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 $(x - \delta, x + \delta)$.