has a converse. That is, one might think that if for any two points $x_1 < x_2$ and for any number c between $f(x_1)$ and $f(x_2)$ there is a point x in (x_1, x_2) such that $f(x) = c$, then f must be continuous.

That this is not so may be concluded from Example 4.27(*d*).

DISCONTINUITIES

If x is a point in the domain of definition of the function f at which f is not continuous, we say that f is *discontinuous* at x, or that f *has a discontinuity* at x. If f is defined on an interval or on a segment, it is customary to divide discontinuities into two types. Before giving this classification, we have to define the *right-hand* and the *left-hand limits* of f at x, which we denote by $f(x+)$ and $f(x-)$, respectively.

4.25 Definition Let f be defined on (a, b). Consider any point x such that $a \leq x < b$. We write

$$f(x+) = q$$

if $f(t_n) \to q$ as $n \to \infty$, for all sequences $\{t_n\}$ in (x, b) such that $t_n \to x$. To obtain the definition of $f(x-)$, for $a < x \leq b$, we restrict ourselves to sequences $\{t_n\}$ in (a, x).

It is clear that any point x of (a, b), $\lim_{t \to x} f(t)$ exists if and only if

$$f(x+) = f(x-) = \lim_{t \to x} f(t).$$

4.26 Definition Let f be defined on (a, b). If f is discontinuous at a point x, and if $f(x+)$ and $f(x-)$ exist, then f is said to have a discontinuity of the *first kind*, or a *simple discontinuity*, at x. Otherwise the discontinuity is said to be of the *second kind.*

There are two ways in which a function can have a simple discontinuity: either $f(x+) \neq f(x-)$ [in which case the value $f(x)$ is immaterial], or $f(x+) = f(x-) \neq f(x)$.

4.27 Examples

(*a*) Define

$$f(x) = \begin{cases} 1 & (x \text{ rational}), \\ 0 & (x \text{ irrational}). \end{cases}$$

Then f has a discontinuity of the second kind at every point x, since neither $f(x+)$ nor $f(x-)$ exists.

(*b*) Define

$$f(x) = \begin{cases} x & (x \text{ rational}), \\ 0 & (x \text{ irrational}). \end{cases}$$

Then f is continuous at $x = 0$ and has a discontinuity of the second kind at every other point.

(c) Define

$$f(x) = \begin{cases} x+2 & (-3 < x < -2), \\ -x-2 & (-2 \le x < 0), \\ x+2 & (0 \le x < 1). \end{cases}$$

Then f has a simple discontinuity at $x = 0$ and is continuous at every other point of $(-3, 1)$.

(d) Define

$$f(x) = \begin{cases} \sin \dfrac{1}{x} & (x \neq 0), \\ 0 & (x = 0). \end{cases}$$

Since neither $f(0+)$ nor $f(0-)$ exists, f has a discontinuity of the second kind at $x = 0$. We have not yet shown that $\sin x$ is a continuous function. If we assume this result for the moment, Theorem 4.7 implies that f is continuous at every point $x \neq 0$.

MONOTONIC FUNCTIONS

We shall now study those functions which never decrease (or never increase) on a given segment.

4.28 Definition Let f be real on (a, b). Then f is said to be *monotonically increasing* on (a, b) if $a < x < y < b$ implies $f(x) \le f(y)$. If the last inequality is reversed, we obtain the definition of a *monotonically decreasing* function. The class of monotonic functions consists of both the increasing and the decreasing functions.

4.29 Theorem *Let f be monotonically increasing on (a, b). Then $f(x+)$ and $f(x-)$ exist at every point of x of (a, b). More precisely,*

$$\sup_{a<t<x} f(t) = f(x-) \le f(x) \le f(x+) = \inf_{x<t<b} f(t). \tag{25}$$

Furthermore, if $a < x < y < b$, then

$$f(x+) \le f(y-). \tag{26}$$

Analogous results evidently hold for monotonically decreasing functions.

Proof By hypothesis, the set of numbers $f(t)$, where $a < t < x$, is bounded above by the number $f(x)$, and therefore has a least upper bound which we shall denote by A. Evidently $A \le f(x)$. We have to show that $A = f(x-)$.

Let $\varepsilon > 0$ be given. It follows from the definition of A as a least upper bound that there exists $\delta > 0$ such that $a < x - \delta < x$ and

$$A - \varepsilon < f(x - \delta) \le A. \tag{27}$$

Since f is monotonic, we have

$$f(x - \delta) \le f(t) \le A \qquad (x - \delta < t < x). \tag{28}$$

Combining (27) and (28), we see that

$$|f(t) - A| < \varepsilon \qquad (x - \delta < t < x).$$

Hence $f(x-) = A$.

The second half of (25) is proved in precisely the same way.

Next, if $a < x < y < b$, we see from (25) that

$$f(x+) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t). \tag{29}$$

The last equality is obtained by applying (25) to (a, y) in place of (a, b). Similarly,

$$f(y-) = \sup_{a<t<y} f(t) = \sup_{x<t<y} f(t). \tag{30}$$

Comparison of (29) and (30) gives (26).

Corollary *Monotonic functions have no discontinuities of the second kind.*

This corollary implies that every monotonic function is discontinuous at a countable set of points at most. Instead of appealing to the general theorem whose proof is sketched in Exercise 17, we give here a simple proof which is applicable to monotonic functions.

4.30 Theorem *Let f be monotonic on (a, b). Then the set of points of (a, b) at which f is discontinuous is at most countable.*

Proof Suppose, for the sake of definiteness, that f is increasing, and let E be the set of points at which f is discontinuous.

With every point x of E we associate a rational number $r(x)$ such that

$$f(x-) < r(x) < f(x+).$$

Since $x_1 < x_2$ implies $f(x_1+) \le f(x_2-)$, we see that $r(x_1) \ne r(x_2)$ if $x_1 \ne x_2$.

We have thus established a 1-1 correspondence between the set E and a subset of the set of rational numbers. The latter, as we know, is countable.

4.31 Remark It should be noted that the discontinuities of a monotonic function need not be isolated. In fact, given any countable subset E of (a, b), which may even be dense, we can construct a function f, monotonic on (a, b), discontinuous at every point of E, and at no other point of (a, b).

To show this, let the points of E be arranged in a sequence $\{x_n\}$, $n = 1, 2, 3, \ldots$. Let $\{c_n\}$ be a sequence of positive numbers such that Σc_n converges. Define

$$f(x) = \sum_{x_n < x} c_n \qquad (a < x < b). \tag{31}$$

The summation is to be understood as follows: Sum over those indices n for which $x_n < x$. If there are no points x_n to the left of x, the sum is empty; following the usual convention, we define it to be zero. Since (31) converges absolutely, the order in which the terms are arranged is immaterial.

We leave the verification of the following properties of f to the reader:

(*a*) f is monotonically increasing on (a, b);
(*b*) f is discontinuous at every point of E; in fact,

$$f(x_n+) - f(x_n-) = c_n.$$

(*c*) f is continuous at every other point of (a, b).

Moreover, it is not hard to see that $f(x-) = f(x)$ at all points of (a, b). If a function satisfies this condition, we say that f is *continuous from the left*. If the summation in (31) were taken over all indices n for which $x_n \le x$, we would have $f(x+) = f(x)$ at every point of (a, b); that is, f would be *continuous from the right*.

Functions of this sort can also be defined by another method; for an example we refer to Theorem 6.16.

INFINITE LIMITS AND LIMITS AT INFINITY

To enable us to operate in the extended real number system, we shall now enlarge the scope of Definition 4.1, by reformulating it in terms of neighborhoods.

For any real number x, we have already defined a neighborhood of x to be any segment $(x - \delta, x + \delta)$.

4.32 Definition For any real c, the set of real numbers x such that $x > c$ is called a neighborhood of $+\infty$ and is written $(c, +\infty)$. Similarly, the set $(-\infty, c)$ is a neighborhood of $-\infty$.

4.33 Definition Let f be a real function defined on $E \subset R$. We say that

$$f(t) \to A \text{ as } t \to x,$$

where A and x are in the extended real number system, if for every neighborhood U of A there is a neighborhood V of x such that $V \cap E$ is not empty, and such that $f(t) \in U$ for all $t \in V \cap E$, $t \neq x$.

A moment's consideration will show that this coincides with Definition 4.1 when A and x are real.

The analogue of Theorem 4.4 is still true, and the proof offers nothing new. We state it, for the sake of completeness.

4.34 Theorem *Let f and g be defined on $E \subset R$. Suppose*

$$f(t) \to A, \qquad g(t) \to B \qquad \text{as } t \to x.$$

Then

(*a*) $f(t) \to A'$ *implies* $A' = A$.
(*b*) $(f+g)(t) \to A + B$,
(*c*) $(fg)(t) \to AB$,
(*d*) $(f/g)(t) \to A/B$,

provided the right members of (*b*), (*c*), *and* (*d*) *are defined.*

Note that $\infty - \infty$, $0 \cdot \infty$, ∞/∞, $A/0$ are not defined (see Definition 1.23).

EXERCISES

1. Suppose f is a real function defined on R^1 which satisfies

$$\lim_{h \to 0} [f(x+h) - f(x-h)] = 0$$

for every $x \in R^1$. Does this imply that f is continuous?

2. If f is a continuous mapping of a metric space X into a metric space Y, prove that

$$f(\bar{E}) \subset \overline{f(E)}$$

for every set $E \subset X$. ($\bar{E}$ denotes the closure of E.) Show, by an example, that $f(\bar{E})$ can be a proper subset of $\overline{f(E)}$.

3. Let f be a continuous real function on a metric space X. Let $Z(f)$ (the *zero set* of f) be the set of all $p \in X$ at which $f(p) = 0$. Prove that $Z(f)$ is closed.

4. Let f and g be continuous mappings of a metric space X into a metric space Y,

and let E be a dense subset of X. Prove that $f(E)$ is dense in $f(X)$. If $g(p) = f(p)$ for all $p \in E$, prove that $g(p) = f(p)$ for all $p \in X$. (In other words, a continuous mapping is determined by its values on a dense subset of its domain.)

5. If f is a real continuous function defined on a closed set $E \subset R^1$, prove that there exist continuous real functions g on R^1 such that $g(x) = f(x)$ for all $x \in E$. (Such functions g are called *continuous extensions* of f from E to R^1.) Show that the result becomes false if the word "closed" is omitted. Extend the result to vector-valued functions. *Hint:* Let the graph of g be a straight line on each of the segments which constitute the complement of E (compare Exercise 29, Chap. 2). The result remains true if R^1 is replaced by any metric space, but the proof is not so simple.

6. If f is defined on E, the *graph* of f is the set of points $(x, f(x))$, for $x \in E$. In particular, if E is a set of real numbers, and f is real-valued, the graph of f is a subset of the plane.

Suppose E is compact, and prove that f is continuous on E if and only if its graph is compact.

7. If $E \subset X$ and if f is a function defined on X, the *restriction* of f to E is the function g whose domain of definition is E, such that $g(p) = f(p)$ for $p \in E$. Define f and g on R^2 by: $f(0, 0) = g(0, 0) = 0$, $f(x, y) = xy^2/(x^2 + y^4)$, $g(x, y) = xy^2/(x^2 + y^6)$ if $(x, y) \neq (0, 0)$. Prove that f is bounded on R^2, that g is unbounded in every neighborhood of $(0, 0)$, and that f is not continuous at $(0, 0)$; nevertheless, the restrictions of both f and g to every straight line in R^2 are continuous!

8. Let f be a real uniformly continuous function on the bounded set E in R^1. Prove that f is bounded on E.

Show that the conclusion is false if boundedness of E is omitted from the hypothesis.

9. Show that the requirement in the definition of uniform continuity can be rephrased as follows, in terms of diameters of sets: To every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\operatorname{diam} f(E) < \varepsilon$ for all $E \subset X$ with $\operatorname{diam} E < \delta$.

10. Complete the details of the following alternative proof of Theorem 4.19: If f is not uniformly continuous, then for some $\varepsilon > 0$ there are sequences $\{p_n\}$, $\{q_n\}$ in X such that $d_X(p_n, q_n) \to 0$ but $d_Y(f(p_n), f(q_n)) > \varepsilon$. Use Theorem 2.37 to obtain a contradiction.

11. Suppose f is a uniformly continuous mapping of a metric space X into a metric space Y and prove that $\{f(x_n)\}$ is a Cauchy sequence in Y for every Cauchy sequence $\{x_n\}$ in X. Use this result to give an alternative proof of the theorem stated in Exercise 13.

12. A uniformly continuous function of a uniformly continuous function is uniformly continuous.

State this more precisely and prove it.

13. Let E be a dense subset of a metric space X, and let f be a uniformly continuous *real* function defined on E. Prove that f has a continuous extension from E to X

(see Exercise 5 for terminology). (Uniqueness follows from Exercise 4.) *Hint:* For each $p \in X$ and each positive integer n, let $V_n(p)$ be the set of all $q \in E$ with $d(p, q) < 1/n$. Use Exercise 9 to show that the intersection of the closures of the sets $f(V_1(p)), f(V_2(p)), \ldots$, consists of a single point, say $g(p)$, of R^1. Prove that the function g so defined on X is the desired extension of f.

Could the range space R^1 be replaced by R^k? By any compact metric space? By any complete metric space? By any metric space?

14. Let $I = [0, 1]$ be the closed unit interval. Suppose f is a continuous mapping of I into I. Prove that $f(x) = x$ for at least one $x \in I$.

15. Call a mapping of X into Y *open* if $f(V)$ is an open set in Y whenever V is an open set in X.

Prove that every continuous open mapping of R^1 into R^1 is monotonic.

16. Let $[x]$ denote the largest integer contained in x, that is, $[x]$ is the integer such that $x - 1 < [x] \leq x$; and let $(x) = x - [x]$ denote the fractional part of x. What discontinuities do the functions $[x]$ and (x) have?

17. Let f be a real function defined on (a, b). Prove that the set of points at which f has a simple discontinuity is at most countable. *Hint:* Let E be the set on which $f(x-) < f(x+)$. With each point x of E, associate a triple (p, q, r) of rational numbers such that

(*a*) $f(x-) < p < f(x+)$,

(*b*) $a < q < t < x$ implies $f(t) < p$,

(*c*) $x < t < r < b$ implies $f(t) > p$.

The set of all such triples is countable. Show that each triple is associated with at most one point of E. Deal similarly with the other possible types of simple discontinuities.

18. Every rational x can be written in the form $x = m/n$, where $n > 0$, and m and n are integers without any common divisors. When $x = 0$, we take $n = 1$. Consider the function f defined on R^1 by

$$f(x) = \begin{cases} 0 & (x \text{ irrational}), \\ \dfrac{1}{n} & \left(x = \dfrac{m}{n}\right). \end{cases}$$

Prove that f is continuous at every irrational point, and that f has a simple discontinuity at every rational point.

19. Suppose f is a real function with domain R^1 which has the intermediate value property: If $f(a) < c < f(b)$, then $f(x) = c$ for some x between a and b.

Suppose also, for every rational r, that the set of all x with $f(x) = r$ is closed.

Prove that f is continuous.

Hint: If $x_n \to x_0$ but $f(x_n) > r > f(x_0)$ for some r and all n, then $f(t_n) = r$ for some t_n between x_0 and x_n; thus $t_n \to x_0$. Find a contradiction. (N. J. Fine, *Amer. Math. Monthly*, vol. 73, 1966, p. 782.)

20. If E is a nonempty subset of a metric space X, define the distance from $x \in X$ to E by

$$\rho_E(x) = \inf_{z \in E} d(x, z).$$

(*a*) Prove that $\rho_E(x) = 0$ if and only if $x \in \bar{E}$.

(*b*) Prove that ρ_E is a uniformly continuous function on X, by showing that

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y)$$

for all $x \in X$, $y \in X$.

Hint: $\rho_E(x) \leq d(x, z) \leq d(x, y) + d(y, z)$, so that

$$\rho_E(x) \leq d(x, y) + \rho_E(y).$$

21. Suppose K and F are disjoint sets in a metric space X, K is compact, F is closed. Prove that there exists $\delta > 0$ such that $d(p, q) > \delta$ if $p \in K$, $q \in F$. *Hint:* ρ_F is a continuous positive function on K.

Show that the conclusion may fail for two disjoint closed sets if neither is compact.

22. Let A and B be disjoint nonempty closed sets in a metric space X, and define

$$f(p) = \frac{\rho_A(p)}{\rho_A(p) + \rho_B(p)} \quad (p \in X).$$

Show that f is a continuous function on X whose range lies in $[0, 1]$, that $f(p) = 0$ precisely on A and $f(p) = 1$ precisely on B. This establishes a converse of Exercise 3: Every closed set $A \subset X$ is $Z(f)$ for some continuous real f on X. Setting

$$V = f^{-1}([0, \tfrac{1}{2})), \qquad W = f^{-1}((\tfrac{1}{2}, 1]),$$

show that V and W are open and disjoint, and that $A \subset V$, $B \subset W$. (Thus pairs of disjoint closed sets in a metric space can be covered by pairs of disjoint open sets. This property of metric spaces is called *normality*.)

23. A real-valued function f defined in (a, b) is said to be *convex* if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

whenever $a < x < b$, $a < y < b$, $0 < \lambda < 1$. Prove that every convex function is continuous. Prove that every increasing convex function of a convex function is convex. (For example, if f is convex, so is e^f.)

If f is convex in (a, b) and if $a < s < t < u < b$, show that

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}.$$

24. Assume that f is a continuous real function defined in (a, b) such that

$$f\left(\frac{x + y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

for all $x, y \in (a, b)$. Prove that f is convex.

25. If $A \subset R^k$ and $B \subset R^k$, define $A + B$ to be the set of all sums $\mathbf{x} + \mathbf{y}$ with $\mathbf{x} \in A$, $\mathbf{y} \in B$.

(*a*) If K is compact and C is closed in R^k, prove that $K + C$ is closed.

Hint: Take $\mathbf{z} \notin K + C$, put $F = \mathbf{z} - C$, the set of all $\mathbf{z} - \mathbf{y}$ with $\mathbf{y} \in C$. Then K and F are disjoint. Choose δ as in Exercise 21. Show that the open ball with center $\mathbf{z}$ and radius δ does not intersect $K + C$.

(*b*) Let α be an irrational real number. Let C_1 be the set of all integers, let C_2 be the set of all $n\alpha$ with $n \in C_1$. Show that C_1 and C_2 are closed subsets of R^1 whose sum $C_1 + C_2$ is *not* closed, by showing that $C_1 + C_2$ is a countable dense subset of R^1.

26. Suppose X, Y, Z are metric spaces, and Y is compact. Let f map X into Y, let g be a continuous one-to-one mapping of Y into Z, and put $h(x) = g(f(x))$ for $x \in X$.

Prove that f is uniformly continuous if h is uniformly continuous.

Hint: g^{-1} has compact domain $g(Y)$, and $f(x) = g^{-1}(h(x))$.

Prove also that f is continuous if h is continuous.

Show (by modifying Example 4.21, or by finding a different example) that the compactness of Y cannot be omitted from the hypotheses, even when X and Z are compact.