3 CHAPTER Continuity

Most of the calculus involves the study of continuous functions. In this chapter we study continuous and uniformly continuous functions.

§17 Continuous Functions

Recall that the salient features of a function f are:

(a) The set on which f is defined, called the *domain of* f and written $\mathrm{dom}(f)$;

(b) The assignment, rule or formula specifying the value $f(x)$ of f at each x in $\mathrm{dom}(f)$.

We will be concerned with functions f such that $\mathrm{dom}(f) \subseteq \mathbb{R}$ and such that f is a *real-valued function*, i.e., $f(x) \in \mathbb{R}$ for all $x \in \mathrm{dom}(f)$. Properly speaking, the symbol f represents the function while $f(x)$ represents the value of the function at x. However, a function is often given by specifying its values and without mentioning its domain. In this case, the domain is understood to be the *natural domain*: the largest subset of $\mathbb{R}$ on which the function is a well defined

K.A. Ross, *Elementary Analysis: The Theory of Calculus*,
Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4614-6271-2_3,

real-valued function. Thus "the function $f(x) = \frac{1}{x}$" is shorthand for "the function f given by $f(x) = \frac{1}{x}$ with natural domain $\{x \in \mathbb{R} : x \neq 0\}$." Similarly, the natural domain of $g(x) = \sqrt{4 - x^2}$ is $[-2, 2]$ and the natural domain of $\csc x = \frac{1}{\sin x}$ is the set of real numbers x not of the form $n\pi, n \in \mathbb{Z}$.

In keeping with the approach in this book, we will define continuity in terms of sequences. We then show our definition is equivalent to the usual ϵ–δ definition.

17.1 Definition.
Let f be a real-valued function whose domain is a subset of $\mathbb{R}$. The function f is *continuous at* x_0 in $\mathrm{dom}(f)$ if, for every sequence (x_n) in $\mathrm{dom}(f)$ converging to x_0, we have $\lim_n f(x_n) = f(x_0)$. If f is continuous at each point of a set $S \subseteq \mathrm{dom}(f)$, then f is said to be *continuous on* S. The function f is said to be *continuous* if it is continuous on $\mathrm{dom}(f)$.

Our definition implies that the values $f(x)$ are close to $f(x_0)$ when the values x are close to x_0. The next theorem says this in another way. In fact, condition (1) of the next theorem is the ϵ–δ definition of continuity given in many calculus books.

17.2 Theorem.
Let f be a real-valued function whose domain is a subset of $\mathbb{R}$. Then f is continuous at x_0 in $\mathrm{dom}(f)$ if and only if

$$\begin{gathered}\textit{for each} \quad \epsilon > 0 \quad \textit{there exists} \quad \delta > 0 \quad \textit{such that} \\ x \in \mathrm{dom}(f) \quad \textit{and} \quad |x - x_0| < \delta \quad \textit{imply} \quad |f(x) - f(x_0)| < \epsilon.\end{gathered} \tag{1}$$

Proof
Suppose (1) holds, and consider a sequence (x_n) in $\mathrm{dom}(f)$ such that $\lim x_n = x_0$. We need to prove $\lim f(x_n) = f(x_0)$. Let $\epsilon > 0$. By (1), there exists $\delta > 0$ such that

$$x \in \mathrm{dom}(f) \quad \text{and} \quad |x - x_0| < \delta \quad \text{imply} \quad |f(x) - f(x_0)| < \epsilon.$$

Since $\lim x_n = x_0$, there exists N so that

$$n > N \quad \text{implies} \quad |x_n - x_0| < \delta.$$

It follows that

$$n > N \quad \text{implies} \quad |f(x_n) - f(x_0)| < \epsilon.$$

This proves $\lim f(x_n) = f(x_0)$.

Now assume f is continuous at x_0, but (1) fails. Then there exists $\epsilon > 0$ so that the implication

$$\text{“}x \in \text{dom}(f) \quad \text{and} \quad |x - x_0| < \delta \quad \text{imply} \quad |f(x) - f(x_0)| < \epsilon\text{”}$$

fails for each $\delta > 0$. In particular, the implication

$$\text{“}x \in \text{dom}(f) \quad \text{and} \quad |x - x_0| < \frac{1}{n} \quad \text{imply} \quad |f(x) - f(x_0)| < \epsilon\text{”}$$

fails for each $n \in \mathbb{N}$. So for each $n \in \mathbb{N}$ there exists x_n in $\text{dom}(f)$ such that $|x_n - x_0| < \frac{1}{n}$ and yet $|f(x_0) - f(x_n)| \geq \epsilon$. Thus we have $\lim x_n = x_0$, but we cannot have $\lim f(x_n) = f(x_0)$ since $|f(x_0) - f(x_n)| \geq \epsilon$ for all n. This shows f cannot be continuous at x_0, contrary to our assumption. ■

As the next example illustrates, it is sometimes easier to work with the sequential definition of continuity in Definition 17.1 than the ϵ–δ property in Theorem 17.2. However, it is important to get comfortable with the ϵ–δ property, partly because the definition of uniform continuity is more closely related to the ϵ–δ property than the sequential definition.

Example 1
Let $f(x) = 2x^2 + 1$ for $x \in \mathbb{R}$. Prove f is continuous on $\mathbb{R}$ by
(a) Using the definition,
(b) Using the ϵ–δ property of Theorem 17.2.

Solution
(a) Suppose $\lim x_n = x_0$. Then we have

$$\lim f(x_n) = \lim[2x_n^2 + 1] = 2[\lim x_n]^2 + 1 = 2x_0^2 + 1 = f(x_0)$$

where the second equality is an application of the limit Theorems 9.2–9.4. Hence f is continuous at each x_0 in $\mathbb{R}$.

(b) Let x_0 be in $\mathbb{R}$ and let $\epsilon > 0$. We want to show $|f(x) - f(x_0)| < \epsilon$ provided $|x - x_0|$ is sufficiently small, i.e., less than some δ. We observe

$$|f(x) - f(x_0)| = |2x^2 + 1 - (2x_0^2 + 1)| = |2x^2 - 2x_0^2|$$
$$= 2|x - x_0| \cdot |x + x_0|.$$

We need to get a bound for $|x + x_0|$ that does not depend on x. We notice that if $|x - x_0| < 1$, say, then $|x| < |x_0| + 1$ and hence $|x + x_0| \le |x| + |x_0| < 2|x_0| + 1$. Thus we have

$$|f(x) - f(x_0)| \le 2|x - x_0|(2|x_0| + 1)$$

provided $|x - x_0| < 1$. To arrange for $2|x - x_0|(2|x_0| + 1) < \epsilon$, it suffices to have $|x - x_0| < \frac{\epsilon}{2(2|x_0|+1)}$ and also $|x - x_0| < 1$. So we put

$$\delta = \min\left\{1, \frac{\epsilon}{2(2|x_0| + 1)}\right\}.$$

The work above shows $|x - x_0| < \delta$ implies $|f(x) - f(x_0)| < \epsilon$, as desired. □

The reason that solution (a) in Example 1 is so much easier is that the careful analysis was done in proving the limit theorems in §9.

Example 2
Let $f(x) = x^2 \sin(\frac{1}{x})$ for $x \neq 0$ and $f(0) = 0$. The graph of f in Fig. 17.1 *looks* continuous. Prove f is continuous at 0.

Solution
Let $\epsilon > 0$. Clearly $|f(x) - f(0)| = |f(x)| \le x^2$ for all x. Since we want this to be less than ϵ, we set $\delta = \sqrt{\epsilon}$. Then $|x - 0| < \delta$ implies $x^2 < \delta^2 = \epsilon$, so

$$|x - 0| < \delta \quad \text{implies} \quad |f(x) - f(0)| < \epsilon.$$

Hence the ϵ–δ property holds and f is continuous at 0. □

In Example 2 it would have been equally easy to show that if $\lim x_n = 0$ then $\lim f(x_n) = 0$. The function f in Example 2 is also continuous at the other points of $\mathbb{R}$; see Example 4.

Example 3
Let $f(x) = \frac{1}{x} \sin(\frac{1}{x^2})$ for $x \neq 0$ and $f(0) = 0$. Show f is *discontinuous*, i.e., not continuous, at 0.

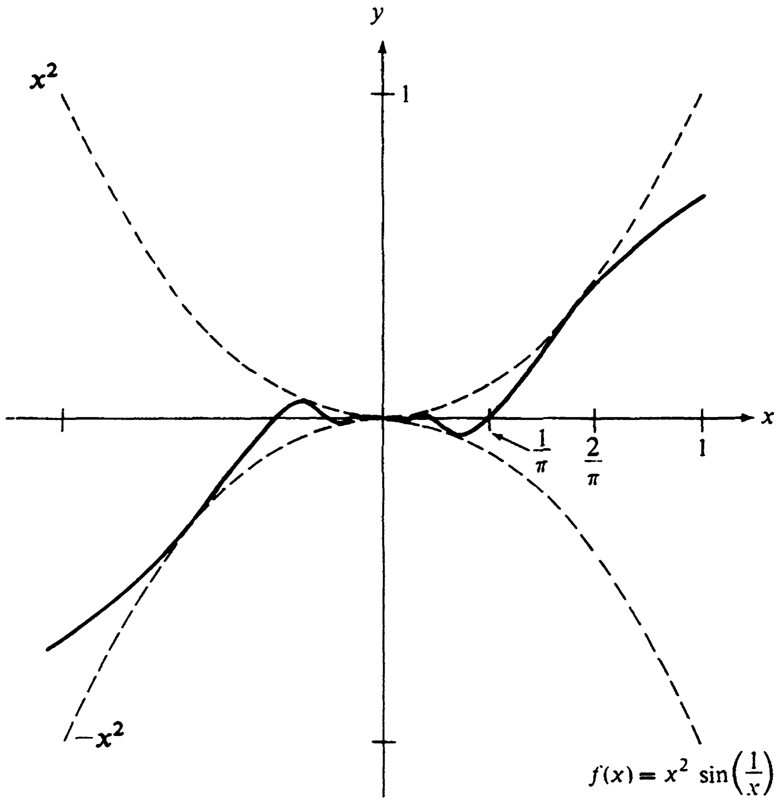

FIGURE 17.1

Solution

It suffices for us to find a sequence (x_n) converging to 0 such that $f(x_n)$ does not converge to $f(0) = 0$. So we will arrange for $\frac{1}{x_n}\sin(\frac{1}{x_n^2}) = \frac{1}{x_n}$ where $x_n \to 0$. Thus we want $\sin(\frac{1}{x_n^2}) = 1$, $\frac{1}{x_n^2} = 2\pi n + \frac{\pi}{2}$, $x_n^2 = \frac{1}{2\pi n + \frac{\pi}{2}}$ or $x_n = \frac{1}{\sqrt{2\pi n + \frac{\pi}{2}}}$. Then $\lim x_n = 0$ while $\lim f(x_n) = \lim \frac{1}{x_n} = +\infty$. □

Let f be a real-valued function. For k in $\mathbb{R}$, kf signifies the function defined by $(kf)(x) = kf(x)$ for $x \in \text{dom}(f)$. Also $|f|$ is the function defined by $|f|(x) = |f(x)|$ for $x \in \text{dom}(f)$. Thus if f is given by $f(x) = \sqrt{x} - 4$ for $x \geq 0$, then $3f$ is given by $(3f)(x) = 3\sqrt{x} - 12$ for $x \geq 0$, and $|f|$ is given by $|f|(x) = |\sqrt{x} - 4|$ for $x \geq 0$. Here is an easy theorem.

17.3 Theorem.
Let f be a real-valued function with $\mathrm{dom}(f) \subseteq \mathbb{R}$*. If f is continuous at x_0 in* $\mathrm{dom}(f)$*, then $|f|$ and kf, $k \in \mathbb{R}$, are continuous at x_0.*

Proof
Consider a sequence (x_n) in $\mathrm{dom}(f)$ converging to x_0. Since f is continuous at x_0, we have $\lim f(x_n) = f(x_0)$. Theorem 9.2 shows $\lim kf(x_n) = kf(x_0)$. This proves kf is continuous at x_0.

To prove $|f|$ is continuous at x_0, we need to prove $\lim |f(x_n)| = |f(x_0)|$. This follows from the inequality

$$||f(x_n)| - |f(x_0)|| \leq |f(x_n) - f(x_0)|; \tag{1}$$

see Exercise 3.5. [In detail, consider $\epsilon > 0$. Since $\lim f(x_n) = f(x_0)$, there exists N such that $n > N$ implies $|f(x_n) - f(x_0)| < \epsilon$. So by (1),

$$n > N \quad \text{implies} \quad ||f(x_n)| - |f(x_0)|| < \epsilon;$$

thus $\lim |f(x_n)| = |f(x_0)|$.] ∎

We remind readers that if f and g are real-valued functions, then we can combine f and g to obtain new functions:

$(f+g)(x) = f(x) + g(x)$;	$fg(x) = f(x)g(x)$;
$(f/g)(x) = \frac{f(x)}{g(x)}$;	$g \circ f(x) = g(f(x))$;
$\max(f,g)(x) = \max\{f(x), g(x)\}$;	$\min(f,g)(x) = \min\{f(x), g(x)\}$.

The function $g \circ f$ is called the *composition of g and f*. Each of these new functions is defined exactly where they make sense. Thus the domains of $f+g$, fg, $\max(f,g)$ and $\min(f,g)$ are $\mathrm{dom}(f) \cap \mathrm{dom}(g)$, the domain of f/g is the set $\mathrm{dom}(f) \cap \{x \in \mathrm{dom}(g) : g(x) \neq 0\}$, and the domain of $g \circ f$ is $\{x \in \mathrm{dom}(f) : f(x) \in \mathrm{dom}(g)\}$. Note $f + g = g + f$ and $fg = gf$ but that in general $f \circ g \neq g \circ f$.

These new functions are continuous if f and g are continuous.

17.4 Theorem.
Let f and g be real-valued functions that are continuous at x_0 in $\mathbb{R}$. Then

(i) *$f + g$ is continuous at x_0;*
(ii) *fg is continuous at x_0;*
(iii) *f/g is continuous at x_0 if $g(x_0) \neq 0$.*

Proof
We are given that $x_0 \in \text{dom}(f) \cap \text{dom}(g)$. Let (x_n) be a sequence in $\text{dom}(f) \cap \text{dom}(g)$ converging to x_0. Then we have $\lim f(x_n) = f(x_0)$ and $\lim g(x_n) = g(x_0)$. Theorem 9.3 shows

$$\begin{aligned}\lim(f+g)(x_n) &= \lim[f(x_n)+g(x_n)] = \lim f(x_n) + \lim g(x_n)\\ &= f(x_0)+g(x_0) = (f+g)(x_0),\end{aligned}$$

so $f+g$ is continuous at x_0. Likewise, Theorem 9.4 implies fg is continuous at x_0.

To handle f/g we assume $g(x_0) \neq 0$ and consider a sequence (x_n) in $\text{dom}(f) \cap \{x \in \text{dom}(g) : g(x) \neq 0\}$ converging to x_0. Then Theorem 9.6 shows

$$\lim \left(\frac{f}{g}\right)(x_n) = \lim \frac{f(x_n)}{g(x_n)} = \frac{f(x_0)}{g(x_0} = \left(\frac{f}{g}\right)(x_0);$$

so f/g is continuous at x_0. ■

17.5 Theorem.
If f is continuous at x_0 and g is continuous at $f(x_0)$, then the composite function $g \circ f$ is continuous at x_0.

Proof
We are given that $x_0 \in \text{dom}(f)$ and $f(x_0) \in \text{dom}(g)$. Let (x_n) be a sequence in $\{x \in \text{dom}(f) : f(x) \in \text{dom}(g)\}$ converging to x_0. Since f is continuous at x_0, we have $\lim f(x_n) = f(x_0)$. Since the sequence $(f(x_n))$ converges to $f(x_0)$ and g is continuous at $f(x_0)$, we also have $\lim g(f(x_n)) = g(f(x_0))$; that is, $\lim g \circ f(x_n) = g \circ f(x_0)$. Hence $g \circ f$ is continuous at x_0. ■

Example 4
For this example, let us accept as known that polynomial functions and the functions $\sin x$, $\cos x$ and e^x are continuous on $\mathbb{R}$. Then $4e^x$ and $|\sin x|$ are continuous on $\mathbb{R}$ by Theorem 17.3. The function $\sin x + 4e^x + x^3$ is continuous on $\mathbb{R}$ by (i) of Theorem 17.4. The function $x^4 \sin x$ is continuous on $\mathbb{R}$ by (ii) of Theorem 17.4, and (iii) of Theorem 17.4 shows $\tan x = \frac{\sin x}{\cos x}$ is continuous wherever $\cos x \neq 0$, i.e., at all x not of the form $n\pi + \frac{\pi}{2}$, $n \in \mathbb{Z}$. Theorem 17.5 tells us $e^{\sin x}$ and $\cos(e^x)$ are continuous on $\mathbb{R}$; for example, $\cos(e^x) =$

$g \circ f(x)$ where $f(x) = e^x$ and $g(x) = \cos x$. Several applications of Theorems 17.3–17.5 will show $x^2 \sin(\frac{1}{x})$ and $\frac{1}{x}\sin(\frac{1}{x^2})$ are continuous at all nonzero x in $\mathbb{R}$. □

Example 5

Let f and g be continuous at x_0 in $\mathbb{R}$. Prove $\max(f, g)$ is continuous at x_0. □

Solution

We could prove this using Definition 17.1 or the ϵ–δ condition in Theorem 17.2. However, we illustrate a useful technique by reducing the problem to results we have already established. First, observe

$$\max(f, g) = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|,$$

which shows that the function $\max(f, g)$ is a combination of functions to which our theorems apply. This equation holds because $\max\{a, b\} = \frac{1}{2}(a + b) + \frac{1}{2}|a - b|$ is true for all $a, b \in \mathbb{R}$, a fact which is easily checked by considering the cases $a \geq b$ and $a < b$. By Theorem 17.4(i), $f + g$ and $f - g$ are continuous at x_0. Hence $|f - g|$ is continuous at x_0 by Theorem 17.3. Then $\frac{1}{2}(f + g)$ and $\frac{1}{2}|f - g|$ are continuous at x_0 by Theorem 17.3, and another application of Theorem 17.4(i) shows $\max(f, g)$ is continuous at x_0. □

Exercises

17.1 Let $f(x) = \sqrt{4 - x}$ for $x \leq 4$ and $g(x) = x^2$ for all $x \in \mathbb{R}$.

- **(a)** Give the domains of $f + g$, fg, $f \circ g$ and $g \circ f$.
- **(b)** Find the values $f \circ g(0)$, $g \circ f(0)$, $f \circ g(1)$, $g \circ f(1)$, $f \circ g(2)$ and $g \circ f(2)$.
- **(c)** Are the functions $f \circ g$ and $g \circ f$ equal?
- **(d)** Are $f \circ g(3)$ and $g \circ f(3)$ meaningful?

17.2 Let $f(x) = 4$ for $x \geq 0$, $f(x) = 0$ for $x < 0$, and $g(x) = x^2$ for all x. Thus $\mathrm{dom}(f) = \mathrm{dom}(g) = \mathbb{R}$.

- **(a)** Determine the following functions: $f + g$, fg, $f \circ g$, $g \circ f$. Be sure to specify their domains.
- **(b)** Which of the functions f, g, $f + g$, fg, $f \circ g$, $g \circ f$ is continuous?

17.3 Accept on faith that the following familiar functions are continuous on their domains: $\sin x$, $\cos x$, e^x, 2^x, $\log_e x$ for $x > 0$, x^p for $x > 0$ [p any real number]. Use these facts and theorems in this section to prove the following functions are also continuous.

(a) $\log_e(1 + \cos^4 x)$

(b) $[\sin^2 x + \cos^6 x]^\pi$

(c) 2^{x^2}

(d) 8^x

(e) $\tan x$ for $x \neq$ odd multiple of $\frac{\pi}{2}$

(f) $x \sin(\frac{1}{x})$ for $x \neq 0$

(g) $x^2 \sin(\frac{1}{x})$ for $x \neq 0$

(h) $\frac{1}{x} \sin(\frac{1}{x^2})$ for $x \neq 0$

17.4 Prove the function $\sqrt{x}$ is continuous on its domain $[0, \infty)$. *Hint*: Apply Example 5 in §8.

17.5 **(a)** Prove that if $m \in \mathbb{N}$, then the function $f(x) = x^m$ is continuous on $\mathbb{R}$.

(b) Prove every *polynomial function* $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is continuous on $\mathbb{R}$.

17.6 A *rational function* is a function f of the form p/q where p and q are polynomial functions. The domain of f is $\{x \in \mathbb{R} : q(x) \neq 0\}$. Prove every rational function is continuous. *Hint*: Use Exercise 17.5.

17.7 **(a)** Observe that if $k \in \mathbb{R}$, then the function $g(x) = kx$ is continuous by Exercise 17.5.

(b) Prove $|x|$ is a continuous function on $\mathbb{R}$.

(c) Use (a) and (b) and Theorem 17.5 to give another proof of Theorem 17.3.

17.8 Let f and g be real-valued functions.

(a) Show $\min(f, g) = \frac{1}{2}(f + g) - \frac{1}{2}|f - g|$.

(b) Show $\min(f, g) = -\max(-f, -g)$.

(c) Use (a) or (b) to prove that if f and g are continuous at x_0 in $\mathbb{R}$, then $\min(f, g)$ is continuous at x_0.

17.9 Prove each of the following functions in continuous at x_0 by verifying the ϵ–δ property of Theorem 17.2.

(a) $f(x) = x^2$, $x_0 = 2$;

(b) $f(x) = \sqrt{x}$, $x_0 = 0$;

(c) $f(x) = x \sin(\frac{1}{x})$ for $x \neq 0$ and $f(0) = 0$, $x_0 = 0$;

(d) $g(x) = x^3$, x_0 arbitrary.
Hint for (d): $x^3 - x_0^3 = (x - x_0)(x^2 + x_0 x + x_0^2)$.

17.10 Prove the following functions are discontinuous at the indicated points. You may use either Definition 17.1 or the ϵ–δ property in Theorem 17.2.

(a) $f(x) = 1$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$, $x_0 = 0$;

(b) $g(x) = \sin(\frac{1}{x})$ for $x \neq 0$ and $g(0) = 0$, $x_0 = 0$;

(c) $\text{sgn}(x) = -1$ for $x < 0$, $\text{sgn}(x) = 1$ for $x > 0$, and $\text{sgn}(0) = 0$, $x_0 = 0$. The function sgn is called the *signum function*; note $\text{sgn}(x) = \frac{x}{|x|}$ for $x \neq 0$.

17.11 Let f be a real-valued function with $\text{dom}(f) \subseteq \mathbb{R}$. Prove f is continuous at x_0 if and only if, for every *monotonic* sequence (x_n) in $\text{dom}(f)$ converging to x_0, we have $\lim f(x_n) = f(x_0)$. *Hint*: Don't forget Theorem 11.4.

17.12 **(a)** Let f be a continuous real-valued function with domain (a, b). Show that if $f(r) = 0$ for each rational number r in (a, b), then $f(x) = 0$ for all $x \in (a, b)$.

(b) Let f and g be continuous real-valued functions on (a, b) such that $f(r) = g(r)$ for each rational number r in (a, b). Prove $f(x) = g(x)$ for all $x \in (a, b)$. *Hint*: Use part (a).

17.13 **(a)** Let $f(x) = 1$ for rational numbers x and $f(x) = 0$ for irrational numbers. Show f is discontinuous at every x in $\mathbb{R}$.

(b) Let $h(x) = x$ for rational numbers x and $h(x) = 0$ for irrational numbers. Show h is continuous at $x = 0$ and at no other point.

17.14 For each nonzero rational number x, write x as $\frac{p}{q}$ where p, q are integers with no common factors and $q > 0$, and then define $f(x) = \frac{1}{q}$. Also define $f(0) = 1$ and $f(x) = 0$ for all $x \in \mathbb{R} \setminus \mathbb{Q}$. Thus $f(x) = 1$ for each integer, $f(\frac{1}{2}) = f(-\frac{1}{2}) = f(\frac{3}{2}) = \cdots = \frac{1}{2}$, etc. Show f is continuous at each point of $\mathbb{R} \setminus \mathbb{Q}$ and discontinuous at each point of $\mathbb{Q}$.

17.15 Let f be a real-valued function whose domain is a subset of $\mathbb{R}$. Show f is continuous at x_0 in $\mathrm{dom}(f)$ if and only if, for every sequence (x_n) in $\mathrm{dom}(f) \setminus \{x_0\}$ converging to x_0, we have $\lim f(x_n) = f(x_0)$.

17.16 The postage-stamp function P is defined by $P(x) = A$ for $0 \leq x < 1$ and $P(x) = A + Bn$ for $n \leq x < n+1$. The definition of P means that P takes the value A on the interval $[0,1)$, the value $A+B$ on the interval $[1,2)$, the value $A+2B$ on the interval $[2,3)$, etc. Here A is postage needed for the first ounce, and B is the postage needed for each additional ounce. Show P is discontinuous at every positive integer. Because postage rates tend to increase over time, A and B are actually functions.

§18 Properties of Continuous Functions

A real-valued function f is said to be *bounded* if $\{f(x) : x \in \mathrm{dom}(f)\}$ is a bounded set, i.e., if there exists a real number M such that $|f(x)| \leq M$ for all $x \in \mathrm{dom}(f)$.

18.1 Theorem.
Let f be a continuous real-valued function on a closed interval $[a,b]$. Then f is a bounded function. Moreover, f assumes its maximum and minimum values on $[a,b]$; that is, there exist x_0, y_0 in $[a,b]$ such that $f(x_0) \leq f(x) \leq f(y_0)$ for all $x \in [a,b]$.

Proof
Assume f is not bounded on $[a,b]$. Then to each $n \in \mathbb{N}$ there corresponds an $x_n \in [a,b]$ such that $|f(x_n)| > n$. By the Bolzano-Weierstrass Theorem 11.5, (x_n) has a subsequence (x_{n_k}) that converges to some real number x_0. The number x_0 also must belong to the closed interval $[a,b]$, as noted in Exercise 8.9. Since f is continuous at x_0, we have $\lim_{k\to\infty} f(x_{n_k}) = f(x_0)$, but we also have $\lim_{k\to\infty} |f(x_{n_k})| = +\infty$, which is a contradiction. It follows that f is bounded.

Now let $M = \sup\{f(x) : x \in [a,b]\}$; M is finite by the preceding paragraph. For each $n \in \mathbb{N}$ there exists $y_n \in [a,b]$ such that $M - \frac{1}{n} < f(y_n) \leq M$. Hence we have $\lim f(y_n) = M$. By the Bolzano-Weierstrass theorem, there is a subsequence (y_{n_k}) of (y_n) converging to a limit y_0 in $[a,b]$. Since f is continuous at y_0, we have $f(y_0) = \lim_{k\to\infty} f(y_{n_k})$. Since $(f(y_{n_k}))_{k\in\mathbb{N}}$ is a subsequence of $(f(y_n))_{n\in\mathbb{N}}$, Theorem 11.3 shows $\lim_{k\to\infty} f(y_{n_k}) = \lim_{n\to\infty} f(y_n) = M$ and therefore $f(y_0) = M$. Thus f assumes its maximum at y_0.

The last paragraph applies to the function $-f$, so $-f$ assumes its maximum at some $x_0 \in [a, b]$. It follows easily that f assumes its minimum at x_0; see Exercise 18.1. ■

Theorem 18.1 is used all the time, at least implicitly, in solving maximum-minimum problems in calculus because it is taken for granted that the problems have solutions, namely that a continuous function on a closed interval actually takes on a maximum and a minimum. If the domain is not a closed interval, one needs to be careful; see Exercise 18.3.

Theorem 18.1 is false if the closed interval $[a, b]$ is replaced by an open interval. For example, $f(x) = \frac{1}{x}$ is continuous but unbounded on $(0, 1)$. The function x^2 is continuous and bounded on $(-1, 1)$, but it does not have a maximum value on $(-1, 1)$.

18.2 Intermediate Value Theorem.
If f is a continuous real-valued function on an interval I, then f has the intermediate value property on I: Whenever $a, b \in I$, $a < b$ and y lies between $f(a)$ and $f(b)$ [i.e., $f(a) < y < f(b)$ or $f(b) < y < f(a)$], there exists at least one x in (a, b) such that $f(x) = y$.

Proof
We assume $f(a) < y < f(b)$; the other case is similar. Let $S = \{x \in [a, b] : f(x) < y\}$; see Fig. 18.1. Since a belongs to S, the set S is nonempty, so $x_0 = \sup S$ represents a number in $[a, b]$. For each $n \in \mathbb{N}$, $x_0 - \frac{1}{n}$ is not an upper bound for S, so there exists $s_n \in S$ such that $x_0 - \frac{1}{n} < s_n \le x_0$. Thus $\lim s_n = x_0$ and, since $f(s_n) < y$ for all n, we have

$$f(x_0) = \lim f(s_n) \le y. \tag{1}$$

Let $t_n = \min\{b, x_0 + \frac{1}{n}\}$. Since $x_0 \le t_n \le x_0 + \frac{1}{n}$ we have $\lim t_n = x_0$. Each t_n belongs to $[a, b]$ but not to S, so $f(t_n) \ge y$ for all n. Therefore we have

$$f(x_0) = \lim f(t_n) \ge y;$$

this and (1) imply $f(x_0) = y$. ■

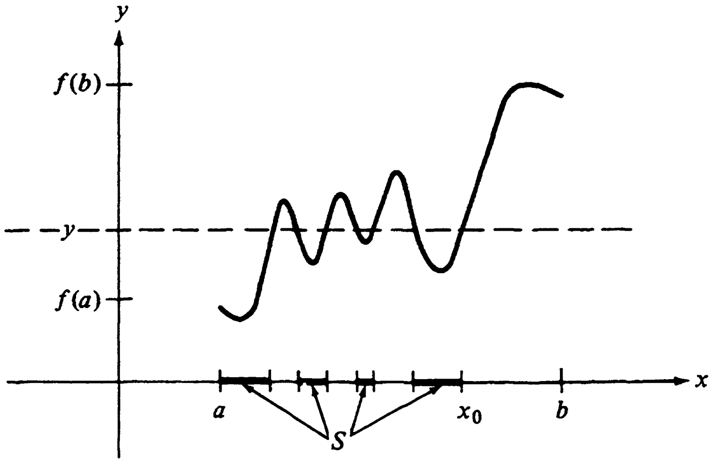

FIGURE 18.1

18.3 Corollary.
If f is a continuous real-valued function on an interval I, then the set $f(I) = \{f(x) : x \in I\}$ is also an interval or a single point.

Proof
The set $J = f(I)$ has the property:

$$y_0, y_1 \in J \quad \text{and} \quad y_0 < y < y_1 \quad \text{imply} \quad y \in J. \tag{1}$$

If $\inf J < \sup J$, then such a set J will be an interval. In fact, we will show

$$\inf J < y < \sup J \quad \text{implies} \quad y \in J, \tag{2}$$

so J is an interval with endpoints $\inf J$ and $\sup J$; $\inf J$ and $\sup J$ may or may not belong to J and they may or may not be finite.

To prove (2) from (1), consider $\inf J < y < \sup J$. Then there exist y_0, y_1 in J so that $y_0 < y < y_1$. Thus $y \in J$ by (1). ■

Example 1
Let f be a continuous function mapping $[0,1]$ into $[0,1]$. In other words, $\text{dom}(f) = [0,1]$ and $f(x) \in [0,1]$ for all $x \in [0,1]$. Show f has a *fixed point*, i.e., a point $x_0 \in [0,1]$ such that $f(x_0) = x_0$; x_0 is left "fixed" by f. □

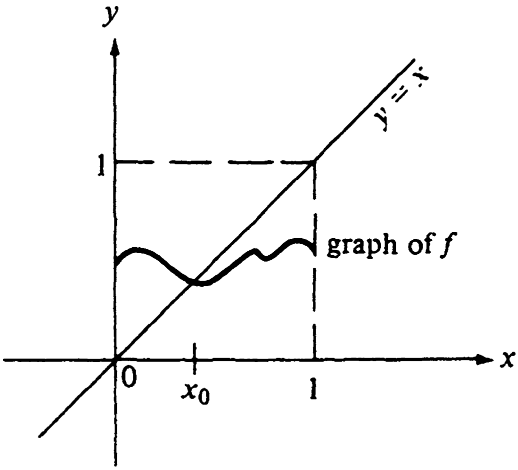

FIGURE 18.2

Solution

The graph of f lies in the unit square; see Fig. 18.2. Our assertion is equivalent to the assertion that the graph of f crosses the $y = x$ line, which is almost obvious.

A rigorous proof involves a little trick. Consider $g(x) = f(x) - x$ which is also a continuous function on $[0, 1]$. Since $g(0) = f(0) - 0 = f(0) \geq 0$ and $g(1) = f(1) - 1 \leq 1 - 1 = 0$, the Intermediate Value theorem shows $g(x_0) = 0$ for some $x_0 \in [0, 1]$. Then obviously we have $f(x_0) = x_0$. □

Example 2

Show that if $y > 0$ and $m \in \mathbb{N}$, then y has a positive mth root. □

Solution

The function $f(x) = x^m$ is continuous [Exercise 17.5]. There exists $b > 0$ so that $y \leq b^m$; in fact, if $y \leq 1$ let $b = 1$ and if $y > 1$ let $b = y$. Thus $f(0) < y \leq f(b)$ and the Intermediate Value theorem implies $f(x) = y$ for some x in $(0, b]$. So $y = x^m$ and x is an mth root of y. □

Let us analyze the function $f(x) = x^m$ in Example 2 more closely. It is a *strictly increasing function* on $[0, \infty)$:

$$0 \leq x_1 < x_2 \quad \text{implies} \quad x_1^m < x_2^m.$$

Therefore f is one-to-one and each nonnegative y has exactly one nonnegative mth root. This assures us that the notation $y^{1/m}$ is unambiguous. In fact, $f^{-1}(y) = y^{1/m}$ is the *inverse function* of f since $f^{-1} \circ f(x) = x$ for $x \in \mathrm{dom}(f)$ and $f \circ f^{-1}(y) = y$ for $y \in \mathrm{dom}(f^{-1})$. Since $f(x) = x^m$ is continuous, the function $f^{-1}(y) = y^{1/m}$ is continuous on $[0, \infty)$ by the next theorem. Note that for $m = 2$ this result appears in Exercise 17.4.

18.4 Theorem.
Let f be a continuous strictly increasing function on some interval I. Then $f(I)$ is an interval J by Corollary 18.3 and f^{-1} represents a function with domain J. The function f^{-1} is a continuous strictly increasing function on J.

Proof
The function f^{-1} is easily shown to be strictly increasing. Since f^{-1} maps J onto I, the next theorem shows f^{-1} is continuous. ■

18.5 Theorem.
Let g be a strictly increasing function on an interval J such that $g(J)$ is an interval I. Then g is continuous on J.

Proof
Consider x_0 in J. We assume x_0 is not an endpoint of J; tiny changes in the proof are needed otherwise. Then $g(x_0)$ is not an endpoint of I, so there exists $\epsilon_0 > 0$ such that $(g(x_0) - \epsilon_0, g(x_0) + \epsilon_0) \subseteq I$.

Let $\epsilon > 0$. Since we only need to verify the ϵ–δ property of Theorem 17.2 for small ϵ, we may assume $\epsilon < \epsilon_0$. Then there exist $x_1, x_2 \in J$ such that $g(x_1) = g(x_0) - \epsilon$ and $g(x_2) = g(x_0) + \epsilon$. Clearly we have $x_1 < x_0 < x_2$. Also, if $x_1 < x < x_2$, then $g(x_1) < g(x) < g(x_2)$, hence $g(x_0) - \epsilon < g(x) < g(x_0) + \epsilon$, and hence $|g(x) - g(x_0)| < \epsilon$. Now if we put $\delta = \min\{x_2 - x_0, x_0 - x_1\}$, then $|x - x_0| < \delta$ implies $x_1 < x < x_2$ and hence $|g(x) - g(x_0)| < \epsilon$. ■

Theorem 18.5 provides a partial converse to the Intermediate Value theorem, since it tells us that a strictly increasing function with the intermediate value property is continuous. However, Exercise 18.12 shows that a function can have the intermediate value property without being continuous.

18.6 Theorem.
Let f be a one-to-one continuous function on an interval I. Then f is strictly increasing $[x_1 < x_2$ implies $f(x_1) < f(x_2)]$ or strictly decreasing $[x_1 < x_2$ implies $f(x_1) > f(x_2)]$.

Proof
First we show

$$\text{if } a < b < c \text{ in } I, \text{ then } f(b) \text{ lies between } f(a) \text{ and } f(c). \tag{1}$$

If not, then $f(b) > \max\{f(a), f(c)\}$, say. Select y so that $f(b) > y > \max\{f(a), f(c)\}$. By the Intermediate Value Theorem 18.2 applied to $[a, b]$ and to $[b, c]$, there exist $x_1 \in (a, b)$ and $x_2 \in (b, c)$ such that $f(x_1) = f(x_2) = y$. This contradicts the one-to-one property of f.

Now select any $a_0 < b_0$ in I and suppose, say, that $f(a_0) < f(b_0)$. We will show f is strictly increasing on I. By (1) we have

$$\begin{array}{llll} f(x) < f(a_0) & \text{for} \quad x < a_0 & [\text{since} \quad x < a_0 < b_0], \\ f(a_0) < f(x) < f(b_0) & \text{for} \quad a_0 < x < b_0, & \\ f(b_0) < f(x) & \text{for} \quad x > b_0 & [\text{since} \quad a_0 < b_0 < x]. \end{array}$$

In particular,

$$f(x) < f(a_0) \quad \text{for all} \quad x < a_0, \tag{2}$$

$$f(a_0) < f(x) \quad \text{for all} \quad x > a_0. \tag{3}$$

Consider any $x_1 < x_2$ in I. If $x_1 \le a_0 \le x_2$, then $f(x_1) < f(x_2)$ by (2) and (3). If $x_1 < x_2 < a_0$, then $f(x_1) < f(a_0)$ by (2), so by (1) we have $f(x_1) < f(x_2)$. Finally, if $a_0 < x_1 < x_2$, then $f(a_0) < f(x_2)$, so that $f(x_1) < f(x_2)$. ■

Exercises

18.1 Let f be as in Theorem 18.1. Show that if $-f$ assumes its maximum at $x_0 \in [a, b]$, then f assumes its minimum at x_0.

18.2 Reread the proof of Theorem 18.1 with $[a, b]$ replaced by (a, b). Where does it break down? Discuss.

18.3 Use calculus to find the maximum and minimum of $f(x) = x^3 - 6x^2 + 9x + 1$ on $[0, 5)$.

18.4 Let $S \subseteq \mathbb{R}$ and suppose there exists a sequence (x_n) in S converging to a number $x_0 \notin S$. Show there exists an unbounded continuous function on S.

18.5 **(a)** Let f and g be continuous functions on $[a, b]$ such that $f(a) \geq g(a)$ and $f(b) \leq g(b)$. Prove $f(x_0) = g(x_0)$ for at least one x_0 in $[a, b]$.

(b) Show Example 1 can be viewed as a special case of part (a).

18.6 Prove $x = \cos x$ for some x in $(0, \frac{\pi}{2})$.

18.7 Prove $xe^x = 2$ for some x in $(0, 1)$.

18.8 Suppose f is a real-valued continuous function on $\mathbb{R}$ and $f(a)f(b) < 0$ for some $a, b \in \mathbb{R}$. Prove there exists x between a and b such that $f(x) = 0$.

18.9 Prove that a polynomial function f of odd degree has at least one real root. *Hint*: It may help to consider first the case of a cubic, i.e., $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ where $a_3 \neq 0$.

18.10 Suppose f is continuous on $[0, 2]$ and $f(0) = f(2)$. Prove there exist x, y in $[0, 2]$ such that $|y - x| = 1$ and $f(x) = f(y)$. *Hint*: Consider $g(x) = f(x + 1) - f(x)$ on $[0, 1]$.

18.11 **(a)** Show that if f is strictly increasing on an interval I, then $-f$ is strictly decreasing on I.

(b) State and prove Theorems 18.4 and 18.5 for strictly decreasing functions.

18.12 Let $f(x) = \sin(\frac{1}{x})$ for $x \neq 0$ and let $f(0) = 0$.

(a) Observe that f is discontinuous at 0 by Exercise 17.10(b).

(b) Show f has the intermediate value property on $\mathbb{R}$.

§19 Uniform Continuity

Let f be a real-valued function whose domain is a subset of $\mathbb{R}$. Theorem 17.2 tells us that f is continuous on a set $S \subseteq \text{dom}(f)$ if and only if

$$\begin{gathered}\text{for each } x_0 \in S \text{ and } \epsilon > 0 \text{ there is } \delta > 0 \text{ so that} \\ x \in \text{dom}(f) \text{ and } |x - x_0| < \delta \text{ imply } |f(x) - f(x_0)| < \epsilon.\end{gathered} \qquad (*)$$

The choice of δ depends on $\epsilon > 0$ and on the point x_0 in S.

Example 1
We verify (*) for the function $f(x) = \frac{1}{x^2}$ on $(0, \infty)$. Let $x_0 > 0$ and $\epsilon > 0$. We need to show $|f(x) - f(x_0)| < \epsilon$ for $|x - x_0|$ sufficiently small. Note that

$$f(x) - f(x_0) = \frac{1}{x^2} - \frac{1}{x_0^2} = \frac{x_0^2 - x^2}{x^2 x_0^2} = \frac{(x_0 - x)(x_0 + x)}{x^2 x_0^2}. \qquad (1)$$

If $|x - x_0| < \frac{x_0}{2}$, then we have $|x| > \frac{x_0}{2}$, $|x| < \frac{3x_0}{2}$ and $|x_0 + x| < \frac{5x_0}{2}$. These observations and (1) show that if $|x - x_0| < \frac{x_0}{2}$, then

$$|f(x) - f(x_0)| < \frac{|x_0 - x| \cdot \frac{5x_0}{2}}{(\frac{x_0}{2})^2 x_0^2} = \frac{10|x_0 - x|}{x_0^3}.$$

Thus if we set $\delta = \min\{\frac{x_0}{2}, \frac{x_0^3 \epsilon}{10}\}$, then

$$|x - x_0| < \delta \quad \text{implies} \quad |f(x) - f(x_0)| < \epsilon.$$

This establishes (*) for f on $(0, \infty)$. Note that δ depends on both ϵ and x_0. Even if ϵ is fixed, δ gets small when x_0 is small. This shows that *our* choice of δ definitely depends on x_0 as well as ϵ, though this may be because we obtained δ via sloppy estimates. As a matter of fact, in this case δ *must* depend on x_0 as well as ϵ; see Example 3. Figure 19.1 shows how a fixed ϵ requires smaller and smaller δ as x_0 approaches 0. [In the figure, δ_1 signifies a δ that works for x_1 and ϵ, δ_2 signifies a δ that works for x_2 and ϵ, etc.] □

It turns out to be very useful to know when the δ in condition (*) can be chosen to depend only on $\epsilon > 0$ and S, so that δ does not depend on the particular point x_0. Such functions are said to be uniformly continuous on S. In the definition, the points x and x_0 play a symmetric role, so we will call them x and y.

19.1 Definition.
Let f be a real-valued function defined on a set $S \subseteq \mathbb{R}$. Then f is *uniformly continuous on S* if

$$\begin{gathered}\text{for each } \epsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ x, y \in S \text{ and } |x - y| < \delta \text{ imply } |f(x) - f(y)| < \epsilon.\end{gathered} \qquad (1)$$

We will say f is *uniformly continuous* if f is uniformly continuous on $\mathrm{dom}(f)$.

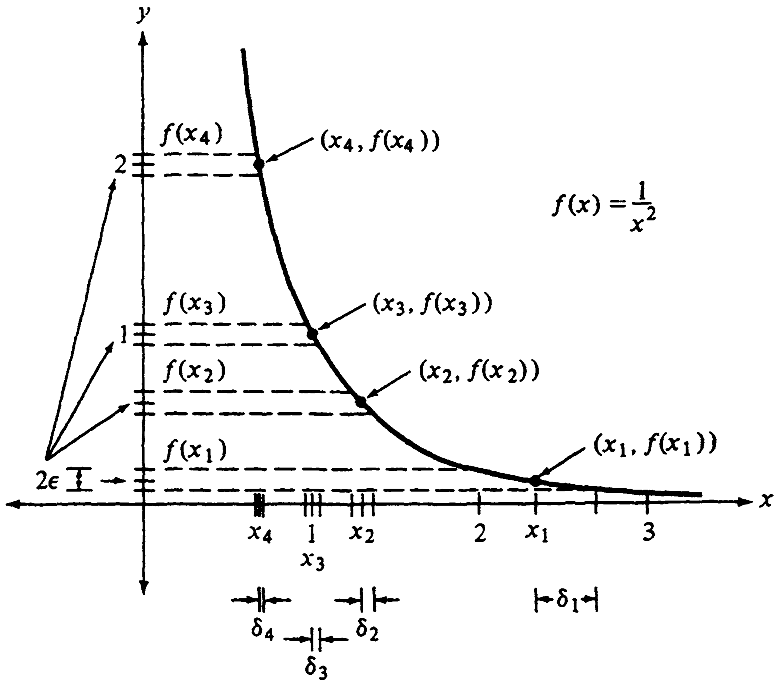

FIGURE 19.1

Note that if a function is uniformly continuous on its domain, then it is continuous on its domain. This should be obvious; if it isn't, Theorem 17.2 and Definition 19.1 should be carefully scrutinized. Note also that uniform continuity is a property concerning a function *and a set* [on which it is defined]. It makes no sense to speak of a function being uniformly continuous at each point.

Example 2

We show $f(x) = \frac{1}{x^2}$ is uniformly continuous on any set of the form $[a, \infty)$ where $a > 0$. Here a is fixed. Let $\epsilon > 0$. We need to show there exists $\delta > 0$ such that

$$x \geq a, \quad y \geq a \quad \text{and} \quad |x - y| < \delta \quad \text{imply} \quad |f(x) - f(y)| < \epsilon. \quad (1)$$

As in formula (1) of Example 1, we have

$$f(x) - f(y) = \frac{(y - x)(y + x)}{x^2 y^2}.$$

If we can show $\frac{y+x}{x^2y^2}$ is bounded on $[a,\infty)$ by a constant M, then we will take $\delta = \frac{\epsilon}{M}$. But we have

$$\frac{y+x}{x^2y^2} = \frac{1}{x^2y} + \frac{1}{xy^2} \le \frac{1}{a^3} + \frac{1}{a^3} = \frac{2}{a^3},$$

so we set $\delta = \frac{\epsilon a^3}{2}$. It is now straightforward to verify (1). In fact, $x \ge a$, $y \ge a$ and $|x-y| < \delta$ imply

$$|f(x) - f(y)| = \frac{|y-x|(y+x)|}{x^2y^2} < \delta\left(\frac{1}{x^2y} + \frac{1}{xy^2}\right) \le \frac{2\delta}{a^3} = \epsilon.$$

We have shown f is uniformly continuous on $[a,\infty)$ since δ depends only on ϵ and the set $[a,\infty)$. $\square$

Example 3

The function $f(x) = \frac{1}{x^2}$ is not uniformly continuous on the set $(0,\infty)$ or even on the set $(0,1)$.

We will prove this by directly violating the definition of uniform continuity. The squeamish reader may skip this demonstration and wait for the easy proof in Example 6. We will show (1) in Definition 19.1 fails for $\epsilon = 1$; that is

$$\begin{gathered}\text{for each } \delta > 0 \text{ there exist } x, y \text{ in } (0,1) \text{ such that}\\ |x-y| < \delta \text{ and yet } |f(x) - f(y)| \ge 1.\end{gathered} \tag{1}$$

[Actually, for this function, (1) in Definition 19.1 fails for all $\epsilon > 0$.] To show (1) it suffices to take $y = x + \frac{\delta}{2}$ and arrange for

$$\left| f(x) - f\left(x + \frac{\delta}{2}\right)\right| \ge 1. \tag{2}$$

[The motivation for this maneuver is to go from two variables, x and y, in (1) to one variable, x, in (2).] By (1) in Example 1, (2) is equivalent to

$$1 \le \frac{(x + \frac{\delta}{2} - x)(x + \frac{\delta}{2} + x)}{x^2(x+\frac{\delta}{2})^2} = \frac{\delta(2x + \frac{\delta}{2})}{2x^2(x+\frac{\delta}{2})^2}. \tag{3}$$

It suffices to prove (1) for $\delta < \frac{1}{2}$. To apply (3), let us try $x = \delta$ for no particular reason. Then

$$\frac{\delta(2\delta + \frac{\delta}{2})}{2\delta^2(\delta + \frac{\delta}{2})^2} = \frac{5\delta^2/2}{9\delta^4/2} = \frac{5}{9\delta^2} > \frac{5}{9(\frac{1}{2})^2} = \frac{20}{9} > 1.$$

We were lucky! To summarize, we have shown that if $0 < \delta < \frac{1}{2}$, then $|f(\delta) - f(\delta + \frac{\delta}{2})| > 1$, so (1) holds with $x = \delta$ and $y = \delta + \frac{\delta}{2}$. □

Example 4

Is the function $f(x) = x^2$ uniformly continuous on $[-7, 7]$? To see if it is, consider $\epsilon > 0$. Note that $|f(x)-f(y)| = |x^2-y^2| = |x-y|\cdot|x+y|$. Since $|x + y| \leq 14$ for x, y in $[-7, 7]$, we have

$$|f(x) - f(y)| \leq 14|x - y| \quad \text{for} \quad x, y \in [-7, 7].$$

Thus if $\delta = \frac{\epsilon}{14}$, then

$$x, y \in [-7, 7] \quad \text{and} \quad |x - y| < \delta \quad \text{imply} \quad |f(x) - f(y)| < \epsilon.$$

We have shown that f is uniformly continuous on $[-7, 7]$. A similar proof would work for $f(x) = x^2$ on any closed interval. However, these results are not accidents as the next important theorem shows. □

19.2 Theorem.

If f is continuous on a closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

Proof

Assume f is not uniformly continuous on $[a, b]$. Then there exists $\epsilon > 0$ such that for each $\delta > 0$ the implication

$$\text{“}|x - y| < \delta \quad \text{implies} \quad |f(x) - f(y)| < \epsilon\text{”}$$

fails. That is, for each $\delta > 0$ there exist $x, y \in [a, b]$ such that $|x-y| < \delta$ and yet $|f(x) - f(y)| \geq \epsilon$. Thus for each $n \in \mathbb{N}$ there exist x_n, y_n in $[a, b]$ such that $|x_n - y_n| < \frac{1}{n}$ and yet $|f(x_n) - f(y_n)| \geq \epsilon$. By the Bolzano-Weierstrass Theorem 11.5, a subsequence (x_{n_k}) of (x_n) converges. Moreover, if $x_0 = \lim_{k\to\infty} x_{n_k}$, then x_0 belongs to $[a, b]$; see Exercise 8.9. Clearly we also have $x_0 = \lim_{k\to\infty} y_{n_k}$. Since f is continuous at x_0, we have

$$f(x_0) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(y_{n_k}),$$

so

$$\lim_{k\to\infty} [f(x_{n_k}) - f(y_{n_k})] = 0.$$

Since $|f(x_{n_k}) - f(y_{n_k})| \geq \epsilon$ for all k, we have a contradiction. We conclude f is uniformly continuous on $[a, b]$. ■

The preceding proof used only two properties of $[a, b]$:

(a) Boundedness, so the Bolzano-Weierstrass theorem applies;
(b) A convergent sequence in $[a, b]$ converges to an element in $[a, b]$.

As noted prior to Theorem 11.9, sets with property (b) are called *closed sets*. Hence Theorem 19.2 has the following generalization. *If f is continuous on a closed and bounded set S, then f is uniformly continuous on S.* See also Theorems 21.4 and 13.12 that appear in enrichment sections.

Example 5
In view of Theorem 19.2, the following functions are uniformly continuous on the indicated sets: x^{73} on $[-13, 13]$, $\sqrt{x}$ on $[0, 400]$, $x^{17}\sin(e^x) - e^{4x}\cos 2x$ on $[-8\pi, 8\pi]$, and $\frac{1}{x^6}$ on $[\frac{1}{4}, 44]$. $\square$

19.3 Discussion.
Example 5 illustrates the power of Theorem 19.2, but it still may not be clear why uniform continuity is worth studying. One of the important applications of uniform continuity concerns the integrability of continuous functions on closed intervals. To see the relevance of uniform continuity, consider a continuous nonnegative real-valued function f on $[0, 1]$. For $n \in \mathbb{N}$ and $i = 0, 1, 2, \ldots, n-1$, let

$$M_{i,n} = \sup\left\{f(x) : x \in \left[\tfrac{i}{n}, \tfrac{i+1}{n}\right]\right\} \quad \text{and} \quad m_{i,n} = \inf\left\{f(x) : x \in \left[\tfrac{i}{n}, \tfrac{i+1}{n}\right]\right\}.$$

Then the sum of the areas of the rectangles in Fig. 19.2a equals

$$U_n = \frac{1}{n}\sum_{i=0}^{n-1} M_{i,n}$$

and the sum of the areas of the rectangles in Fig. 19.2b equals

$$L_n = \frac{1}{n}\sum_{i=0}^{n-1} m_{i,n}.$$

The function f would turn out to be Riemann integrable provided the numbers U_n and L_n are close together for large n, i.e., if

$$\lim_{n\to\infty}(U_n - L_n) = 0; \tag{1}$$

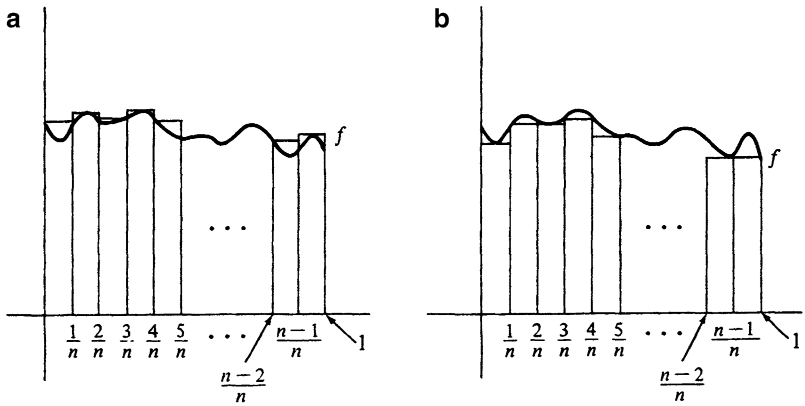

FIGURE 19.2

see Exercise 32.6. Moreover, we would have $\int_0^1 f(x)dx = \lim U_n = \lim L_n$. Relation (1) may appear obvious from Fig. 19.2, but uniform continuity is needed to prove it. First note

$$0 \le U_n - L_n = \frac{1}{n}\sum_{i=0}^{n-1}(M_{i,n} - m_{i,n})$$

for all n. Let $\epsilon > 0$. By Theorem 19.2, f is uniformly continuous on $[0,1]$, so there exists $\delta > 0$ such that

$$x, y \in [0,1] \quad \text{and} \quad |x-y| < \delta \quad \text{imply} \quad |f(x) - f(y)| < \epsilon. \tag{2}$$

Select N so that $\frac{1}{N} < \delta$. Consider $n > N$; for $i = 0, 1, 2, \ldots, n-1$, Theorem 18.1 shows there exist x_i, y_i in $[\frac{i}{n}, \frac{i+1}{n}]$ satisfying $f(x_i) = m_{i,n}$ and $f(y_i) = M_{i,n}$. Since $|x_i - y_i| \le \frac{1}{n} < \frac{1}{N} < \delta$, (2) shows

$$M_{i,n} - m_{i,n} = f(y_i) - f(x_i) < \epsilon,$$

so

$$0 \le U_n - L_n = \frac{1}{n}\sum_{i=0}^{n-1}(M_{i,n} - m_{i,n}) < \frac{1}{n}\sum_{i=0}^{n-1}\epsilon = \epsilon.$$

This proves (1) as desired. □

The next two theorems show uniformly continuous functions have nice properties.

19.4 Theorem.
If f is uniformly continuous on a set S and (s_n) is a Cauchy sequence in S, then $(f(s_n))$ is a Cauchy sequence.

Proof
Let (s_n) be a Cauchy sequence in S and let $\epsilon > 0$. Since f is uniformly continuous on S, there exists $\delta > 0$ so that

$$x, y \in S \quad \text{and} \quad |x - y| < \delta \quad \text{imply} \quad |f(x) - f(y)| < \epsilon. \tag{1}$$

Since (s_n) is a Cauchy sequence, there exists N so that

$$m, n > N \quad \text{implies} \quad |s_n - s_m| < \delta.$$

From (1) we see that

$$m, n > N \quad \text{implies} \quad |f(s_n) - f(s_m)| < \epsilon.$$

This proves $(f(s_n))$ is also a Cauchy sequence. ■

Example 6
We show $f(x) = \frac{1}{x^2}$ is not uniformly continuous on $(0, 1)$. Let $s_n = \frac{1}{n}$ for $n \in \mathbb{N}$. Then (s_n) is obviously a Cauchy sequence in $(0, 1)$. Since $f(s_n) = n^2$, $(f(s_n))$ is not a Cauchy sequence. Therefore f cannot be uniformly continuous on $(0, 1)$ by Theorem 19.4. □

The next theorem involves extensions of functions. We say a function $\tilde{f}$ is an *extension of a function* f if

$$\text{dom}(f) \subseteq \text{dom}(\tilde{f}) \quad \text{and} \quad f(x) = \tilde{f}(x) \quad \text{for all} \quad x \in \text{dom}(f).$$

Example 7
Let $f(x) = x \sin(\frac{1}{x})$ for $x \in (0, \frac{1}{\pi}]$. The function defined by

$$\tilde{f}(x) = \begin{cases} x \sin(\frac{1}{x}) & \text{for} \quad 0 < x \leq \frac{1}{\pi} \\ 0 & \text{for} \quad x = 0 \end{cases}$$

is an extension of f. Note that $\text{dom}(f) = (0, \frac{1}{\pi}]$ and $\text{dom}(\tilde{f}) = [0, \frac{1}{\pi}]$. In this case, $\tilde{f}$ is a continuous extension of f. See Fig. 19.3 as well as Exercises 17.3(f) and 17.9(c). □

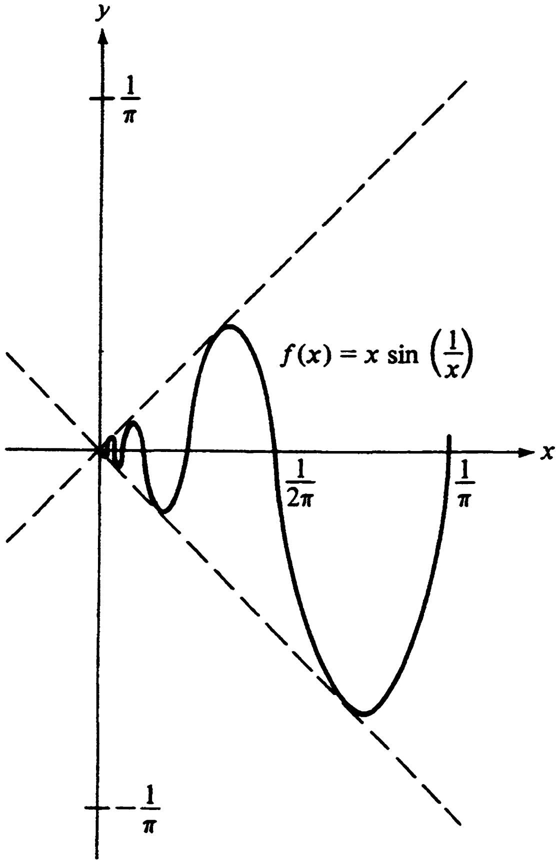

FIGURE 19.3

Example 8
Let $g(x) = \sin(\frac{1}{x})$ for $x \in (0, \frac{1}{\pi}]$. The function g can be extended to a function $\tilde{g}$ with domain $[0, \frac{1}{\pi}]$ in many ways, but $\tilde{g}$ will not be continuous. See Fig. 19.4. □

The function f in Example 7 is uniformly continuous [since $\tilde{f}$ is], and f extends to a continuous function on the closed interval. The function g in Example 8 does not extend to a continuous function on the closed interval, and it turns out that g is not uniformly continuous. These examples illustrate the next theorem.

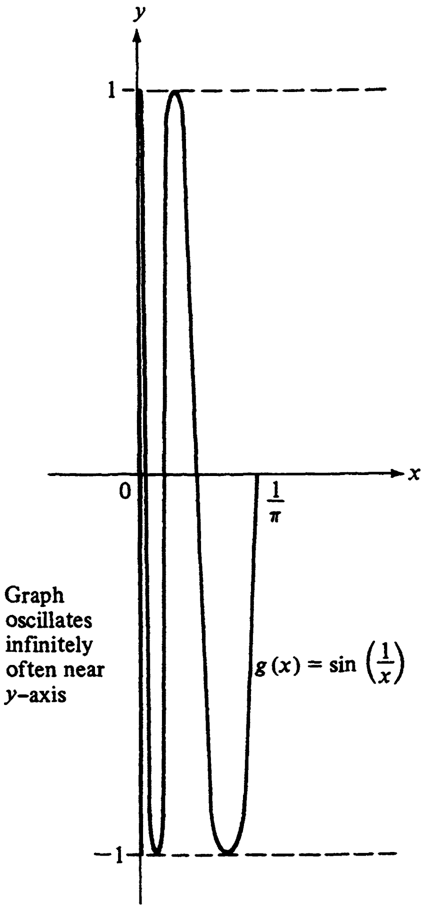

FIGURE 19.4

19.5 Theorem.

A real-valued function f on (a, b) is uniformly continuous on (a, b) if and only if it can be extended to a continuous function $\tilde{f}$ on $[a, b]$.

Proof

First suppose f can be extended to a continuous function $\tilde{f}$ on $[a, b]$. Then $\tilde{f}$ is uniformly continuous on $[a, b]$ by Theorem 19.2, so clearly f is uniformly continuous on (a, b).

Suppose now that f is uniformly continuous on (a, b). We need to define $\tilde{f}(a)$ and $\tilde{f}(b)$ so that the extended function will be continuous.

It suffices for us to deal with $\tilde{f}(a)$. We make two claims:

$$\begin{gathered}\text{if } (s_n) \text{ is a sequence in } (a,b) \text{ converging} \\ \text{to } a, \text{ then } (f(s_n)) \text{ converges,}\end{gathered} \tag{1}$$

and

$$\begin{gathered}\text{if } (s_n) \text{ and } (t_n) \text{ are sequences in } (a,b) \text{ converging} \\ \text{to } a, \text{ then } \lim f(s_n) = \lim f(t_n).\end{gathered} \tag{2}$$

Momentarily accepting (1) and (2) as valid, we define

$$\begin{gathered}\tilde{f}(a) = \lim f(s_n) \text{ for any sequence} \\ (s_n) \text{ in } (a,b) \text{ converging to } a.\end{gathered} \tag{3}$$

Assertion (1) guarantees the limit exists, and assertion (2) guarantees this definition is unambiguous. The continuity of $\tilde{f}$ at a follows directly from (3); see Exercise 17.15.

To prove (1), note that (s_n) is a Cauchy sequence, so $(f(s_n))$ is also a Cauchy sequence by Theorem 19.4. Hence $(f(s_n))$ converges by Theorem 10.11. To prove (2) we create a third sequence (u_n) such that (s_n) and (t_n) are both subsequences of (u_n). In fact, we simply interleaf (s_n) and (t_n):

$$(u_n)_{n=1}^{\infty} = (s_1, t_1, s_2, t_2, s_3, t_3, s_4, t_4, s_5, t_5, \ldots).$$

It is evident that $\lim u_n = a$, so $\lim f(u_n)$ exists by (1). Theorem 11.3 shows the subsequences $(f(s_n))$ and $(f(t_n))$ of $(f(u_n))$ both converge to $\lim f(u_n)$, so $\lim f(s_n) = \lim f(t_n)$. ■

Example 9
Let $h(x) = \frac{\sin x}{x}$ for $x \neq 0$. The function $\tilde{h}$ defined on $\mathbb{R}$ by

$$\tilde{h}(x) = \begin{cases} \frac{\sin x}{x} & \text{for} \quad x \neq 0 \\ 1 & \text{for} \quad x = 0 \end{cases}$$

is an extension of h. Clearly h and $\tilde{h}$ are continuous at all $x \neq 0$. It turns out that $\tilde{h}$ is continuous at $x = 0$ [see below], so h is uniformly continuous on $(a, 0)$ and $(0, b)$ for any $a < 0 < b$ by Theorem 19.5. In fact, $\tilde{h}$ is uniformly continuous on $\mathbb{R}$ [Exercise 19.11].

We cannot prove the continuity of $\tilde{h}$ at 0 in this book because we do not give a definition of $\sin x$. The continuity of $\tilde{h}$ at 0 reflects

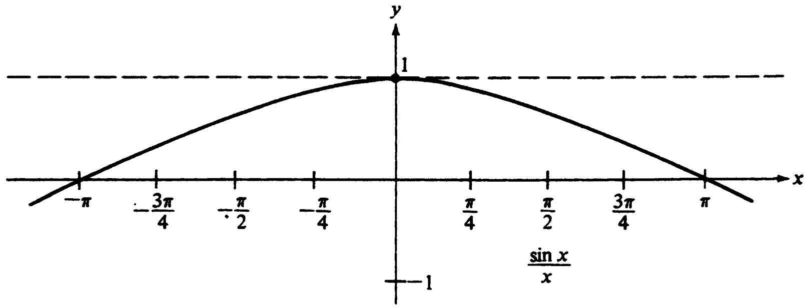

FIGURE 19.5

the fact that $\sin x$ is differentiable at 0 and its derivative there is $\cos(0) = 1$, i.e.,

$$1 = \lim_{x \to 0} \frac{\sin x - \sin 0}{x - 0} = \lim_{x \to 0} \frac{\sin x}{x};$$

see Fig. 19.5. The proof of this depends on how $\sin x$ is defined; see the brief discussion in 37.12. For a discussion of this limit and L'Hospital's rule, see Example 1 in §30. □

Here is another useful criterion that implies uniform continuity.

19.6 Theorem.
Let f be a continuous function on an interval I [I may be bounded or unbounded]. Let I° be the interval obtained by removing from I any endpoints that happen to be in I. If f is differentiable on I° and if f' is bounded on I°, then f is uniformly continuous on I.

Proof
For this proof we need the Mean Value theorem, which can be found in most calculus texts or later in this book [Theorem 29.3].

Let M be a bound for f' on I so that $|f'(x)| \leq M$ for all x. Let $\epsilon > 0$ and let $\delta = \frac{\epsilon}{M}$. Consider $a, b \in I$ where $a < b$ and $|b - a| < \delta$. By the Mean Value theorem, there exists $x \in (a, b)$ such that $f'(x) = \frac{f(b) - f(a)}{b - a}$, so

$$|f(b) - f(a)| = |f'(x)| \cdot |b - a| \leq M|b - a| < M\delta = \epsilon.$$

This proves the uniform continuity of f on I. ■

Example 10
Let $a > 0$ and consider $f(x) = \frac{1}{x^2}$. Since $f'(x) = -\frac{2}{x^3}$ we have $|f'(x)| \le \frac{2}{a^3}$ on $[a, \infty)$. Hence f is uniformly continuous on $[a, \infty)$ by Theorem 19.6. For a direct proof of this fact, see Example 2. □

Exercises

19.1 Which of the following continuous functions are uniformly continuous on the specified set? Justify your answers. Use any theorems you wish.

(a) $f(x) = x^{17} \sin x - e^x \cos 3x$ on $[0, \pi]$,

(b) $f(x) = x^3$ on $[0, 1]$,

(c) $f(x) = x^3$ on $(0, 1)$,

(d) $f(x) = x^3$ on $\mathbb{R}$,

(e) $f(x) = \frac{1}{x^3}$ on $(0, 1]$,

(f) $f(x) = \sin \frac{1}{x^2}$ on $(0, 1]$,

(g) $f(x) = x^2 \sin \frac{1}{x}$ on $(0, 1]$.

19.2 Prove each of the following functions is uniformly continuous on the indicated set by directly verifying the ϵ–δ property in Definition 19.1.

(a) $f(x) = 3x + 11$ on $\mathbb{R}$,

(b) $f(x) = x^2$ on $[0, 3]$,

(c) $f(x) = \frac{1}{x}$ on $[\frac{1}{2}, \infty)$.

19.3 Repeat Exercise 19.2 for the following.

(a) $f(x) = \frac{x}{x+1}$ on $[0, 2]$,

(b) $f(x) = \frac{5x}{2x-1}$ on $[1, \infty)$.

19.4 **(a)** Prove that if f is uniformly continuous on a bounded set S, then f is a bounded function on S. *Hint*: Assume not. Use Theorems 11.5 and 19.4.

(b) Use (a) to give yet another proof that $\frac{1}{x^2}$ is not uniformly continuous on $(0, 1)$.

19.5 Which of the following continuous functions is uniformly continuous on the specified set? Justify your answers, using appropriate theorems or Exercise 19.4(a).

(a) $\tan x$ on $[0, \frac{\pi}{4}]$,

(b) $\tan x$ on $[0, \frac{\pi}{2})$,

(c) $\frac{1}{x}\sin^2 x$ on $(0, \pi]$,

(d) $\frac{1}{x-3}$ on $(0, 3)$,

(e) $\frac{1}{x-3}$ on $(3, \infty)$,

(f) $\frac{1}{x-3}$ on $(4, \infty)$.

19.6 **(a)** Let $f(x) = \sqrt{x}$ for $x \geq 0$. Show f' is unbounded on $(0, 1]$ but f is nevertheless uniformly continuous on $(0, 1]$. Compare with Theorem 19.6.

(b) Show f is uniformly continuous on $[1, \infty)$.

19.7 **(a)** Let f be a continuous function on $[0, \infty)$. Prove that if f is uniformly continuous on $[k, \infty)$ for some k, then f is uniformly continuous on $[0, \infty)$.

(b) Use (a) and Exercise 19.6(b) to prove $\sqrt{x}$ is uniformly continuous on $[0, \infty)$.

19.8 **(a)** Use the Mean Value theorem to prove

$$|\sin x - \sin y| \leq |x - y|$$

for all x, y in $\mathbb{R}$; see the proof of Theorem 19.6.

(b) Show $\sin x$ is uniformly continuous on $\mathbb{R}$.

19.9 Let $f(x) = x \sin(\frac{1}{x})$ for $x \neq 0$ and $f(0) = 0$.

(a) Observe f is continuous on $\mathbb{R}$; see Exercises 17.3(f) and 17.9(c).

(b) Why is f uniformly continuous on any bounded subset of $\mathbb{R}$?

(c) Is f uniformly continuous on $\mathbb{R}$?

19.10 Repeat Exercise 19.9 for the function g where $g(x) = x^2 \sin(\frac{1}{x})$ for $x \neq 0$ and $g(0) = 0$.

19.11 Accept the fact that the function $\tilde{h}$ in Example 9 is continuous on $\mathbb{R}$; prove it is uniformly continuous on $\mathbb{R}$.

19.12 Let f be a continuous function on $[a,b]$. Show the function f^* defined by $f^*(x) = \sup\{f(y) : a \le y \le x\}$, for $x \in [a,b]$, is an increasing continuous function on $[a,b]$.

§20 Limits of Functions

A function f is continuous at a point a provided the values $f(x)$ are near the value $f(a)$ for x near a [and $x \in \mathrm{dom}(f)$]. See Definition 17.1 and Theorem 17.2. It would be reasonable to view $f(a)$ as the limit of the values $f(x)$, for x near a, and to write $\lim_{x\to a} f(x) = f(a)$. In this section we formalize this notion. This section is needed for our careful study of derivatives in Chap. 5, but it may be deferred until then.

We will be interested in ordinary limits, left-handed and right-handed limits and infinite limits. In order to handle these various concepts efficiently and also to emphasize their common features, we begin with a very general definition, which is not a standard definition.

20.1 Definition.
Let S be a subset of $\mathbb{R}$, let a be a real number or symbol ∞ or $-\infty$ that is the limit of some sequence in S, and let L be a real number or symbol $+\infty$ or $-\infty$. We write $\lim_{x\to a^S} f(x) = L$ if

$$f \text{ is a function defined on } S, \tag{1}$$

and

$$\begin{aligned}&\text{for every sequence } (x_n) \text{ in } S \text{ with limit } a,\\ &\text{we have } \lim_{n\to\infty} f(x_n) = L.\end{aligned} \tag{2}$$

The expression "$\lim_{x\to a^S} f(x)$" is read "limit, as x tends to a along S, of $f(x)$."

20.2 Remarks.

(a) From Definition 17.1 we see that a function f is continuous at a in $\mathrm{dom}(f) = S$ if and only if $\lim_{x\to a^S} f(x) = f(a)$.

(b) Observe that limits, when they exist, are unique. This follows from (2) of Definition 20.1, since limits of sequences are unique, a fact that is verified at the end of §7. □

We now define the various standard limit concepts for functions.

20.3 Definition.

(a) For $a \in \mathbb{R}$ and a function f we write $\lim_{x\to a} f(x) = L$ provided $\lim_{x\to a^S} f(x) = L$ for some set $S = J \setminus \{a\}$ where J is an open interval containing a. $\lim_{x\to a} f(x)$ is called the [*two-sided*] *limit of f at a*. Note f need not be defined at a and, even if f is defined at a, the value $f(a)$ need not equal $\lim_{x\to a} f(x)$. In fact, $f(a) = \lim_{x\to a} f(x)$ if and only if f is defined on an open interval containing a and f is continuous at a.

(b) For $a \in \mathbb{R}$ and a function f we write $\lim_{x\to a^+} f(x) = L$ provided $\lim_{x\to a^S} f(x) = L$ for some open interval $S = (a, b)$. $\lim_{x\to a^+} f(x)$ is the *right-hand limit of f at a*. Again f need not be defined at a.

(c) For $a \in \mathbb{R}$ and a function f we write $\lim_{x\to a^-} f(x) = L$ provided $\lim_{x\to a^S} f(x) = L$ for some open interval $S = (c, a)$. $\lim_{x\to a^-} f(x)$ is the *left-hand limit of f at a*.

(d) For a function f we write $\lim_{x\to\infty} f(x) = L$ provided $\lim_{x\to\infty^S} f(x) = L$ for some interval $S = (c, \infty)$. Likewise, we write $\lim_{x\to-\infty} f(x) = L$ provided $\lim_{x\to-\infty^S} f(x) = L$ for some interval $S = (-\infty, b)$.

The limits defined above are unique; i.e., they do not depend on the exact choice of the set S [Exercise 20.19].

Example 1

We have $\lim_{x\to 4} x^3 = 64$ and $\lim_{x\to 2} \frac{1}{x} = \frac{1}{2}$ because the functions x^3 and $\frac{1}{x}$ are continuous at 4 and 2, respectively. It is easy to show $\lim_{x\to 0^+} \frac{1}{x} = +\infty$ and $\lim_{x\to 0^-} \frac{1}{x} = -\infty$; see Exercise 20.14. It follows that $\lim_{x\to 0} \frac{1}{x}$ does *not* exist; see Theorem 20.10. □

Example 2

Consider $\lim_{x\to 2} \frac{x^2-4}{x-2}$. This is not like Example 1, because the function under the limit is not even defined at $x = 2$. However, we can rewrite the function as

$$\frac{x^2-4}{x-2} = \frac{(x-2)(x+2)}{x-2} = x+2 \quad \text{for} \quad x \neq 2.$$

Now it is clear that $\lim_{x\to 2} \frac{x^2-4}{x-2} = \lim_{x\to 2}(x+2) = 4$. We should emphasize that the functions $\frac{x^2-4}{x-2}$ and $x+2$ are *not* identical. The domain of $f(x) = \frac{x^2-4}{x-2}$ is $(-\infty, 2) \cup (2, \infty)$ while the domain of $\tilde{f}(x) = x+2$ is $\mathbb{R}$, so that $\tilde{f}$ is an extension of f. This seems like nitpicking and this example may appear silly, but the function f, not $\tilde{f}$, arises naturally in computing the derivative of $g(x) = x^2$ at $x = 2$. Indeed, using the definition of derivative we have

$$g'(2) = \lim_{x\to 2} \frac{g(x)-g(2)}{x-2} = \lim_{x\to 2} \frac{x^2-4}{x-2},$$

so our modest computation above shows $g'(2) = 4$. Of course, this is obvious from the formula $g'(x) = 2x$, but we are preparing the foundations of limits and derivatives, so we are beginning with simple examples. □

Example 3

Consider $\lim_{x\to 1} \frac{\sqrt{x}-1}{x-1}$. We employ a trick that should be familiar by now; we multiply the numerator and denominator by $\sqrt{x}+1$ and obtain

$$\frac{\sqrt{x}-1}{x-1} = \frac{(\sqrt{x}-1)(\sqrt{x}+1)}{(x-1)(\sqrt{x}+1)} = \frac{x-1}{(x-1)(\sqrt{x}+1)} = \frac{1}{\sqrt{x}+1}$$

for $x \neq 1$. Hence $\lim_{x\to 1} \frac{\sqrt{x}-1}{x-1} = \lim_{x\to 1} \frac{1}{\sqrt{x}+1} = \frac{1}{2}$. We have just laboriously verified that if $h(x) = \sqrt{x}$, then $h'(1) = \frac{1}{2}$. □

Example 4

Let $f(x) = \frac{1}{(x-2)^3}$ for $x \neq 2$. Then $\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0$, $\lim_{x\to 2^+} f(x) = +\infty$ and $\lim_{x\to 2^-} f(x) = -\infty$.

To verify $\lim_{x\to\infty} f(x) = 0$, we consider a sequence (x_n) such that $\lim_{n\to\infty} x_n = +\infty$ and show $\lim_{n\to\infty} f(x_n) = 0$. This will show $\lim_{x\to\infty^S} f(x) = 0$ for $S = (2,\infty)$, for example. Exercise 9.11 and Theorem 9.9 show $\lim_{n\to\infty}(x_n-2)^3 = +\infty$, and then Theorem 9.10

shows

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} (x_n - 2)^{-3} = 0. \tag{1}$$

Here is a direct proof of (1). Let $\epsilon > 0$. For large n, we need $|x_n - 2|^{-3} < \epsilon$ or $\epsilon^{-1} < |x_n - 2|^3$ or $\epsilon^{-1/3} < |x_n - 2|$. The last inequality holds if $x_n > \epsilon^{-1/3} + 2$. Since $\lim_{n\to\infty} x_n = +\infty$, there exists N so that

$$n > N \quad \text{implies} \quad x_n > \epsilon^{-1/3} + 2.$$

Reversing the algebraic steps above, we find

$$n > N \quad \text{implies} \quad |x_n - 2|^{-3} < \epsilon.$$

This establishes (1).

Similar arguments prove $\lim_{x\to-\infty} f(x) = 0$ and $\lim_{x\to 2^+} f(x) = +\infty$. To prove $\lim_{x\to 2^-} f(x) = -\infty$, consider a sequence (x_n) such that $x_n < 2$ for all n and $\lim_{n\to\infty} x_n = 2$. Then $2 - x_n > 0$ for all n and $\lim_{n\to\infty}(2 - x_n) = 0$. Hence $\lim_{n\to\infty}(2 - x_n)^3 = 0$ by Theorem 9.4, and 9.10 implies $\lim_{n\to\infty}(2 - x_n)^{-3} = +\infty$. It follows [Exercise 9.10(b)] that

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} (x_n - 2)^{-3} = -\infty. \tag{2}$$

This proves that $\lim_{x\to 2^S} f(x) = -\infty$ for $S = (-\infty, 2)$, so that $\lim_{x\to 2^-} f(x) = -\infty$. Of course, a direct proof of (2) also can be given.

The limits discussed above are confirmed in Fig. 20.1. $\square$

We will discuss the various limits defined in Definition 20.3 further at the end of this section. First we prove some limit theorems in considerable generality.

20.4 Theorem.
Let f_1 and f_2 be functions for which the limits $L_1 = \lim_{x\to a^S} f_1(x)$ and $L_2 = \lim_{x\to a^S} f_2(x)$ exist and are finite. Then

(i) $\lim_{x\to a^S}(f_1 + f_2)(x)$ *exists and equals* $L_1 + L_2$*;*

(ii) $\lim_{x\to a^S}(f_1 f_2)(x)$ *exists and equals* $L_1 L_2$*;*

(iii) $\lim_{x\to a^S}(f_1/f_2)(x)$ *exists and equals* L_1/L_2 *provided* $L_2 \neq 0$ *and* $f_2(x) \neq 0$ *for* $x \in S$*.*

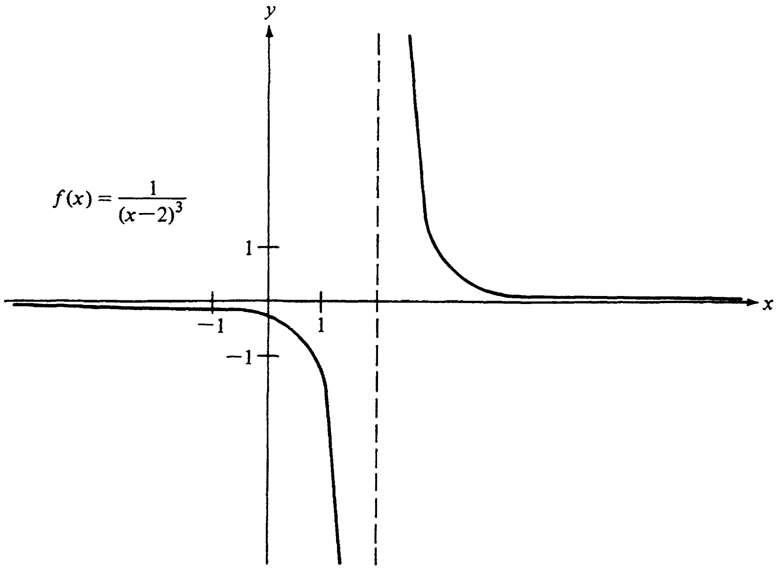

FIGURE 20.1

Proof

The hypotheses imply both f_1 and f_2 are defined on S and a is the limit of some sequence in S. Clearly the functions $f_1 + f_2$ and $f_1 f_2$ are defined on S and so is f_1/f_2 if $f_2(x) \neq 0$ for $x \in S$.

Consider a sequence (x_n) in S with limit a. By hypotheses we have $L_1 = \lim_{n\to\infty} f_1(x_n)$ and $L_2 = \lim_{n\to\infty} f_2(x_n)$. Theorems 9.3 and 9.4 now show

$$\lim_{n\to\infty}(f_1 + f_2)(x_n) = \lim_{n\to\infty} f_1(x_n) + \lim_{n\to\infty} f_2(x_n) = L_1 + L_2$$

and

$$\lim_{n\to\infty}(f_1 f_2)(x_n) = \left[\lim_{n\to\infty} f_1(x_n)\right] \cdot \left[\lim_{n\to\infty} f_2(x_n)\right] = L_1 L_2.$$

Thus (2) in Definition 20.1 holds for $f_1 + f_2$ and $f_1 f_2$, so that (i) and (ii) hold. Likewise (iii) follows by an application of Theorem 9.6. ■

Some of the infinite variations of Theorem 20.4 appear in Exercise 20.20. The next theorem is less general than might have been expected; Example 7 shows why.

20.5 Theorem.
Let f be a function for which the limit $L = \lim_{x \to a^S} f(x)$ exists and is finite. If g is a function defined on $\{f(x) : x \in S\} \cup \{L\}$ that is continuous at L, then $\lim_{x \to a^S} g \circ f(x)$ exists and equals $g(L)$.

Proof
Note $g \circ f$ is defined on S by our assumptions. Consider a sequence (x_n) in S with limit a. Then we have $L = \lim_{n \to \infty} f(x_n)$. Since g is continuous at L, it follows that

$$g(L) = \lim_{n \to \infty} g(f(x_n)) = \lim_{n \to \infty} g \circ f(x_n).$$

Hence $\lim_{x \to a^S} g \circ f(x) = g(L)$. ■

Example 5
If f is a function for which the limit $L = \lim_{x \to a} f(x)$ exists and is finite, then we have $\lim_{x \to a} |f(x)| = |L|$. This follows immediately from Theorem 20.5 with $g(x) = |x|$. Similarly, we have $\lim_{x \to a} e^{f(x)} = e^L$ if we accept the fact that $g(x) = e^x$ is continuous on $\mathbb{R}$. □

Example 6
If f is a function for which $\lim_{x \to 0^+} f(x) = 0$ and $\lim_{x \to \infty} f(x) = \frac{\pi}{2}$, then we have $\lim_{x \to 0^+} e^{f(x)} = e^0 = 1$, $\lim_{x \to \infty} e^{f(x)} = e^{\frac{\pi}{2}}$, $\lim_{x \to 0^+} \sin(f(x)) = \sin(0) = 0$ and $\lim_{x \to \infty} \sin(f(x)) = \sin \frac{\pi}{2} = 1$. □

Example 7
We give an example to show continuity of g is needed in Theorem 20.5. Explicitly, we give examples of functions f and g such that $\lim_{x \to 0} f(x) = 1$, $\lim_{x \to 1} g(x) = 4$ and yet $\lim_{x \to 0} g \circ f(x)$ does not exist. One would expect this limit to exist and to equal 4, but in the example $f(x)$ will *equal* 1 for arbitrarily small x while $g(1) \neq 4$. The functions f and g are defined by $f(x) = 1 + x \sin \frac{\pi}{x}$ for $x \neq 0$, $g(x) = 4$ for $x \neq 1$, and $g(1) = -4$. Clearly $\lim_{x \to 0} f(x) = 1$ and $\lim_{x \to 1} g(x) = 4$. Let $x_n = \frac{2}{n}$ for $n \in \mathbb{N}$. Then $f(x_n) = 1 + \frac{2}{n} \sin(\frac{n\pi}{2})$; hence $f(x_n) = 1$ for even n and $f(x_n) \neq 1$ for odd n. Therefore $g \circ f(x_n) = -4$ for even n and $g \circ f(x_n) = 4$ for odd n. Since $\lim_{n \to \infty} x_n = 0$, we conclude $\lim_{x \to 0} g \circ f(x)$ does not exist. □

As in Theorem 17.2, the limits defined in Definitions 20.1 and 20.3 can be recast to avoid sequences. First we state and prove a typical result of this sort. Then, after Corollary 20.8, we give a general scheme without proof.

20.6 Theorem.
Let f be a function defined on a subset S of $\mathbb{R}$, let a be a real number that is the limit of some sequence in S, and let L be a real number. then $\lim_{x\to a^S} f(x) = L$ *if and only if*

$$\begin{array}{l} \text{for each } \epsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ x \in S \text{ and } |x-a| < \delta \text{ imply } |f(x) - L| < \epsilon. \end{array} \tag{1}$$

Proof
We imitate our proof of Theorem 17.2. Suppose (1) holds, and consider a sequence (x_n) in S such that $\lim_{n\to\infty} x_n = a$. To show $\lim_{n\to\infty} f(x_n) = L$, consider $\epsilon > 0$. By (1) there exists $\delta > 0$ such that

$$x \in S \quad \text{and} \quad |x-a| < \delta \quad \text{imply} \quad |f(x) - L| < \epsilon.$$

Since $\lim_{n\to\infty} x_n = a$, there exists a number N such that $n > N$ implies $|x_n - a| < \delta$. Since $x_n \in S$ for all n, we conclude

$$n > N \quad \text{implies} \quad |f(x_n) - L| < \epsilon.$$

Thus $\lim_{n\to\infty} f(x_n) = L$.

Now assume $\lim_{x\to a^S} f(x) = L$, but (1) fails. Then for some $\epsilon > 0$ the implication

$$\text{“}x \in S \quad \text{and} \quad |x-a| < \delta \quad \text{imply} \quad |f(x) - L| < \epsilon\text{”}$$

fails for each $\delta > 0$. Then for each $n \in \mathbb{N}$ there exists x_n in S where $|x_n - a| < \frac{1}{n}$ while $|f(x_n) - L| \geq \epsilon$. Hence (x_n) is a sequence in S with limit a for which $\lim_{n\to\infty} f(x_n) = L$ fails. Consequently $\lim_{x\to a^S} f(x) = L$ fails to hold. ■

20.7 Corollary.
Let f be a function defined on $J \setminus \{a\}$ for some open interval J containing a, and let L be a real number. Then $\lim_{x\to a} f(x) = L$ *if*

and only if

$$\text{for each } \epsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ 0 < |x-a| < \delta \text{ implies } |f(x)-L| < \epsilon. \tag{1}$$

20.8 Corollary.
Let f be a function defined on some interval (a,b), and let L be a real number. Then $\lim_{x\to a^+} f(x) = L$ if and only if

$$\text{for each } \epsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ 0 < x < a+\delta \text{ implies } |f(x)-L| < \epsilon. \tag{1}$$

20.9 Discussion.
We now consider $\lim_{x\to s} f(x) = L$ where L can be finite, $+\infty$ or $-\infty$, and s is a symbol a, a^+, a^-, ∞ or $-\infty$ [here $a \in \mathbb{R}$]. Note we have 15 [$= 3 \cdot 5$] different sorts of limits here. It turns out that $\lim_{x\to s} f(x) = L$ if and only if

$$\text{for each } \underline{\qquad\quad} \text{ there exists } \underline{\qquad\quad} \text{ such that} \\ \underline{\qquad\quad} \text{ implies } \underline{\qquad\quad}. \tag{1}$$

For finite limits L, the first and last blanks are filled in by "$\epsilon > 0$" and "$|f(x)-L| < \epsilon$." For $L = +\infty$, the first and last blanks are filled in by "$M > 0$" and "$f(x) > M$," while for $L = -\infty$ they are filled in by "$M < 0$" and "$f(x) < M$." When we consider $\lim_{x\to a} f(x)$, then f is defined on $J \setminus \{a\}$ for some open interval J containing a, and the second and third blanks are filled in by "$\delta > 0$" and "$0 < |x-a| < \delta$." For $\lim_{x\to a^+} f(x)$ we require f to be defined on an interval (a,b) and the second and third blanks are filled in by "$\delta > 0$" and "$a < x < a+\delta$." For $\lim_{x\to a^-} f(x)$ we require f to be defined on an interval (c,a) and the second and third blanks are filled in by "$\delta > 0$" and "$a-\delta < x < a$." For $\lim_{x\to\infty} f(x)$ we require f to be defined on an interval (c,∞) and the second and third blanks are filled in by "$\alpha < \infty$ and "$\alpha < x$." A similar remark applies to $\lim_{x\to-\infty} f(x)$.

The assertions above with L finite and s equal to a or a^+ are contained in Corollaries 20.7 and 20.8. □

20.10 Theorem.
Let f be a function defined on $J \setminus \{a\}$ for some open interval J containing a. Then $\lim_{x\to a} f(x)$ exists if and only if the limits

$\lim_{x\to a^+} f(x)$ and $\lim_{x\to a^-} f(x)$ both exist and are equal, in which case all three limits are equal.

Proof

Suppose $\lim_{x\to a} f(x) = L$ and L is finite. Then (1) in Corollary 20.7 holds, so (1) in Corollary 20.8 obviously holds. Thus we have $\lim_{x\to a^+} f(x) = L$; similarly $\lim_{x\to a^-} f(x) = L$.

Now suppose $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L$ where L is finite. Consider $\epsilon > 0$; we apply Corollary 20.8 and its analogue for a^- to obtain $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$a < x < a + \delta_1 \quad \text{implies} \quad |f(x) - L| < \epsilon$$

and

$$a - \delta_2 < x < a \quad \text{implies} \quad |f(x) - L| < \epsilon.$$

If $\delta = \min\{\delta_1, \delta_2\}$, then

$$0 < |x - a| < \delta \quad \text{implies} \quad |f(x) - L| < \epsilon,$$

so $\lim_{x\to a} f(x) = L$ by Corollary 20.7.

Similar arguments apply if the limits L are infinite. For example, suppose $\lim_{x\to a} f(x) = +\infty$ and consider $M > 0$. There exists $\delta > 0$ such that

$$0 < |x - a| < \delta \quad \text{implies} \quad f(x) > M. \tag{1}$$

Then clearly

$$a < x < a + \delta \quad \text{implies} \quad f(x) > M \tag{2}$$

and

$$a - \delta < x < a \quad \text{implies} \quad f(x) > M, \tag{3}$$

so that $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = +\infty$.

As a last example, suppose $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = +\infty$. For each $M > 0$ there exists $\delta_1 > 0$ so that (2) holds, and there exists $\delta_2 > 0$ so that (3) holds. Then (1) holds with $\delta = \min\{\delta_1, \delta_2\}$. We conclude $\lim_{x\to a} f(x) = +\infty$. ■

20.11 Remark.
Note $\lim_{x\to-\infty} f(x)$ is very similar to the right-hand limits $\lim_{x\to a^+} f(x)$. For example, if L is finite, then $\lim_{x\to a^+} f(x) = L$ if and only if

$$\begin{aligned}&\text{for each } \epsilon > 0 \text{ there exists } \alpha > a \text{ such that}\\ &a < x < \alpha \text{ implies } |f(x) - L| < \epsilon,\end{aligned} \tag{1}$$

since $\alpha > a$ if and only if $\alpha = a + \delta$ for some $\delta > 0$; see Corollary 20.8. If we set $a = -\infty$ in (1), we obtain the condition (1) in Discussion 20.9 equivalent to $\lim_{x\to-\infty} f(x) = L$.

In the same way, the limits $\lim_{x\to\infty} f(x)$ and $\lim_{x\to a^-} f(x)$ will equal L [L finite] if and only if

$$\begin{aligned}&\text{for each } \epsilon > 0 \text{ there exists } \alpha < a \text{ such that}\\ &\alpha < x < a \text{ implies } |f(x) - L| < \epsilon.\end{aligned} \tag{2}$$

Obvious changes are needed if L is infinite. □

Exercises

20.1 Sketch the function $f(x) = \frac{x}{|x|}$. Determine, by inspection, the limits $\lim_{x\to\infty} f(x)$, $\lim_{x\to0^+} f(x)$, $\lim_{x\to0^-} f(x)$, $\lim_{x\to-\infty} f(x)$ and $\lim_{x\to0} f(x)$ *when they exist*. Also indicate when they do not exist.

20.2 Repeat Exercise 20.1 for $f(x) = \frac{x^3}{|x|}$.

20.3 Repeat Exercise 20.1 for $f(x) = \frac{\sin x}{x}$. See Example 9 of §19.

20.4 Repeat Exercise 20.1 for $f(x) = x \sin \frac{1}{x}$.

20.5 Prove the limit assertions in Exercise 20.1.

20.6 Prove the limit assertions in Exercise 20.2.

20.7 Prove the limit assertions in Exercise 20.3.

20.8 Prove the limit assertions in Exercise 20.4.

20.9 Repeat Exercise 20.1 for $f(x) = \frac{1-x^2}{x}$.

20.10 Prove the limit assertions in Exercise 20.9.

20.11 Find the following limits.

(a) $\lim_{x\to a} \frac{x^2-a^2}{x-a}$ **(b)** $\lim_{x\to b} \frac{\sqrt{x}-\sqrt{b}}{x-b}$, $b > 0$

(c) $\lim_{x\to a} \frac{x^3-a^3}{x-a}$ *Hint*: $x^3-a^3 = (x-a)(x^2+ax+a^2)$.

20.12 **(a)** Sketch the function $f(x) = (x-1)^{-1}(x-2)^{-2}$.

(b) Determine $\lim_{x\to 2^+} f(x)$, $\lim_{x\to 2^-} f(x)$, $\lim_{x\to 1^+} f(x)$ and $\lim_{x\to 1^-} f(x)$.

(c) Determine $\lim_{x\to 2} f(x)$ and $\lim_{x\to 1} f(x)$ if they exist.

20.13 Prove that if $\lim_{x\to a} f(x) = 3$ and $\lim_{x\to a} g(x) = 2$, then

(a) $\lim_{x\to a}[3f(x) + g(x)^2] = 13$,

(b) $\lim_{x\to a} \frac{1}{g(x)} = \frac{1}{2}$,

(c) $\lim_{x\to a} \sqrt{3f(x) + 8g(x)} = 5$.

20.14 Prove $\lim_{x\to 0^+} \frac{1}{x} = +\infty$ and $\lim_{x\to 0^-} \frac{1}{x} = -\infty$.

20.15 Prove $\lim_{x\to -\infty} f(x) = 0$ and $\lim_{x\to 2^+} f(x) = +\infty$ for the function f in Example 4.

20.16 Suppose the limits $L_1 = \lim_{x\to a^+} f_1(x)$ and $L_2 = \lim_{x\to a^+} f_2(x)$ exist.

(a) Show if $f_1(x) \le f_2(x)$ for all x in some interval (a, b), then $L_1 \le L_2$.

(b) Suppose that, in fact, $f_1(x) < f_2(x)$ for all x in some interval (a, b). Can you conclude $L_1 < L_2$?

20.17 Show that if $\lim_{x\to a^+} f_1(x) = \lim_{x\to a^+} f_3(x) = L$ and if $f_1(x) \le f_2(x) \le f_3(x)$ for all x in some interval (a, b), then $\lim_{x\to a^+} f_2(x) = L$. This is called the squeeze lemma. *Warning*: This is not immediate from Exercise 20.16(a), because we are not assuming $\lim_{x\to a^+} f_2(x)$ exists; this must be proved.

20.18 Let $f(x) = \frac{\sqrt{1+3x^2}-1}{x^2}$ for $x \ne 0$. Show $\lim_{x\to 0} f(x)$ exists and determine its value. Justify all claims.

20.19 The limits defined in Definition 20.3 do not depend on the choice of the set S. As an example, consider $a < b_1 < b_2$ and suppose f is defined on (a, b_2). Show that if the limit $\lim_{x\to a^S} f(x)$ exists for either $S = (a, b_1)$ or $S = (a, b_2)$, then the limit exists for the other choice of S and these limits are identical. Their common value is what we write as $\lim_{x\to a^+} f(x)$.

20.20 Let f_1 and f_2 be functions such that $\lim_{x\to a^S} f_1(x) = +\infty$ and such that the limit $L_2 = \lim_{x\to a^S} f_2(x)$ exists.

(a) Prove $\lim_{x\to a^S}(f_1 + f_2)(x) = +\infty$ if $L_2 \ne -\infty$. *Hint*: Use Exercise 9.11.

(b) Prove $\lim_{x\to a^S}(f_1f_2)(x) = +\infty$ if $0 < L_2 \le +\infty$. *Hint*: Use Theorem 9.9.

(c) Prove $\lim_{x\to a^S}(f_1f_2)(x) = -\infty$ if $-\infty \le L_2 < 0$.

(d) What can you say about $\lim_{x\to a^S}(f_1f_2)(x)$ if $L_2 = 0$?

§21 * More on Metric Spaces: Continuity

In this section and the next section we continue the introduction to metric space ideas initiated in §13. More thorough treatments appear in [33, 53] and [62]. In particular, for this brief introduction we avoid the technical and somewhat confusing matter of relative topologies that is not, and should not be, avoided in the more thorough treatments.

We are interested in functions between metric spaces (S, d) and (S^*, d^*). We will write "$f\colon S \to S^*$" to signify $\mathrm{dom}(f) = S$ and f takes values in S^*, i.e., $f(x) \in S^*$ for all $s \in S$.

21.1 Definition.
Consider metric spaces (S, d) and (S^*, d^*). A function $f\colon S \to S^*$ is *continuous at* s_0 in S if

$$\begin{array}{l}\text{for each } \epsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ d(s, s_0) < \delta \text{ implies } d^*(f(s), f(s_0)) < \epsilon.\end{array} \tag{1}$$

We say f is *continuous on a subset* E of S if f is continuous at each point of E. The function f is *uniformly continuous on a subset* E of S if

$$\begin{array}{l}\text{for each } \epsilon > 0 \text{ there exists } \delta > 0 \text{ such that} \\ s, t \in E \text{ and } d(s, t) < \delta \text{ imply } d^*(f(s), f(t)) < \epsilon.\end{array} \tag{2}$$

Example 1
Let $S = S^* = \mathbb{R}$ and $d = d^* = \mathrm{dist}$ where, as usual, $\mathrm{dist}(a, b) = |a-b|$. The definition of continuity given above is equivalent to that in §17 in view of Theorem 17.2. The definition of uniform continuity is equivalent to that in Definition 19.1. □

Example 2
In several variable calculus, real-valued functions with domain $\mathbb{R}^2$ or $\mathbb{R}^3$, or even $\mathbb{R}^k$, are extensively studied. This corresponds to the case $S = \mathbb{R}^k$,

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

$S^* = \mathbb{R}$ and $d^* = \text{dist}$. We will not develop the theory, but generally speaking, functions that look continuous will be. Some examples on $\mathbb{R}^2$ are $f(x_1, x_2) = x_1^2 + x_2^2$, $f(x_1, x_2) = x_1 x_2\sqrt{x_1^2 + x_2^2 + 1}$, $f(x_1, x_2) = \cos(x_1 - x_2^5)$. Some examples on $\mathbb{R}^3$ are $g(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$, $g(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3$, $g(x_1, x_2, x_3) = e^{x_1 + x_2}\log(x_3^2 + 2)$. □

Example 3
Functions with domain $\mathbb{R}$ and values in $\mathbb{R}^2$ or $\mathbb{R}^3$, or generally $\mathbb{R}^k$, are also studied in several variable calculus. This corresponds to the case $S = \mathbb{R}$, $d = \text{dist}$, $S^* = \mathbb{R}^k$ and

$$d^*(\boldsymbol{x}, \boldsymbol{y}) = \left[\sum_{j=1}^{k}(x_j - y_j)^2\right]^{1/2}.$$

The images of such functions are what nonmathematicians often call a "curve" or "path." In order to distinguish a function from its image, we will adhere to the following terminology. Suppose $\gamma: \mathbb{R} \to \mathbb{R}^k$ is continuous. Then we will call γ a *path*; its image $\gamma(\mathbb{R})$ in $\mathbb{R}^k$ will be called a *curve*. We will also use this terminology if γ is defined and continuous on some subinterval of $\mathbb{R}$, such as $[a, b]$; see Exercise 21.7.

As an example, consider γ where $\gamma(t) = (\cos t, \sin t)$. This function maps $\mathbb{R}$ onto the circle in $\mathbb{R}^2$ about $(0, 0)$ with radius 1. More generally $\gamma_0(t) = (a\cos t, b\sin t)$ maps $\mathbb{R}$ onto the ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$; see Fig. 21.1.

The graph of an ordinary continuous function $f: \mathbb{R} \to \mathbb{R}$ looks like a curve, and it is! It is the curve for the path $\gamma(t) = (t, f(t))$.

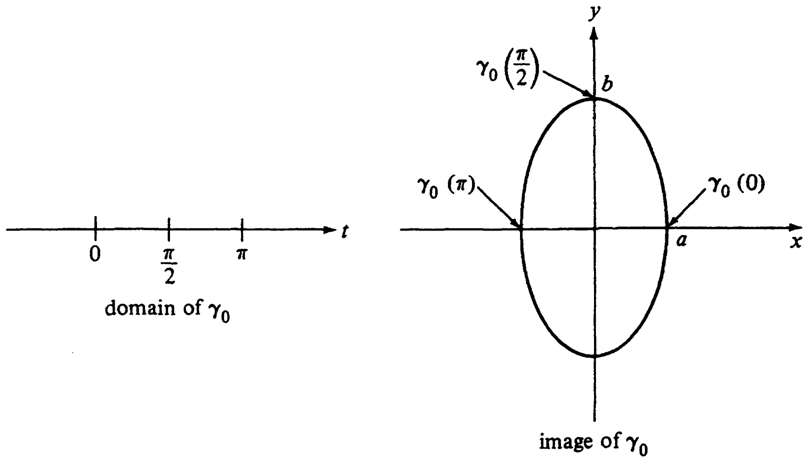

FIGURE 21.1

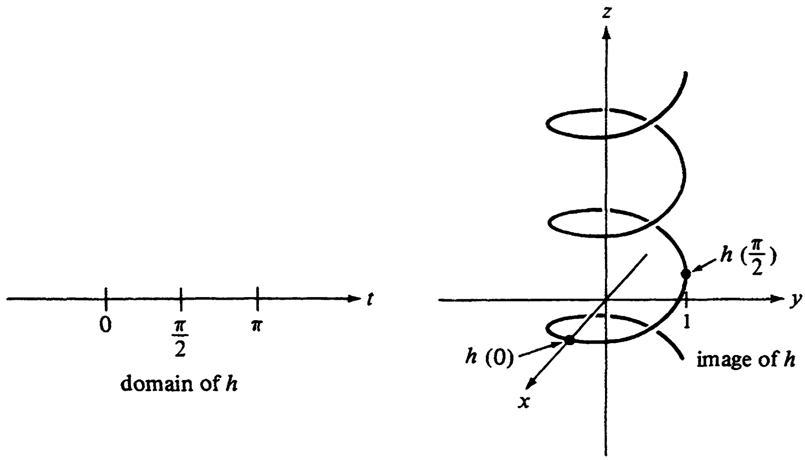

FIGURE 21.2

Curves in $\mathbb{R}^3$ can be quite exotic. For example, the curve for the path $h(t) = (\cos t, \sin t, \frac{t}{4})$ is a *helix*. See Fig. 21.2. □

We did not prove that any of the paths above are continuous, because we can easily prove the following general fact.

21.2 Proposition.
Consider a function $\gamma : [a,b] \to \mathbb{R}^k$ and write

$$\gamma(t) = (f_1(t), f_2(t), \ldots, f_k(t)) \quad \textit{for} \quad t \in [a,b];$$

thus $f_1, f_2, \ldots, f_k$ are real-valued functions on $[a,b]$. Then γ is continuous, i.e., a path, if and only if each f_j is continuous.

Proof
Both implications follow from formula (1) in the proof of Lemma 13.3 and Exercise 13.2:

$$|x_j - y_j| \le d^*(\boldsymbol{x}, \boldsymbol{y}) \le \sqrt{k} \max\{|x_j - y_j| : j = 1, 2, \ldots, k\}. \tag{1}$$

For t, t_0 in $[a,b]$, (1) implies

$$|f_j(t) - f_j(t_0)| \le d^*(\gamma(t), \gamma(t_0)) \quad \text{for} \quad j = 1, \ldots, k, \tag{2}$$

$$\text{and } d^*(\gamma(t), \gamma(t_0)) \le \sqrt{k} \max\{|f_j(t) - f_j(t_0)| : j = 1, 2, \ldots, k\}. \tag{3}$$

Suppose γ is continuous on $[a,b]$, and consider some j in $\{1, 2, \ldots, k\}$. To show f_j is continuous at t_0 in $[a,b]$, given $\epsilon > 0$, select $\delta > 0$ so that $d^*(\gamma(t), \gamma(t_0)) < \epsilon$ for $|t - t_0| < \delta$. Then by (2), $|t - t_0| < \delta$ implies $|f_j(t) - f_j(t_0)| < \epsilon$. So f_j is continuous on $[a,b]$.

Now suppose that each f_j is continuous on $[a,b]$, and consider $\epsilon > 0$. For each j, there is $\delta_j > 0$ such that

$$|t - t_0| < \delta_j \quad \text{implies} \quad |f_j(t) - f_j(t_0)| < \frac{\epsilon}{\sqrt{k}}.$$

For $\delta = \min\{\delta_1, \delta_2, \ldots, \delta_k\}$ and $|t - t_0| < \delta$, we have

$$\max\{|f_j(t) - f_j(t_0)| : j = 1, 2, \ldots, k\} < \frac{\epsilon}{\sqrt{k}},$$

so by (3) we have $d^*(\gamma(t), \gamma(t_0)) < \epsilon$. Thus γ is continuous at t_0. ■

The next theorem shows continuity is a topological property; see Discussion 13.7.

21.3 Theorem.
Consider metric spaces (S, d) and (S^*, d^*). A function $f : S \to S^*$ is continuous on S if and only if

$$\begin{array}{l} f^{-1}(U) \textit{ is an open subset of } S \\ \textit{for every open subset } U \textit{ of } S^*. \end{array} \tag{1}$$

Recall $f^{-1}(U) = \{s \in S : f(s) \in U\}$.

Proof

Suppose f is continuous on S. Let U be an open subset of S^*, and consider $s_0 \in f^{-1}(U)$. We need to show s_0 is interior to $f^{-1}(U)$. Since $f(s_0) \in U$ and U is open, we have

$$\{s^* \in S^* : d^*(s^*, f(s_0)) < \epsilon\} \subseteq U \tag{2}$$

for some $\epsilon > 0$. Since f is continuous at s_0, there exists $\delta > 0$ such that

$$d(s, s_0) < \delta \quad \text{implies} \quad d^*(f(s), f(s_0)) < \epsilon. \tag{3}$$

From (2) and (3) we conclude $d(s, s_0) < \delta$ implies $f(s)$ is in U; hence $s \in f^{-1}(U)$. That is,

$$\{s \in S : d(s, s_0) < \delta\} \subseteq f^{-1}(U),$$

so that s_0 is interior to $f^{-1}(U)$.

Conversely, suppose (1) holds, and consider $s_0 \in S$ and $\epsilon > 0$. Then $U = \{s^* \in S^* : d^*(s^*, f(s_0)) < \epsilon\}$ is open in S^*, so $f^{-1}(U)$ is open in S. Since $s_0 \in f^{-1}(U)$, for some $\delta > 0$ we have

$$\{s \in S : d(s, s_0) < \delta\} \subseteq f^{-1}(U).$$

It follows that

$$d(s, s_0) < \delta \quad \text{implies} \quad d^*(f(s), f(s_0)) < \epsilon.$$

Thus f is continuous at s_0. ■

Continuity at a point is also a topological property; see Exercise 21.2. Uniform continuity is a topological property, too, but if we made this precise we would be led to a special class of topologies given by so-called "uniformities."

We will show continuous functions preserve two important topological properties: compactness, and connectedness which will be defined in the next section. The next theorem and corollary illustrate the power of compactness.

21.4 Theorem.
Consider metric spaces (S, d), (S^*, d^*) and a continuous function $f: S \to S^*$. Let E be a compact subset of S. Then
(i) *$f(E)$ is a compact subset of S^*, and*
(ii) *f is uniformly continuous on E.*

Proof
To prove (i), let $\mathcal{U}$ be an open cover of $f(E)$. For each $U \in \mathcal{U}$, $f^{-1}(U)$ is open in S. Moreover, $\{f^{-1}(U) : U \in \mathcal{U}\}$ is a cover of E. Hence there exist $U_1, U_2, \ldots, U_m$ in $\mathcal{U}$ such that

$$E \subseteq f^{-1}(U_1) \cup f^{-1}(U_2) \cup \cdots \cup f^{-1}(U_m).$$

Then

$$f(E) \subseteq U_1 \cup U_2 \cup \cdots \cup U_m,$$

so $\{U_1, U_2, \ldots, U_m\}$ is the desired finite subcover of $\mathcal{U}$ for $f(E)$. This proves (i).

To establish (ii), let $\epsilon > 0$. For each $s \in E$ there exists $\delta_s > 0$ [this δ depends on s] such that

$$d(s,t) < \delta_s \quad \text{implies} \quad d^*(f(s), f(t)) < \frac{\epsilon}{2}. \qquad (1)$$

For each $s \in E$, let $V_s = \{t \in S : d(s,t) < \frac{1}{2}\delta_s\}$. Then the family $\mathcal{V} = \{V_s : s \in E\}$ is an open cover of E. By compactness, there exist finitely many points $s_1, s_2, \ldots, s_n$ in E such that

$$E \subseteq V_{s_1} \cup V_{s_2} \cup \cdots \cup V_{s_n}.$$

Let $\delta = \frac{1}{2}\min\{\delta_{s_1}, \delta_{s_2}, \ldots, \delta_{s_n}\}$. We complete the proof by showing

$$s, t \in E \quad \text{and} \quad d(s,t) < \delta \quad \text{imply} \quad d^*(f(s), f(t)) < \epsilon. \qquad (2)$$

For some k in $\{1, 2, \ldots, n\}$ we have $s \in V_{s_k}$, i.e., $d(s, s_k) < \frac{1}{2}\delta_{s_k}$. Also we have

$$d(t, s_k) \le d(t,s) + d(s, s_k) < \delta + \frac{1}{2}\delta_{s_k} \le \delta_{s_k}.$$

Therefore applying (1) twice we have

$$d^*(f(s), f(s_k)) < \frac{\epsilon}{2} \quad \text{and} \quad d^*(f(t), f(s_k)) < \frac{\epsilon}{2}.$$

Hence $d^*(f(s), f(t)) < \epsilon$ as desired. ■

Assertion (ii) in Theorem 21.4 generalizes Theorem 19.2. The next corollary should be compared with Theorem 18.1.

21.5 Corollary.
Let f be a continuous real-valued function on a metric space (S, d). If E is a compact subset of S, then

(i) *f is bounded on E,*
(ii) *f assumes its maximum and minimum on E.*

Proof
Since $f(E)$ is compact in $\mathbb{R}$, the set $f(E)$ is bounded by Theorem 13.12. This implies (i).

Since $f(E)$ is compact, it contains $\sup f(E)$ by Exercise 13.13. Thus there exists $s_0 \in E$ so that $f(s_0) = \sup f(E)$. This tells us f assumes its maximum value on E at the point s_0. Similarly, f assumes its minimum on E. ■

Example 4
All the functions f in Example 2 are bounded on any compact subset of $\mathbb{R}^2$, i.e., on any closed and bounded set in $\mathbb{R}^2$. Likewise, all the functions g in Example 2 are bounded on each closed and bounded set in $\mathbb{R}^3$. □

Example 5
Let γ be any path in $\mathbb{R}^k$; see Example 3. For $-\infty < a < b < \infty$, the image $\gamma([a, b])$ is closed and bounded in $\mathbb{R}^k$ by Theorem 21.4. Note Corollary 21.5 does not apply in this case, since the set S^* in Theorem 21.4 is $\mathbb{R}^k$, not $\mathbb{R}$. Theorem 21.4 also tells us γ is uniformly continuous on $[a, b]$. Thus if $\epsilon > 0$, there exists $\delta > 0$ such that

$$s, t \in [a, b] \quad \text{and} \quad |s - t| < \delta \quad \text{imply} \quad d(\gamma(s), \gamma(t)) < \epsilon.$$

This fact is useful in several variable calculus, where one integrates along paths γ; compare Discussion 19.3. □

Example 6
Consider a function $f : E \to \mathbb{R}$, where E be a compact subset of $\mathbb{R}$. We will show f is continuous if and only if its graph $G = \{(t, f(t)) : t \in E\}$ is a compact subset of $\mathbb{R}^2$. Suppose first that f is continuous. The proof of Proposition 21.2 shows $\gamma(s) = (s, f(s))$ is a continuous function on E, so its image $\gamma(E) = G$ is compact by Theorem 21.4(i).

Now assume G is compact, but f is not continuous at t_0 in E. Then there exists an $\epsilon > 0$ and a sequence (t_n) in E such that

$$\lim_n t_n = t_0 \quad \text{and} \quad |f(t_n) - f(t_0)| \geq \epsilon \quad \text{for all} \quad n. \tag{1}$$

Since G is a bounded set, the Bolzano-Weierstrass Theorem 13.5 shows $(t_n, f(t_n))$ has a convergent subsequence $(t_{n_k}, f(t_{n_k}))$. Its limit is in G, because G is closed in $\mathbb{R}^2$ by the Heine-Borel Theorem 13.12. Thus there exists $\bar{t}$ in E so that $\lim_k (t_{n_k}, f(t_{n_k})) = (\bar{t}, f(\bar{t}))$. Since $\lim_n t_n = t_0$, we have $\bar{t} = t_0$, and therefore $\lim_k (t_{n_k}, f(t_{n_k})) = (t_0, f(t_0))$. In particular, $\lim_k f(t_{n_k}) = f(t_0)$, which contradicts the inequality in (1).

Mark Lynch [44] uses the fact above to give a very interesting construction of a continuous nowhere-differentiable function on a closed interval $[0, 1]$; see his construction on page 361. $\square$

21.6 Remark.
The remainder of this section will be devoted to the Baire Category Theorem and some interesting consequences. We will need some metric-space terminology, some of it from §13. Let (S, d) be a metric space. For a subset E of S, its complement is $S \backslash E = \{s \in S : s \notin E\}$. A point s is in the interior of a set E, or interior to E, if $s \in V \subseteq E$ for some open subset V of S. A subset E of S is closed if $S \setminus E$ is open. A point s is in the closure of a set E if every open set containing s also contains an element of E. We write E^- for the set of points in the closure of E.

A subset D of S is dense in S if every nonempty open set U intersects D, i.e., $D \cap U \neq \varnothing$. For example, $\mathbb{Q}$ is dense in $\mathbb{R}$, since every nonempty open interval in $\mathbb{R}$ contains rationals; see 4.7 on page 25. Finally, a totally new definition: a subset E of S is *nowhere dense* in S if its closure E^- has empty interior. Obviously E is nowhere dense if and only if E^- is nowhere dense. Also, E is nowhere dense in S if and only if $S \setminus E^-$ is dense in S.

Each part of the next theorem is a variant of what is called the Baire Category Theorem. After that theorem, we will give the ultimate version of this theorem and explain the terminology "category." Just as in Example 4 on page 88, open balls $B_r(x_0) = \{x \in S : d(x, x_0) < r\}$ in S are open sets, and closed balls $C_r(x_0) = \{x \in S : d(x, x_0) \le r\}$ are closed sets. Note $C_s(x_0) \subseteq B_r(x_0)$ for $s < r$.

21.7 Theorem.
Let (S, d) be a complete metric space. Then

(a) *If (U_n) is a sequence of dense open subsets of S, then the intersection $X = \cap_{n=1}^{\infty} U_n$ is dense in S.*
(b) *If (F_n) is a sequence of closed subsets of S and if the union $F = \cup_{n=1}^{\infty} F_n$ contains a nonempty open set, then so does at least one of the sets F_n.*
(c) *The union of a sequence of nowhere dense subsets of S has dense complement.*
(d) *The space S is not a union of a sequence of nowhere dense subsets of S.*

Proof
All the hard work will be in proving: Part (a). Consider a nonempty open subset V of S; we need to show $X \cap V \neq \varnothing$. Since U_1 is open and dense and V is open, $U_1 \cap V$ is open and nonempty. So there exists $x_1 \in U_1 \cap V$ and $r_1 \le 1$ so that $C_{r_1}(x_1) \subseteq U_1 \cap V$. Since U_2 is open and dense and $B_{r_1}(x_1)$ is open, there exists $x_2 \in B_{r_1}(x_1) \cap U_2$ and $r_2 \le \frac{1}{2}$ so that $C_{r_2}(x_2) \subseteq B_{r_1}(x_1) \cap U_2$. Continuing, we obtain sequences (x_k) in S and positive numbers (r_k) so that

$$x_{k+1} \in B_{r_k}(x_k), \quad r_k \le \frac{1}{2^{k-1}} \quad \text{and} \quad C_{r_{k+1}}(x_{k+1}) \subseteq B_{r_k}(x_k) \cap U_{k+1}.$$

Note $d(x_k, x_{k+1}) < r_k$ for all k, so for $m < n$ we have

$$d(x_m, x_n) \le \sum_{k=m}^{n-1} d(x_k, x_{k+1}) < \sum_{k=m}^{n-1} \frac{1}{2^{k-1}} < \sum_{k=m}^{\infty} \frac{1}{2^{k-1}} = \frac{1}{2^{m-2}}.$$

Thus the sequence (x_k) is a Cauchy sequence in S. Since S is complete, we have $\lim_k x_k = x$ for some $x \in S$. Since we have

$$x_k \in C_{r_k}(x_k) \subseteq C_{r_n}(x_n) \quad \text{for} \quad k \ge n,$$

and since $C_{r_n}(x_n)$ is closed, x also belongs to C_{r_n}. This is true for all n, so

$$x \in \bigcap_{n=1}^{\infty} C_{r_n}(x_n) \subseteq \bigcap_{n=1}^{\infty} U_n = X.$$

Since we also have $x \in C_{r_1}(x_1) \subset V$, we conclude $x \in X \cap V$, as desired.

Part (b). Suppose $\cup_{n=1}^{\infty} F_n$ has nonempty interior. Then $S \setminus \cup_{n=1}^{\infty} F_n = \cap_{n=1}^{\infty}(S \setminus F_n)$ is not dense in S. Since the sets $S \setminus F_n$ are open, part (a) shows that at least one of them, say $S \setminus F_m$, is also not dense in S. Thus its complement F_m contains a nonempty open set.

Part (c). Let $A = \cup_{n=1}^{\infty} A_n$ be the union of a sequence of nowhere dense sets A_n. We need to show the complement of A, $S \setminus A$, is dense in S. If we replace each A_n by its closure, they will still be nowhere dense and $S \setminus A$ only gets smaller. In other words, we may assume each A_n is closed and nowhere dense. Part (b) implies that A contains no nonempty open set. Thus every nonempty open set intersects $S \setminus A$, and $S \setminus A$ is dense in S.

Part (d). This is obvious from part (c), since the complement of S is the empty set $\varnothing$. ■

Theorem 21.7 tells us in different ways that the union of a sequence of nowhere dense subsets of a complete metric space is small relative to the size of S. For this reason, it is common to divide the subsets of S into two categories. Category 1 consists of sets that are unions of sequences of nowhere dense subsets of S, and Category 2 consists of the other subsets of S. Here is part (d) of Theorem 21.7 restated using this language.

21.8 Baire Category Theorem.
A complete metric space (S, d) is of the second category in itself.

The expression "in itself" here deserves a comment. Many topological properties about sets $E \subseteq S$ depend on the set S. For example, the set $(0, 1]$ is not closed in $S = \mathbb{R}$, but it is closed in

$S = (0, \infty)$ viewed as a metric space. Category is another property that depends on S, so "in itself" in Theorem 21.8 stresses the reference to the complete metric space S.[1]

21.9 Corollary.
Each of the spaces $\mathbb{R}^k$ is of second category in itself.

Example 7
Any set that can be listed as a sequence is said to be "countable." Other sets are "uncountable."

- **(a)** Countable unions of sets of first category are sets of first category. For example, countable subsets of $\mathbb{R}$, and more generally of $\mathbb{R}^k$, are of first category, since single points $\{x\}$ in these spaces are nowhere dense. This and Corollary 21.9 give another (rather heavy-handed) proof that $\mathbb{R}$ is uncountable; compare Exercise 16.8.
- **(b)** Example 3 on page 70 shows the set $\mathbb{Q}$ of rationals is countable. Thus $\mathbb{Q}$ is of first category in $\mathbb{R}$, and the set of irrationals is of second category in $\mathbb{R}$.
- **(c)** Each straight line in the plane $\mathbb{R}^2$ is of first category in $\mathbb{R}^2$, so a countable union of straight lines is of first category in $\mathbb{R}^2$. $\square$

21.10 Discussion.
A closed set E in a metric space is said to be *perfect* if every point x in E is the limit of a sequence of points in $E \setminus \{x\}$; compare Proposition 13.9(b). This proposition also shows closed subsets of complete metric spaces are themselves complete metric spaces. In a perfect subset E of a complete metric space, single points are nowhere dense in E, so countable subsets of E are of first category. Therefore, nonempty perfect subsets of complete metric spaces are uncountable. The Cantor set (in Example 5 on page 89) is an interesting compact perfect set. $\square$

Recall Exercise 17.14, where a function on $\mathbb{R}$ is described that is continuous at each irrational and discontinuous at each rational. In

[1]One of the fine properties of compactness is: If a set is compact in some set, then it is compact in all sets.

the next example, we will show there are no functions on $\mathbb{R}$ that are continuous at *only* the rationals.

Example 8

(a) Consider any function f on $\mathbb{R}$. For each x in $\mathbb{R}$, let

$$\omega_f(x) = \inf_{\delta>0} \sup\{|f(y) - f(z)| : y, z \in (x - \delta, x + \delta)\}.$$

The function ω_f is the "oscillation function" of f. The function f is continuous at x if and only if $\omega_f(x) = 0$; see Exercise 21.13.

We show each set $\{x \in \mathbb{R} : \omega_f(x) < \epsilon\}$ is open in $\mathbb{R}$; $\epsilon > 0$. Suppose $\omega_f(x_0) < \epsilon$; we need to show this inequality holds in an interval containing x_0. By the definition, there exists $\delta > 0$ such that

$$\sup\{|f(y) - f(z)| : y, z \in (x_0 - \delta, x_0 + \delta)\} < \epsilon.$$

For $x \in (x_0 - \frac{\delta}{2}, x_0 + \frac{\delta}{2})$ we have $(x - \frac{\delta}{2}, x + \frac{\delta}{2}) \subseteq (x_0 - \delta, x_0 + \delta)$, and therefore

$$\begin{aligned}\omega_f(x) &\le \sup\left\{|f(y) - f(z)| : y, z \in \left(x - \frac{\delta}{2}, x + \frac{\delta}{2}\right)\right\} \\ &\le \sup\{|f(y) - f(z)| : y, z \in (x_0 - \delta, x_0 + \delta)\} < \epsilon.\end{aligned}$$

(b) There is no function f on $\mathbb{R}$ that is continuous only on the set $\mathbb{Q}$ of rational numbers. Otherwise, by part (a) the equality

$$\mathbb{Q} = \{x \in \mathbb{R} : \omega_f(x) = 0\} = \bigcap_{n=1}^{\infty} \left\{x \in \mathbb{R} : \omega_f(x) < \frac{1}{n}\right\}$$

would express $\mathbb{Q}$ as the intersection of a sequence of open subsets of $\mathbb{R}$, contrary to Exercise 21.11. □

21.11 Theorem.

No nondegenerate interval I in $\mathbb{R}$ can be written as the disjoint union of two or more nondegenerate closed intervals. (An interval is "nondegenerate" if it has more than one point.)

Proof

We prove the theorem for $I = [a, b]$, $a < b$, and leave the other cases to Exercise 21.14. Assume I is the disjoint union of two or more nondegenerate closed intervals. It is clear there are infinitely many such intervals, because otherwise they could be listed in order, and there would be points in I between the first and second interval. It is also clear that the family can be listed as a sequence (i.e., is countable), since there is a one-to-one function from the collection of disjoint intervals into $\mathbb{Q}$.

So we assume $I = \cup_{k=1}^{\infty}[a_k, b_k]$. By Discussion 13.7(iii) on page 87, the union $\cup_{k=1}^{\infty}(a_k, b_k)$ is open, so that $K = I \setminus \cup_{k=1}^{\infty}(a_k, b_k)$ is closed. Then K is the set of endpoints of the removed intervals, so K is nonempty and countable. We will show K is perfect, and this will be a contradiction by Discussion 21.10. Consider x in K; we may assume $x \neq b$. Consider an interval $(x, x+h)$ where $h > 0$. Some closed interval $[a_m, b_m]$ must intersect $(x, x+h)$, and $[a_m, b_m]$ cannot contain x in K. Thus $x < a_m < x+h$. Since a_m is in K, the interval $(x, x+h)$ intersects K, and since $h > 0$ is arbitrary, x is a limit of points in $K \setminus \{x\}$. This shows K is perfect, completing the proof. ■

A close examination of Example 8(b) and Theorem 21.11 shows they both depend on the Baire Category Theorem 21.8. Other applications of this theorem appear in Theorem 38.3 and twice in Theorem 38.5.

Exercises

21.1 Show that if the functions $f_1, f_2, \ldots, f_k$ in Proposition 21.2 are uniformly continuous, then so is γ.

21.2 Consider $f: S \to S^*$ where (S, d) and (S^*, d^*) are metric spaces. Show that f is continuous at $s_0 \in S$ if and only if

> for every open set U in S^* containing $f(s_0)$, there is an open set V in S containing s_0 such that $f(V) \subseteq U$.

21.3 Let (S, d) be a metric space and choose $s_0 \in S$. Show $f(s) = d(s, s_0)$ defines a uniformly continuous real-valued function f on S.

21.4 Consider $f: S \to \mathbb{R}$ where (S, d) is a metric space. Show the following are equivalent:

(i) f is continuous;
(ii) $f^{-1}((a, b))$ is open in S for all $a < b$;
(iii) $f^{-1}((a, b))$ is open in S for all rational $a < b$.

21.5 Let E be a noncompact subset of $\mathbb{R}^k$.

(a) Show there is an unbounded continuous real-valued function on E. *Hint*: Either E is unbounded or else its closure E^- contains $\boldsymbol{x}_0 \notin E$. In the latter case, use $\frac{1}{g}$ where $g(\boldsymbol{x}) = d(\boldsymbol{x}, \boldsymbol{x}_0)$.

(b) Show there is a bounded continuous real-valued function on E that does not assume its maximum on E.

21.6 For metric spaces (S_1, d_1), (S_2, d_2), (S_3, d_3), prove that if $f\colon S_1 \to S_2$ and $g\colon S_2 \to S_3$ are continuous, then $g \circ f$ is continuous from S_1 into S_3. *Hint*: It is somewhat easier to use Theorem 21.3 than to use the definition.

21.7 **(a)** Observe that if $E \subseteq S$ where (S, d) is a metric space, then (E, d) is also a metric space. In particular, if $E \subseteq \mathbb{R}$, then $d(a, b) = |a - b|$ for $a, b \in E$ defines a metric on E.

(b) For $\gamma\colon [a, b] \to \mathbb{R}^k$, give the definition of continuity of γ.

21.8 Let (S, d) and (S^*, d^*) be metric spaces. Show that if $f\colon S \to S^*$ is uniformly continuous, and if (s_n) is a Cauchy sequence in S, then $(f(s_n))$ is a Cauchy sequence in S^*.

21.9 We say a function f *maps* a set E onto a set F provided $f(E) = F$.

(a) Show there is a continuous function mapping the unit square

$$\{(x_1, x_2) \in \mathbb{R}^2 : 0 \le x_1 \le 1, 0 \le x_2 \le 1\}$$

onto $[0, 1]$.

(b) Do you think there is a continuous function mapping $[0, 1]$ onto the unit square?

21.10 Show there exist continuous functions

(a) Mapping $(0, 1)$ onto $[0, 1]$,

(b) Mapping $(0, 1)$ onto $\mathbb{R}$,

(c) Mapping $[0, 1] \cup [2, 3]$ onto $[0, 1]$.

(d) Explain why there are no continuous functions mapping $[0, 1]$ onto $(0, 1)$ or $\mathbb{R}$.

21.11 Show the set $\mathbb{Q}$ of rational numbers is not the intersection of a sequence of open sets, so that the set $\mathbb{R} \setminus \mathbb{Q}$ of irrational numbers is not the union of a sequence of closed subsets of $\mathbb{R}$.

21.12 Give an example of an infinite disjoint sequence of subsets of $\mathbb{R}$, each of which is of second category in $\mathbb{R}$.

21.13 Let ω_f be the oscillation function in Example 8(a). Show that f is continuous at x in $\mathbb{R}$ if and only if $\omega_f(x) = 0$.

21.14 Complete the proof of Theorem 21.11. *Hint*: Assume a nondegenerate interval I in $\mathbb{R}$ is the union of two or more disjoint nondegenerate closed intervals. Select points $x < y$ in I, so that they are interior to two distinct such subintervals. Show $[x, y]$ is also the union of two or more disjoint nondegenerate closed intervals, contradicting the case covered in the proof of Theorem 21.11.

§22 * More on Metric Spaces: Connectedness

Consider a subset E of $\mathbb{R}$ that is not an interval. As noted in the proof of Corollary 18.3, the property

$$y_1, y_2 \in E \quad \text{and} \quad y_1 < y < y_2 \quad \text{imply} \quad y \in E$$

must fail. So there exist y_1, y_2, y in $\mathbb{R}$ such that

$$y_1 < y < y_2, \qquad y_1, y_2 \in E, \qquad y \notin E. \tag{*}$$

The set E is not "connected" because y separates E into two pieces. Put another way, if we set $U_1 = (-\infty, y)$ and $U_2 = (y, \infty)$, then we obtain open sets so that $E \cap U_1$ and $E \cap U_2$ are disjoint nonempty sets who union is E. The last observation can be promoted to a useful general definition.

22.1 Definition.
Let E be a subset of a metric space (S, d). The set E is *disconnected* if one of the following two equivalent conditions holds.
(a)[2] There are open subsets U_1 and U_2 of S such that

$$(E \cap U_1) \cap (E \cap U_2) = \varnothing \quad \text{and} \quad E = (E \cap U_1) \cup (E \cap U_2), \tag{1}$$
$$E \cap U_1 \neq \varnothing \quad \text{and} \quad E \cap U_2 \neq \varnothing. \tag{2}$$

[2]Readers familiar with relative topologies will recognize this as stating that E is a disjoint union of nonempty relatively-open subsets of E.

(b) There are disjoint nonempty subsets A and B of E such that $E = A \cup B$ and neither set intersects the closure of the other set, i.e.,

$$A^- \cap B = \varnothing \quad \text{and} \quad A \cap B^- = \varnothing. \tag{3}$$

A set E is *connected* if it is not disconnected.

Remark. We show that conditions (a) and (b) above are equivalent.

If (a) holds, set $A = E \cap U_1$ and $B = E \cap U_2$. To check (b), we need only verify (3). If, for example, $A^- \cap B \neq \varnothing$, then there exists $b \in A^- \cap B \subseteq U_2$, so there is $r > 0$ so that $\{s \in S : d(s,b) < r\} \subseteq U_2$. Since b is in A^-, Proposition 13.9(c) shows there is a in A such that $d(a,b) < r$. But then $a \in U_2 \cap E = B$, a contradiction, since A and B are disjoint. Thus (3) holds, and condition (b) holds.

Now suppose that (b) holds. Let $U_1 = S \setminus B^-$ and $U_2 = S \setminus A^-$. To verify (a), it suffices to show $E \cap U_1 = A$ and $E \cap U_2 = B$. For example, $x \in E \cap U_1$ implies $x \in S \setminus B$ implies $x \notin B$. Since $x \in A \cup B$, we conclude $x \in A$. Finally, if $x \in A$, then $x \notin B^-$ by (3), so $x \in S \setminus B^- = U_1$. Therefore x is in $E \cap U_1$. We've shown $E \cap U_1 = A$.

Example 1
As noted before the definition, sets in $\mathbb{R}$ that are not intervals are disconnected. Conversely, intervals in $\mathbb{R}$ are connected. To prove this from the definition, we will assume the contrary and obtain a contradiction. So we assume there is an interval I and open sets U_1 and U_2 disconnecting I as in Definition 22.1(a). Select $a \in I \cap U_1$ and $b \in I \cap U_2$. We may suppose $a < b$. Now the interval $[a,b]$ satisfies the same conditions in Definition 22.1(a) as I. So we may assume $I = [a,b]$, $a \in U_1$ and $b \in U_2$.

Let $t = \sup [a,b) \cap U_1$, so that $t > a$. Since $(b-\epsilon, b] \subseteq U_2$ for some $\epsilon > 0$, we also have $t < b$. If t is in U_1, then $[t, t+\epsilon) \subseteq U_1 \cap [a,b]$ for some $\epsilon > 0$, but then $t = \sup[a,b) \cap U_1 \geq t + \epsilon$. If $t \in U_2$, then $(t-\epsilon, t] \subseteq U_2 \cap [a,b]$, which would imply $t = \sup[a,b) \cap U_1 \leq t - \epsilon$. Each possibility leads to a contradiction, so $[a,b]$ and the original interval I are connected. $\square$

22.2 Theorem.
Consider metric spaces (S, d), (S^*, d^*), and let $f: S \to S^*$ be continuous. If E is a connected subset of S, then $f(E)$ is a connected subset of S^*.

Proof
Assume $f(E)$ is not connected in S^*. Then there exist open sets V_1 and V_2 in S^* such that

$$(f(E) \cap V_1) \cap (f(E) \cap V_2) = \varnothing, \quad f(E) = (f(E) \cap V_1) \cup (f(E) \cap V_2),$$
$$f(E) \cap V_1 \neq \varnothing \quad \text{and} \quad f(E) \cap V_2 \neq \varnothing.$$

Let $U_1 = f^{-1}(V_1)$ and $U_2 = f^{-1}(V_2)$. Then U_1 and U_2 are open sets in S that separate E as in Definition 22.1(a). Thus E is not connected, a contradiction. ■

The next corollary generalizes Theorem 18.2 and its corollary.

22.3 Corollary.
Let f be a continuous real-valued function on a metric space (S, d). If E is a connected subset of S, then $f(E)$ is an interval in $\mathbb{R}$. In particular, f has the intermediate value property.

Example 2
Curves are connected. That is, if γ is a path in $\mathbb{R}^k$ as described in Example 3 of §21, page 165 and I is a subinterval of $\mathbb{R}$, then the image $\gamma(I)$ is connected in $\mathbb{R}^k$. □

22.4 Definition.
A subset E of a metric space (S, d) is said to be *path-connected* if, for each pair s, t of points in E, there exists a continuous function $\gamma: [a, b] \to E$ such that $\gamma(a) = s$ and $\gamma(b) = t$. We call γ a *path*.

22.5 Theorem.
If E in (S, d) is path-connected, then E is connected. [*The failure of the converse is illustrated in Exercise 22.4.*]

Proof
Assume E is disconnected by open sets U_1 and U_2:

$$(E \cap U_1) \cap (E \cap U_2) = \varnothing \quad \text{and} \quad E = (E \cap U_1) \cup (E \cap U_2), \tag{1}$$
$$E \cap U_1 \neq \varnothing \quad \text{and} \quad E \cap U_2 \neq \varnothing. \tag{2}$$

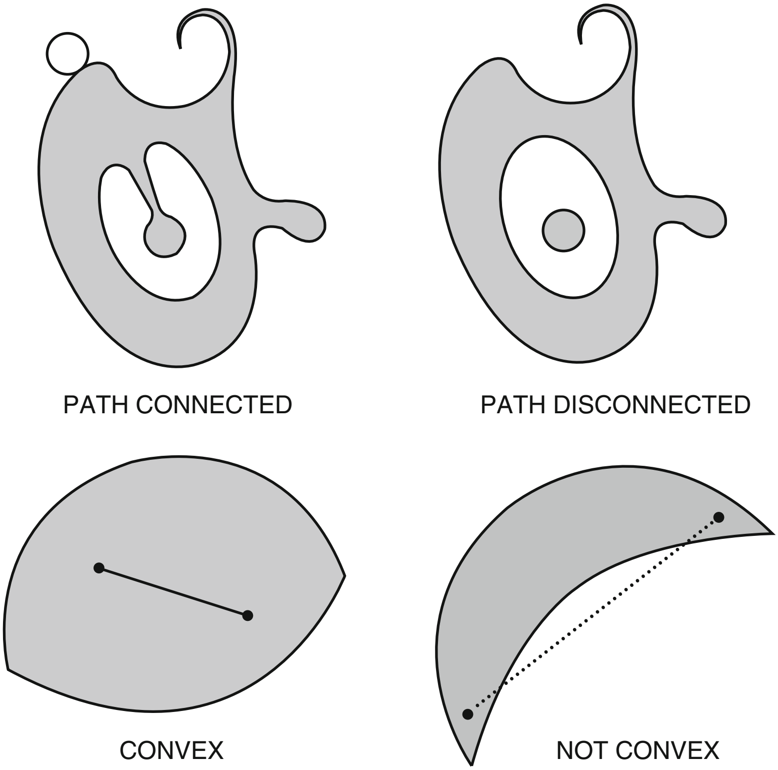

FIGURE 22.1

Select $s \in E \cap U_1$ and $t \in E \cap U_2$. Let $\gamma\colon [a, b] \to E$ be a path where $\gamma(a) = s$ and $\gamma(b) = t$. Let $F = \gamma([a, b])$. Then (1) and (2) hold with F in place of E. Thus F is disconnected, but F is connected by Theorem 22.2. ■

Figure 22.1 gives a path-connected set and a disconnected set in $\mathbb{R}^2$.

Example 3
Many familiar sets in $\mathbb{R}^k$ such as the open ball $\{\boldsymbol{x} : d(\boldsymbol{x}, \mathbf{0}) < r\}$, the closed ball $\{\boldsymbol{x} : d(\boldsymbol{x}, \mathbf{0}) \leq r\}$ and the k-dimensional cube

$$\{\boldsymbol{x} : \max\{|x_j| : j = 1, 2, \ldots, k\} \leq 1\}$$

are convex. A subset E of $\mathbb{R}^k$ is *convex* if

$$\boldsymbol{x}, \boldsymbol{y} \in E \quad \text{and} \quad 0 < t < 1 \quad \text{imply} \quad t\boldsymbol{x} + (1-t)\boldsymbol{y} \in E,$$

i.e., whenever E contains two points it contains the line segment connecting them. See Fig. 22.1. Convex sets E in $\mathbb{R}^k$ are always path-connected. This is because $\gamma(t) = t\boldsymbol{x} + (1-t)\boldsymbol{y}$ defines a path $\gamma\colon [0,1] \to E$ such that $\gamma(0) = \boldsymbol{y}$ and $\gamma(1) = \boldsymbol{x}$. For more details, see any book on several variable calculus. □

Example 4

This example is adapted from [15]. Suppose I is an interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ is a function with graph $G = \{(x, f(x)) : x \in I\}$. Then f is continuous if and only if G is path-connected. Note that "path-connected" here cannot be changed to "connected;" see Exercise 22.13.

If f is continuous, then for each $(a, f(a))$ and $(b, f(b))$ in G, $\gamma(t) = (t, f(t))$ is a path connecting these points. Note that γ is continuous by Proposition 21.2.

For the converse, assume G is path-connected. Consider a, b in I where $a < b$. Since G is path-connected, there is a continuous function $\gamma : [0,1] \to G$ satisfying $\gamma(0) = (a, f(a))$ and $\gamma(1) = (b, f(b))$. By Proposition 21.2, we can write $\gamma(t) = (x(t), y(t))$ where x and y are continuous real-valued functions on $[0,1]$. Of course, $y(t) = f(x(t))$ for all t. First, we claim

$$\{(z, f(z)) : z \in [a,b]\} \subseteq \gamma([0,1]). \tag{1}$$

Since $x(t)$ is continuous on $[0,1]$, $x(0) = a$ and $x(1) = b$, the Intermediate Value Theorem 18.2 for continuous functions shows that if z is in (a,b), then $x(t) = z$ for some $t \in (0,1)$. Thus $(z, f(z)) = (x(t), f(x(t))) = \gamma(t)$. Since the cases $z = a$ and $z = b$ are trivial, this implies (1).

Next we show f satisfies IVP, the intermediate value property:

(IVP) If $a, b \in I$, $a < b$ and c is between $f(a)$ and $f(b)$, then $f(x) = c$ for some $x \in [a,b]$.

Let $\alpha = \sup\{t \in [0,1] : x(t) \le a\}$ and $\beta = \inf\{t \in [0,1] : x(t) \ge b\}$. Thus $0 \le \alpha < \beta \le 1$. Since $x(t)$ is continuous, we see that $x([\alpha, \beta]) = [a,b]$. Now we apply the Intermediate Value Theorem again, this time to $f(x(t))$ on $[\alpha, \beta]$. Since $f(x(\alpha)) = f(a)$ and

$f(x(\beta)) = f(b)$, we have $f(x(t)) = c$ for some t in (α, β). For this t we have $x(t) \in [a, b]$, so we have verified IVP.

Now assume f is not continuous on I. Then there exists $x_0 \in I$, $\epsilon > 0$, and a sequence (x_n) in I that converges to x_0 satisfying $|f(x_n) - f(x_0)| > \epsilon$ for all n. Then $f(x_n) > f(x_0) + \epsilon > f(x_0)$ for infinitely many n, or else $f(x_n) < f(x_0) - \epsilon < f(x_0)$ for infinitely many n. Passing to a subsequence, we may assume $f(x_n) > f(x_0) + \epsilon > f(x_0)$ for all n, say. By the IVP, for each n there is y_n between x_n and x_0 so that $f(y_n) = f(x_0) + \epsilon$. Then $(y_n, f(y_n))$ is a sequence in G satisfying

$$\lim_n y_n = x_0 \quad \text{and} \quad \lim(y_n, f(y_n)) = (x_0, f(x_0) + \epsilon). \tag{2}$$

Moreover, there exist $a < b$ in I such that x_0 and all y_n belong to $[a, b]$. We apply (1) to this a and b and γ described there. So, for each $n \in \mathbb{N}$ there exists $t_n \in [0, 1]$ so that $\gamma(t_n) = (y_n, f(y_n))$. Passing to a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ of (t_n), we may assume that $\lim_k t_{n_k} = t_0$ for some t_0 in $[0, 1]$. Since γ is continuous, we conclude from (2) that

$$\gamma(t_0) = \lim_k \gamma(t_{n_k}) = \lim_k (y_{n_k}, f(y_{n_k})) = (x_0, f(x_0) + \epsilon).$$

Since this limit is not in G, and $\gamma([0, 1]) \subset G$, we have a contradiction and f is continuous on I. □

We end this section with a discussion of some very different metric spaces. The points in these spaces are actually functions themselves.

22.6 Definition.
Let S be a subset of $\mathbb{R}$. Let $C(S)$ be the set of all bounded continuous real-valued functions on S and, for $f, g \in C(S)$, let

$$d(f, g) = \sup\{|f(x) - g(x)| : x \in S\}.$$

With this definition, $C(S)$ becomes a metric space [Exercise 22.6]. In this metric space, a sequence (f_n) converges to a point [function!] f provided $\lim_{n\to\infty} d(f_n, f) = 0$, that is

$$\lim_{n\to\infty} \sup\{|f_n(x) - f(x)| : x \in S\} = 0. \tag{*}$$

Put another way, for each $\epsilon > 0$ there exists a number N such that

$$|f_n(x) - f(x)| < \epsilon \quad \text{for all} \quad x \in S \quad \text{and} \quad n > N.$$

We will study this important concept in the next chapter, but without using metric space terminology. See Definition 24.2 and Remark 24.4 where (*) is called *uniform convergence*.

A sequence (f_n) in $C(S)$ is a Cauchy sequence with respect to the metric d exactly when it is *uniformly Cauchy* as defined in Definition 25.3. In our metric space terminology, Theorems 25.4 and 24.3 tell us that $C(S)$ is a *complete* metric space.

Exercises

22.1 Show there do not exist continuous functions

(a) Mapping $[0,1]$ onto $[0,1] \cup [2,3]$,

(b) Mapping $(0,1)$ onto $\mathbb{Q}$.

22.2 Show $\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ is a connected subset of $\mathbb{R}^2$.

22.3 Prove that if E is a connected subset of a metric space (S, d), then its closure E^- is also connected.

22.4 Consider the following subset of $\mathbb{R}^2$:

$$E = \left\{ \left(x, \sin \frac{1}{x} \right) : x \in (0,1] \right\};$$

E is simply the graph of $g(x) = \sin \frac{1}{x}$ along the interval $(0,1]$.

(a) Determine its closure E^-. See Fig. 19.4.

(b) Show E^- is connected.

(c) Show E^- is not path-connected.

22.5 Let E and F be connected sets in some metric space.

(a) Prove that if $E \cap F \neq \varnothing$, then $E \cup F$ is connected.

(b) Give an example to show $E \cap F$ need not be connected. Incidentally, the empty set *is* connected.

22.6 **(a)** Show $C(S)$ given in Definition 22.6 is a metric space.

(b) Why did we require the functions in $C(S)$ to be bounded when no such requirement appears in Definition 24.2?

22.7 Explain why the metric space B in Exercise 13.3 can be regarded as $C(\mathbb{N})$.

22.8 Consider $C(S)$ for a subset S of $\mathbb{R}$. For a fixed s_0 in S, define $F(f) = f(s_0)$. Show F is a uniformly continuous real-valued function on the metric space $C(S)$.

22.9 Consider $f, g \in C(S)$ where $S \subseteq \mathbb{R}$. Let $F(t) = tf + (1-t)g$. Show F is a uniformly continuous function from $\mathbb{R}$ into $C(S)$.

22.10 Let f be a (fixed) uniformly continuous function in $C(\mathbb{R})$. For each $x \in \mathbb{R}$, let f_x be the function defined by $f_x(y) = f(x+y)$ for $y \in \mathbb{R}$. Let $F(x) = f_x$; show F is uniformly continuous from $\mathbb{R}$ into $C(\mathbb{R})$.

22.11 Consider $C(S)$ where $S \subseteq \mathbb{R}$, and let $\mathcal{E}$ consist of all f in $C(S)$ such that $\sup\{|f(x)| : x \in S\} \leq 1$.

(a) Show $\mathcal{E}$ is closed in $C(S)$.

(b) Show $C(S)$ is connected.

(c) Show $\mathcal{E}$ is connected.

22.12 Consider a subset $\mathcal{E}$ of $C(S)$, $S \subseteq \mathbb{R}$. For this exercise, we say a function f_0 in $\mathcal{E}$ is *interior* to $\mathcal{E}$ if there exists a finite subset F of S and an $\epsilon > 0$ such that

$$\{f \in C(S) : |f(x) - f_0(x)| < \epsilon \text{ for } x \in F\} \subseteq \mathcal{E}.$$

The set $\mathcal{E}$ is *open* if every function in $\mathcal{E}$ is interior to $\mathcal{E}$.

(a) Reread Discussion 13.7.

(b) Show the family of open sets defined above forms a topology for $C(S)$. *Remarks.* This topology is different from the one given by the metric in Definition 22.6. In fact, this topology does not come from any metric at all! It is called the *topology of pointwise convergence* and can be used to study the convergence in Definition 24.1 just as the metric in Definition 22.6 can be used to study the convergence in Definition 24.2.

22.13 Let f be the function defined on $[0, \infty)$ by $f(x) = (x, \sin(1/x))$ for $x > 0$ and $f(0) = 0$. Show that the graph of f is connected but not path-connected; compare Exercise 22.4. Is f continuous on $[0, \infty)$?

22.14 Consider a bounded sequence (s_n) in $\mathbb{R}$ and its set S of subsequential limits as in Theorem 11.8. Show that if $\lim_n(s_{n+1} - s_n) = 0$, then S is a connected interval in $\mathbb{R}$. This is from [11]. *Hint*: Consider $a < c < b$ where a and b are in S, and show c is in S. Use Theorem 11.2.