Chapter 4

Functional Limits and Continuity

4.1 Discussion: Examples of Dirichlet and Thomae

Although it is a common practice in calculus courses to discuss continuity before differentiation, historically mathematicians' attention to the concept of continuity came long after the derivative was in wide use. Pierre de Fermat (1601–1665) was using tangent lines to solve optimization problems as early as 1629. On the other hand, it was not until around 1820 that Cauchy, Bolzano, Weierstrass, and others began to characterize continuity in terms more rigorous than prevailing intuitive notions such as "unbroken curves" or "functions which have no jumps or gaps." The basic reason for this two-hundred year waiting period lies in the fact that, for most of this time, the very notion of *function* did not really permit discontinuities. Functions were entities such as polynomials, sines, and cosines, always smooth and continuous over their relevant domains. The gradual liberation of the term function to its modern understanding—a rule associating a unique output with a given input—was simultaneous with 19th century investigations into the behavior of infinite series. Extensions of the power of calculus were intimately tied to the ability to represent a function $f(x)$ as a limit of polynomials (called a *power series*) or as a limit of sums of sines and cosines (called a *trigonometric* or *Fourier* series). A typical question for Cauchy and his contemporaries was whether the continuity of the limiting polynomials or trigonometric functions necessarily implied that the limit f would also be continuous.

Sequences and series of functions are the topics of Chapter 6. What is relevant at this moment is that we realize why the issue of finding a rigorous

S. Abbott, *Understanding Analysis*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4939-2712-8_4

Figure 4.1: Dirichlet's Function, $g(x)$.

definition for continuity finally made its way to the fore. Any significant progress on the question of whether the limit of continuous functions is continuous (for Cauchy and for us) necessarily depends on a definition of continuity that does not rely on imprecise notions such as "no holes" or "gaps." With a mathematically unambiguous definition for the limit of a sequence in hand, we are well on our way toward a rigorous understanding of continuity.

Given a function f with domain $A \subseteq \mathbf{R}$, we want to define continuity at a point $c \in A$ to mean that if $x \in A$ is chosen *near* c, then $f(x)$ will be *near* $f(c)$. Symbolically, we will say f is continuous at c if

$$\lim_{x \to c} f(x) = f(c).$$

The problem is that, at present, we only have a definition for the limit of a sequence, and it is not entirely clear what is meant by $\lim_{x \to c} f(x)$. The subtleties that arise as we try to fashion such a definition are well-illustrated via a family of examples, all based on an idea of the prominent German mathematician, Peter Lejeune Dirichlet. Dirichlet's idea was to define a function g in a piecewise manner based on whether or not the input variable x is rational or irrational. Specifically, let

$$g(x) = \begin{cases} 1 & \text{if } x \in \mathbf{Q} \\ 0 & \text{if } x \notin \mathbf{Q}. \end{cases}$$

The intricate way that $\mathbf{Q}$ and $\mathbf{I}$ fit inside of $\mathbf{R}$ makes an accurate graph of g technically impossible to draw, but Figure 4.1 illustrates the basic idea.

Does it make sense to attach a value to the expression $\lim_{x \to 1/2} g(x)$? One idea is to consider a sequence $(x_n) \to 1/2$. Using our notion of the limit of a sequence, we might try to define $\lim_{x \to 1/2} g(x)$ as simply the limit of the sequence $g(x_n)$. But notice that this limit depends on how the sequence (x_n) is chosen. If each x_n is rational, then

$$\lim_{n \to \infty} g(x_n) = 1.$$

On the other hand, if x_n is irrational for each n, then

$$\lim_{n \to \infty} g(x_n) = 0.$$

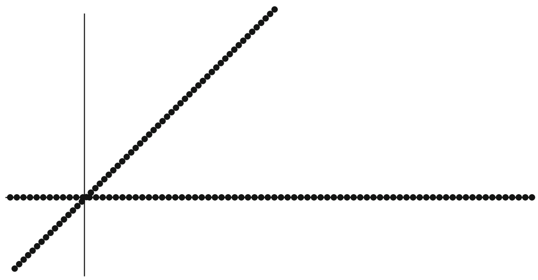


Figure 4.2: MODIFIED DIRICHLET FUNCTION, $h(x)$.

This unacceptable situation demands that we work harder on our definition of functional limits. Generally speaking, we want the value of $\lim_{x\to c} g(x)$ to be independent of how we approach c. In this particular case, the definition of a functional limit that we agree on should lead to the conclusion that

$$\lim_{x\to 1/2} g(x) \quad \text{does not exist.}$$

Postponing the search for formal definitions for the moment, we should nonetheless realize that Dirichlet's function is not continuous at $c = 1/2$. In fact, the real significance of this function is that there is nothing unique about the point $c = 1/2$. Because both $\mathbf{Q}$ and $\mathbf{I}$ (the set of irrationals) are dense in the real line, it follows that for any $z \in \mathbf{R}$ we can find sequences $(x_n) \subseteq \mathbf{Q}$ and $(y_n) \subseteq \mathbf{I}$ such that

$$\lim x_n = \lim y_n = z.$$

(See Example 3.2.9 (iii).) Because

$$\lim g(x_n) \neq \lim g(y_n),$$

the same line of reasoning reveals that $g(x)$ is not continuous at z. In the jargon of analysis, Dirichlet's function is a *nowhere-continuous* function on $\mathbf{R}$.

What happens if we adjust the definition of $g(x)$ in the following way? Define a new function h (Fig. 4.2) on $\mathbf{R}$ by setting

$$h(x) = \begin{cases} x & \text{if } x \in \mathbf{Q} \\ 0 & \text{if } x \notin \mathbf{Q}. \end{cases}$$

If we take c different from zero, then just as before we can construct sequences $(x_n) \to c$ of rationals and $(y_n) \to c$ of irrationals so that

$$\lim h(x_n) = c \quad \text{and} \quad \lim h(y_n) = 0.$$

Thus, h is not continuous at every point $c \neq 0$.

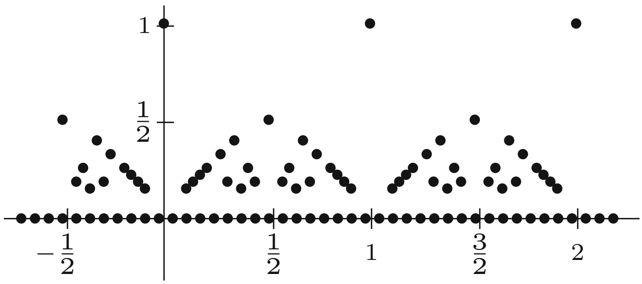


Figure 4.3: THOMAE'S FUNCTION, $t(x)$.

If $c = 0$, however, then these two limits are both equal to $h(0) = 0$. In fact, it appears as though no matter how we construct a sequence (z_n) converging to zero, it will always be the case that $\lim h(z_n) = 0$. This observation goes to the heart of what we want functional limits to entail. To assert that

$$\lim_{x \to c} h(x) = L$$

should imply that

$$h(z_n) \to L \quad \textit{for all} \text{ sequences} \quad (z_n) \to c.$$

For reasons not yet apparent, it is beneficial to fashion the definition for functional limits in terms of neighborhoods constructed around c and L. We will quickly see, however, that this topological formulation is equivalent to the sequential characterization we have arrived at here.

To this point, we have been discussing continuity of a function at a particular point in its domain. This is a significant departure from thinking of continuous functions as curves that can be drawn without lifting the pen from the paper, and it leads to some fascinating questions. In 1875, K.J. Thomae discovered the function

$$t(x) = \begin{cases} 1 & \text{if } x = 0 \\ 1/n & \text{if } x = m/n \in \mathbf{Q}\backslash\{0\} \text{ is in lowest terms with } n > 0 \\ 0 & \text{if } x \notin \mathbf{Q}. \end{cases}$$

If $c \in \mathbf{Q}$, then $t(c) > 0$. Because the set of irrationals is dense in $\mathbf{R}$, we can find a sequence (y_n) in $\mathbf{I}$ converging to c. The result is that

$$\lim t(y_n) = 0 \neq t(c),$$

and Thomae's function (Fig. 4.3) fails to be continuous at any rational point.

The twist comes when we try this argument on some irrational point in the domain such as $c = \sqrt{2}$. All irrational values get mapped to zero by t, so the natural thing would be to consider a sequence (x_n) of rational numbers that converges to $\sqrt{2}$. Now, $\sqrt{2} \approx 1.414213\ldots$, so a good start on a particular sequence of rational approximations for $\sqrt{2}$ might be

$$\left(1, \frac{14}{10}, \frac{141}{100}, \frac{1414}{1000}, \frac{14142}{10000}, \frac{141421}{100000}, \ldots\right).$$

But notice that the denominators of these fractions are getting larger. In this case, the sequence $t(x_n)$ begins,

$$\left(1, \frac{1}{5}, \frac{1}{100}, \frac{1}{500}, \frac{1}{5000}, \frac{1}{100000}, \ldots\right)$$

and is fast approaching $0 = t(\sqrt{2})$. We will see that this always happens. The closer a rational number is chosen to a fixed irrational number, the larger its denominator must necessarily be. As a consequence, Thomae's function has the bizarre property of being continuous at every irrational point on $\mathbf{R}$ and discontinuous at every rational point.

Is there an example of a function with the opposite property? In other words, does there exist a function defined on all of $\mathbf{R}$ that is continuous on $\mathbf{Q}$ but fails to be continuous on $\mathbf{I}$? Can the set of discontinuities of a particular function be arbitrary? If we are given some set $A \subseteq \mathbf{R}$, is it always possible to find a function that is continuous only on the set A^c? In each of the examples in this section, the functions were defined to have erratic oscillations around points in the domain. What conclusions can we draw if we restrict our attention to functions that are somewhat less volatile? One such class is the set of so-called *monotone* functions, which are either increasing or decreasing on a given domain. What might we be able to say about the set of discontinuities of a monotone function on $\mathbf{R}$?

4.2 Functional Limits

Consider a function $f : A \to \mathbf{R}$. Recall that a limit point c of A is a point with the property that every ϵ-neighborhood $V_\epsilon(c)$ intersects A in some point other than c. Equivalently, c is a limit point of A if and only if $c = \lim x_n$ for some sequence $(x_n) \subseteq A$ with $x_n \neq c$. It is important to remember that limit points of A do not necessarily belong to the set A unless A is closed.

If c is a limit point of the domain of f, then, intuitively, the statement

$$\lim_{x \to c} f(x) = L$$

is intended to convey that values of $f(x)$ get arbitrarily close to L as x is chosen closer and closer to c. The issue of what happens when $x = c$ is irrelevant from the point of view of functional limits. In fact, c need not even be in the domain of f.

The structure of the definition of functional limits follows the "challenge–response" pattern established in the definition for the limit of a sequence. Recall that given a sequence (a_n), the assertion $\lim a_n = L$ implies that for every

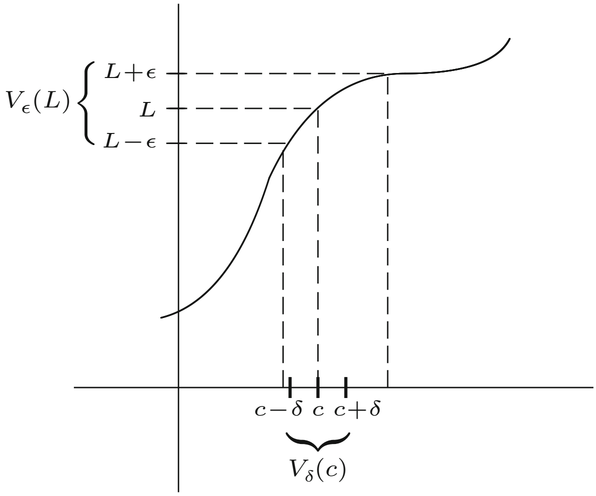


Figure 4.4: DEFINITION OF FUNCTIONAL LIMIT.

ϵ-neighborhood $V_\epsilon(L)$ centered at L, there is a point in the sequence—call it a_N—after which all of the terms a_n fall in $V_\epsilon(L)$. Each ϵ-neighborhood represents a particular challenge, and each N is the respective response. For functional limit statements such as $\lim_{x\to c} f(x) = L$, the challenges are still made in the form of an arbitrary ϵ-neighborhood around L, but the response this time is a δ-neighborhood centered at c.

Definition 4.2.1 (Functional Limit). Let $f : A \to \mathbf{R}$, and let c be a limit point of the domain A. We say that $\lim_{x\to c} f(x) = L$ provided that, for all $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $0 < |x - c| < \delta$ (and $x \in A$) it follows that $|f(x) - L| < \epsilon$.

This is often referred to as the "ϵ–δ version" of the definition for functional limits. Recall that the statement

$$|f(x) - L| < \epsilon \text{ is equivalent to } f(x) \in V_\epsilon(L).$$

Likewise, the statement

$$|x - c| < \delta \text{ is satisfied if and only if } x \in V_\delta(c).$$

The additional restriction $0 < |x - c|$ is just an economical way of saying $x \neq c$. Recasting Definition 4.2.1 in terms of neighborhoods—just as we did for the definition of convergence of a sequence in Section 2.2—amounts to little more than a change of notation, but it does help emphasize the geometrical nature of what is happening (Fig. 4.4).

Definition 4.2.1B (Functional Limit: Topological Version). Let c be a limit point of the domain of $f : A \to \mathbf{R}$. We say $\lim_{x\to c} f(x) = L$ provided that, for every ϵ-neighborhood $V_\epsilon(L)$ of L, there exists a δ-neighborhood $V_\delta(c)$ around c with the property that for all $x \in V_\delta(c)$ different from c (with $x \in A$) it follows that $f(x) \in V_\epsilon(L)$.

The parenthetical reminder "$(x \in A)$" present in both versions of the definition is included to ensure that x is an allowable input for the function in question. When no confusion is likely, we may omit this reminder with the understanding that the appearance of $f(x)$ carries with it the implicit assumption that x is in the domain of f. On a related note, there is no reason to discuss functional limits at isolated points of the domain. Thus, functional limits will only be considered as x tends toward a limit point of the function's domain.

Example 4.2.2. (i) To familiarize ourselves with Definition 4.2.1, let's prove that if $f(x) = 3x + 1$, then

$$\lim_{x \to 2} f(x) = 7.$$

Let $\epsilon > 0$. Definition 4.2.1 requires that we produce a $\delta > 0$ so that $0 < |x - 2| < \delta$ leads to the conclusion $|f(x) - 7| < \epsilon$. Notice that

$$|f(x) - 7| = |(3x + 1) - 7| = |3x - 6| = 3|x - 2|.$$

Thus, if we choose $\delta = \epsilon/3$, then $0 < |x - 2| < \delta$ implies $|f(x) - 7| < 3\,(\epsilon/3) = \epsilon$.

(ii) Let's show

$$\lim_{x \to 2} g(x) = 4,$$

where $g(x) = x^2$. Given an arbitrary $\epsilon > 0$, our goal this time is to make $|g(x) - 4| < \epsilon$ by restricting $|x - 2|$ to be smaller than some carefully chosen δ. As in the previous problem, a little algebra reveals

$$|g(x) - 4| = |x^2 - 4| = |x + 2||x - 2|.$$

We can make $|x - 2|$ as small as we like, but we need an upper bound on $|x+2|$ in order to know how small to choose δ. The presence of the variable x causes some initial confusion, but keep in mind that we are discussing the limit as x approaches 2. If we agree that our δ-neighborhood around $c = 2$ must have radius no bigger than $\delta = 1$, then we get the upper bound $|x + 2| \le |3 + 2| = 5$ for all $x \in V_\delta(c)$.

Now, choose $\delta = \min\{1, \epsilon/5\}$. If $0 < |x - 2| < \delta$, then it follows that

$$|x^2 - 4| = |x + 2||x - 2| < (5)\frac{\epsilon}{5} = \epsilon,$$

and the limit is proved.

Sequential Criterion for Functional Limits

We worked very hard in Chapter 2 to derive an impressive list of properties enjoyed by sequential limits. In particular, the Algebraic Limit Theorem (Theorem 2.3.3) and the Order Limit Theorem (Theorem 2.3.4) proved invaluable in a large number of the arguments that followed. Not surprisingly, we are going to need analogous statements for functional limits. Although it is not difficult to generate independent proofs for these statements, all of them will follow quite naturally from their sequential analogs once we derive the sequential criterion for functional limits motivated in the opening discussion of this chapter.

Theorem 4.2.3 (Sequential Criterion for Functional Limits). *Given a function $f : A \to \mathbf{R}$ and a limit point c of A, the following two statements are equivalent:*

(i) $\lim_{x \to c} f(x) = L$.

(ii) *For all sequences $(x_n) \subseteq A$ satisfying $x_n \neq c$ and $(x_n) \to c$, it follows that $f(x_n) \to L$.*

Proof. ($\Rightarrow$) Let's first assume that $\lim_{x\to c} f(x) = L$. To prove (ii), we consider an arbitrary sequence (x_n), which converges to c and satisfies $x_n \neq c$. Our goal is to show that the image sequence $f(x_n)$ converges to L. This is most easily seen using the topological formulation of the definition.

Let $\epsilon > 0$. Because we are assuming (i), Definition 4.2.1B implies that there exists $V_\delta(c)$ with the property that all $x \in V_\delta(c)$ different from c satisfy $f(x) \in V_\epsilon(L)$. All we need to do then is argue that our particular sequence (x_n) is eventually in $V_\delta(c)$. But we are assuming that $(x_n) \to c$. This implies that there exists a point x_N after which $x_n \in V_\delta(c)$. It follows that $n \geq N$ implies $f(x_n) \in V_\epsilon(L)$, as desired.

($\Leftarrow$) For this implication we give a contrapositive proof, which is essentially a proof by contradiction. Thus, we assume that statement (ii) is true, and carefully negate statement (i). To say that

$$\lim_{x \to c} f(x) \neq L$$

means that there exists at least one particular $\epsilon_0 > 0$ for which no δ is a suitable response. In other words, no matter what $\delta > 0$ we try, there will always be at least one point

$$x \in V_\delta(c) \quad \text{with} \quad x \neq c \quad \text{for which} \quad f(x) \notin V_{\epsilon_0}(L).$$

Now consider $\delta_n = 1/n$. From the preceding discussion, it follows that for each $n \in \mathbf{N}$ we may pick an $x_n \in V_{\delta_n}(c)$ with $x_n \neq c$ and $f(x_n) \notin V_{\epsilon_0}(L)$. But now notice that the result of this is a sequence $(x_n) \to c$ with $x_n \neq c$, where the image sequence $f(x_n)$ certainly does *not* converge to L.

Because this contradicts (ii), which we are assuming is true for this argument, we may conclude that (i) must also hold. □

Theorem 4.2.3 has several useful corollaries. In addition to the previously advertised benefit of granting us some short proofs of statements about how functional limits interact with algebraic combinations of functions, we also get an economical way of establishing that certain limits do not exist.

Corollary 4.2.4 (Algebraic Limit Theorem for Functional Limits). *Let f and g be functions defined on a domain $A \subseteq \mathbf{R}$, and assume $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} g(x) = M$ for some limit point c of A. Then,*

(i) $\lim_{x\to c} kf(x) = kL$ *for all* $k \in \mathbf{R}$,

(ii) $\lim_{x\to c}[f(x) + g(x)] = L + M$,

(iii) $\lim_{x\to c}[f(x)g(x)] = LM$, *and*

(iv) $\lim_{x\to c} f(x)/g(x) = L/M$, *provided* $M \neq 0$.

Proof. These follow from Theorem 4.2.3 and the Algebraic Limit Theorem for sequences. The details are requested in Exercise 4.2.1. □

Corollary 4.2.5 (Divergence Criterion for Functional Limits). *Let f be a function defined on A, and let c be a limit point of A. If there exist two sequences (x_n) and (y_n) in A with $x_n \neq c$ and $y_n \neq c$ and*

$$\lim x_n = \lim y_n = c \quad \textit{but} \quad \lim f(x_n) \neq \lim f(y_n),$$

then we can conclude that the functional limit $\lim_{x\to c} f(x)$ does not exist.

Example 4.2.6. Assuming the familiar properties of the sine function, let's show that $\lim_{x\to 0} \sin(1/x)$ does not exist (Fig. 4.5).

If $x_n = 1/2n\pi$ and $y_n = 1/(2n\pi + \pi/2)$, then $\lim(x_n) = \lim(y_n) = 0$. However, $\sin(1/x_n) = 0$ for all $n \in \mathbf{N}$ while $\sin(1/y_n) = 1$. Thus,

$$\lim \sin(1/x_n) \neq \lim \sin(1/y_n),$$

so by Corollary 4.2.5, $\lim_{x\to 0} \sin(1/x)$ does not exist.

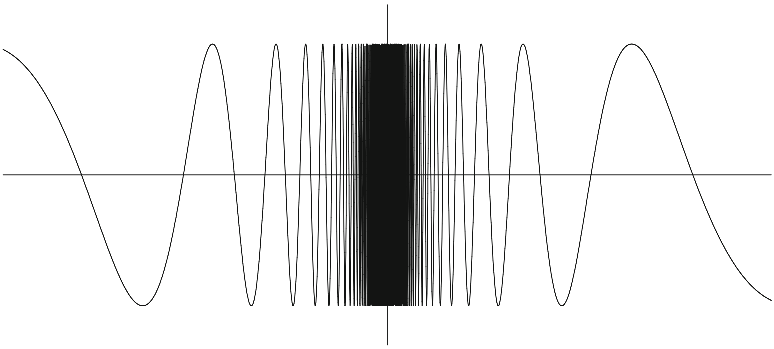


Figure 4.5: The function $\sin(1/x)$ near zero.

Exercises

Exercise 4.2.1. (a) Supply the details for how Corollary 4.2.4 part (ii) follows from the Sequential Criterion for Functional Limits in Theorem 4.2.3 and the Algebraic Limit Theorem for sequences proved in Chapter 2.

(b) Now, write another proof of Corollary 4.2.4 part (ii) directly from Definition 4.2.1 without using the sequential criterion in Theorem 4.2.3.

(c) Repeat (a) and (b) for Corollary 4.2.4 part (iii).

Exercise 4.2.2. For each stated limit, find the largest possible δ-neighborhood that is a proper response to the given ϵ challenge.

(a) $\lim_{x\to 3}(5x-6) = 9,$ where $\epsilon = 1$.

(b) $\lim_{x\to 4}\sqrt{x} = 2,$ where $\epsilon = 1$.

(c) $\lim_{x\to\pi}[[x]] = 3,$ where $\epsilon = 1$. (The function $[[x]]$ returns the greatest integer less than or equal to x.)

(d) $\lim_{x\to\pi}[[x]] = 3,$ where $\epsilon = .01$.

Exercise 4.2.3. Review the definition of Thomae's function $t(x)$ from Section 4.1.

(a) Construct three different sequences (x_n), (y_n), and (z_n), each of which converges to 1 without using the number 1 as a term in the sequence.

(b) Now, compute $\lim t(x_n)$, $\lim t(y_n)$, and $\lim t(z_n)$.

(c) Make an educated conjecture for $\lim_{x\to 1} t(x)$, and use Definition 4.2.1B to verify the claim. (Given $\epsilon > 0$, consider the set of points $\{x \in \mathbf{R} : t(x) \geq \epsilon\}$. Argue that all the points in this set are isolated.)

Exercise 4.2.4. Consider the reasonable but erroneous claim that

$$\lim_{x\to 10} 1/[[x]] = 1/10.$$

(a) Find the largest δ that represents a proper response to the challenge of $\epsilon = 1/2$.

(b) Find the largest δ that represents a proper response to $\epsilon = 1/50$.

(c) Find the largest ϵ challenge for which there is no suitable δ response possible.

Exercise 4.2.5. Use Definition 4.2.1 to supply a proper proof for the following limit statements.

(a) $\lim_{x\to 2}(3x+4) = 10$.

(b) $\lim_{x\to 0} x^3 = 0$.

(c) $\lim_{x\to 2}(x^2+x-1) = 5$.

(d) $\lim_{x\to 3} 1/x = 1/3$.

Exercise 4.2.6. Decide if the following claims are true or false, and give short justifications for each conclusion.

(a) If a particular δ has been constructed as a suitable response to a particular ϵ challenge, then any smaller positive δ will also suffice.

(b) If $\lim_{x\to a} f(x) = L$ and a happens to be in the domain of f, then $L = f(a)$.

(c) If $\lim_{x\to a} f(x) = L$, then $\lim_{x\to a} 3[f(x)-2]^2 = 3(L-2)^2$.

(d) If $\lim_{x\to a} f(x) = 0$, then $\lim_{x\to a} f(x)g(x) = 0$ for any function g (with domain equal to the domain of f.)

Exercise 4.2.7. Let $g : A \to \mathbf{R}$ and assume that f is a bounded function on A in the sense that there exists $M > 0$ satisfying $|f(x)| \le M$ for all $x \in A$.

Show that if $\lim_{x\to c} g(x) = 0$, then $\lim_{x\to c} g(x)f(x) = 0$ as well.

Exercise 4.2.8. Compute each limit or state that it does not exist. Use the tools developed in this section to justify each conclusion.

(a) $\lim_{x\to 2} \frac{|x-2|}{x-2}$

(b) $\lim_{x\to 7/4} \frac{|x-2|}{x-2}$

(c) $\lim_{x\to 0}(-1)^{[[1/x]]}$

(d) $\lim_{x\to 0} \sqrt[3]{x}(-1)^{[[1/x]]}$

Exercise 4.2.9 (Infinite Limits). The statement $\lim_{x\to 0} 1/x^2 = \infty$ certainly makes intuitive sense. To construct a rigorous definition in the challenge–response style of Definition 4.2.1 for an infinite limit statement of this form, we replace the (arbitrarily small) $\epsilon > 0$ challenge with an (arbitrarily large) $M > 0$ challenge:

Definition: $\lim_{x\to c} f(x) = \infty$ means that for all $M > 0$ we can find a $\delta > 0$ such that whenever $0 < |x-c| < \delta$, it follows that $f(x) > M$.

(a) Show $\lim_{x\to 0} 1/x^2 = \infty$ in the sense described in the previous definition.

(b) Now, construct a definition for the statement $\lim_{x\to\infty} f(x) = L$. Show $\lim_{x\to\infty} 1/x = 0$.

(c) What would a rigorous definition for $\lim_{x\to\infty} f(x) = \infty$ look like? Give an example of such a limit.

Exercise 4.2.10 (Right and Left Limits). Introductory calculus courses typically refer to the *right-hand limit* of a function as the limit obtained by "letting x approach a from the right-hand side."

(a) Give a proper definition in the style of Definition 4.2.1 for the right-hand and left-hand limit statements:

$$\lim_{x\to a^+} f(x) = L \quad \text{and} \quad \lim_{x\to a^-} f(x) = M.$$

(b) Prove that $\lim_{x\to a} f(x) = L$ if and only if both the right and left-hand limits equal L.

Exercise 4.2.11 (Squeeze Theorem). Let f, g, and h satisfy $f(x) \le g(x) \le h(x)$ for all x in some common domain A. If $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} h(x) = L$ at some limit point c of A, show $\lim_{x\to c} g(x) = L$ as well.

4.3 Continuous Functions

We now come to a significant milestone in our progress toward a rigorous theory of real-valued functions—a proper definition of the seminal concept of continuity that avoids any intuitive appeals to "unbroken curves" or functions without "jumps" or "holes."

Definition 4.3.1 (Continuity). A function $f : A \to \mathbf{R}$ is *continuous at a point* $c \in A$ if, for all $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - c| < \delta$ (and $x \in A$) it follows that $|f(x) - f(c)| < \epsilon$.

If f is continuous at every point in the domain A, then we say that f is *continuous on* A.

The definition of continuity looks much like the definition for functional limits, with a few subtle differences. The most important is that we require the point c to be in the domain of f. The value $f(c)$ then becomes the value of $\lim_{x\to c} f(x)$. With this observation in mind, it is tempting to shorten Definition 4.3.1 to say that f is continuous at $c \in A$ if

$$\lim_{x\to c} f(x) = f(c).$$

This is fine as long as c is a limit point of A. If c is an isolated point of A, then $\lim_{x\to c} f(x)$ isn't defined but Definition 4.3.1 can still be applied. An unremarkable but noteworthy consequence of this definition is that functions are continuous at isolated points of their domains (Exercise 4.3.5).

We saw in the previous section that, in addition to the standard ϵ–δ definition, functional limits have a useful formulation in terms of sequences. The same is true of continuity. The next theorem summarizes these various equivalent ways to characterize the continuity of a function at a given point.

Theorem 4.3.2 (Characterizations of Continuity). *Let $f : A \to \mathbf{R}$, and let $c \in A$. The function f is continuous at c if and only if any one of the following three conditions is met:*

(i) *For all $\epsilon > 0$, there exists a $\delta > 0$ such that $|x - c| < \delta$ (and $x \in A$) implies $|f(x) - f(c)| < \epsilon$;*

(ii) *For all $V_\epsilon(f(c))$, there exists a $V_\delta(c)$ with the property that $x \in V_\delta(c)$ (and $x \in A$) implies $f(x) \in V_\epsilon(f(c))$;*

(iii) *For all $(x_n) \to c$ (with $x_n \in A$), it follows that $f(x_n) \to f(c)$.*

If c is a limit point of A, then the above conditions are equivalent to

(iv) $\lim_{x \to c} f(x) = f(c)$.

Proof. Statement (i) is just Definition 4.3.1, and statement (ii) is the standard rewording of (i) using topological neighborhoods in place of the absolute value notation. Statement (iii) is equivalent to (i) via an argument nearly identical to that of Theorem 4.2.3, with some slight modifications for when $x_n = c$. Finally, statement (iv) is seen to be equivalent to (i) by considering Definition 4.2.1 and observing that the case $x = c$ (which is excluded in the definition of functional limits) leads to the requirement $f(c) \in V_\epsilon(f(c))$, which is trivially true. □

The length of this list is somewhat deceiving. Statements (i), (ii), and (iv) are closely related and essentially remind us that functional limits have an ϵ–δ formulation as well as a topological description. Statement (iii), however, is qualitatively different from the others. As a general rule, the sequential characterization of continuity is typically the most useful for demonstrating that a function is *not* continuous at some point.

Corollary 4.3.3 (Criterion for Discontinuity). *Let $f : A \to \mathbf{R}$, and let $c \in A$ be a limit point of A. If there exists a sequence $(x_n) \subseteq A$ where $(x_n) \to c$ but such that $f(x_n)$ does not converge to $f(c)$, we may conclude that f is not continuous at c.*

The sequential characterization of continuity is also important for the other reasons that it was important for functional limits. In particular, it allows us to bring our catalog of results about the behavior of sequences to bear on the study of continuous functions. The next theorem should be compared to Corollary 4.2.3 as well as to Theorem 2.3.3.

Theorem 4.3.4 (Algebraic Continuity Theorem). *Assume $f : A \to \mathbf{R}$ and $g : A \to \mathbf{R}$ are continuous at a point $c \in A$. Then,*

(i) *$kf(x)$ is continuous at c for all $k \in \mathbf{R}$;*

(ii) *$f(x) + g(x)$ is continuous at c;*

(iii) *$f(x)g(x)$ is continuous at c; and*

(iv) *$f(x)/g(x)$ is continuous at c, provided the quotient is defined.*

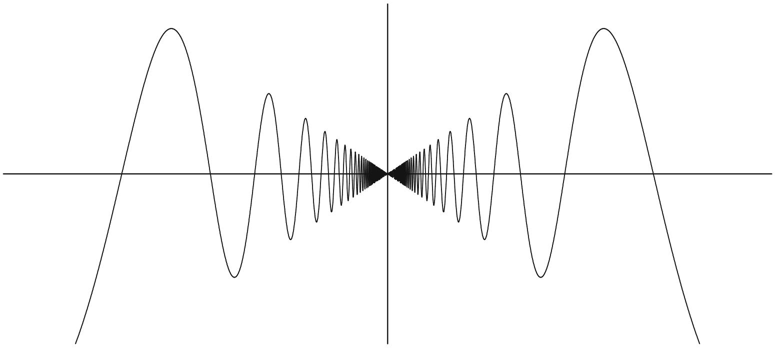


Figure 4.6: The function $x\sin(1/x)$ near zero.

Proof. All of these statements can be quickly derived from Corollary 4.2.4 and Theorem 4.3.2. □

These results provide us with the tools we need to firm up our arguments in the opening section of this chapter about the behavior of Dirichlet's function and Thomae's function. The details are requested in Exercise 4.3.7. Here are some more examples of arguments for and against continuity of some familiar functions.

Example 4.3.5. All polynomials are continuous on $\mathbf{R}$. In fact, rational functions (i.e., quotients of polynomials) are continuous wherever they are defined.

To see why this is so, consider the identity function $g(x) = x$. Because $|g(x) - g(c)| = |x - c|$, we can respond to a given $\epsilon > 0$ by choosing $\delta = \epsilon$, and it follows that g is continuous on all of $\mathbf{R}$. It is even simpler to show that a constant function $f(x) = k$, is continuous. (Letting $\delta = 1$ regardless of the value of ϵ does the trick.) Because an arbitrary polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

consists of sums and products of $g(x)$ with different constant functions, we may conclude from Theorem 4.3.4 that $p(x)$ is continuous.

Likewise, Theorem 4.3.4 implies that quotients of polynomials are continuous as long as the denominator is not zero.

Example 4.3.6. In Example 4.2.6, we saw that the oscillations of $\sin(1/x)$ are so rapid near the origin that $\lim_{x\to 0}\sin(1/x)$ does not exist. Now, consider the function

$$g(x) = \begin{cases} x\sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

To investigate the continuity of g at $c = 0$ (Fig. 4.6), we can estimate

$$|g(x) - g(0)| = |x\sin(1/x) - 0| \leq |x|.$$

Given $\epsilon > 0$, set $\delta = \epsilon$, so that whenever $|x - 0| = |x| < \delta$ it follows that $|g(x) - g(0)| < \epsilon$. Thus, g is continuous at the origin.

Example 4.3.7. Throughout the exercises we have been using the greatest integer function $h(x) = [[x]]$ which for each $x \in \mathbf{R}$ returns the largest integer $n \in \mathbf{Z}$ satisfying $n \leq x$. This familiar step function certainly has discontinuous "jumps" at each integer value of its domain, but it is a useful exercise to try and articulate this observation in the language of analysis.

Given $m \in \mathbf{Z}$, define the sequence (x_n) by $x_n = m - 1/n$. It follows that $(x_n) \to m$, but

$$h(x_n) \to (m-1),$$

which does not equal $m = h(m)$. By Corollary 4.3.3, we see that h fails to be continuous at each $m \in \mathbf{Z}$.

Now let's see why h is continuous at a point $c \notin \mathbf{Z}$. Given $\epsilon > 0$, we must find a δ-neighborhood $V_\delta(c)$ such that $x \in V_\delta(c)$ implies $h(x) \in V_\epsilon(h(c))$. We know that $c \in \mathbf{R}$ falls between consecutive integers $n < c < n+1$ for some $n \in \mathbf{Z}$. If we take $\delta = \min\{c - n, (n+1) - c\}$, then it follows from the definition of h that $h(x) = h(c)$ for all $x \in V_\delta(c)$. Thus, we certainly have

$$h(x) \in V_\epsilon(h(c))$$

whenever $x \in V_\delta(c)$.

This latter proof is quite different from the typical situation in that the value of δ does not actually depend on the choice of ϵ. Usually, a smaller ϵ requires a smaller δ in response, but here the same value of δ works no matter how small ϵ is chosen.

Example 4.3.8. Consider $f(x) = \sqrt{x}$ defined on $A = \{x \in \mathbf{R} : x \geq 0\}$. Exercise 2.3.1 outlines a sequential proof that f is continuous on A. Here, we give an ϵ–δ proof of the same fact.

Let $\epsilon > 0$. We need to argue that $|f(x) - f(c)|$ can be made less than ϵ for all values of x in some δ neighborhood around c. If $c = 0$, this reduces to the statement $\sqrt{x} < \epsilon$, which happens as long as $x < \epsilon^2$. Thus, if we choose $\delta = \epsilon^2$, we see that $|x - 0| < \delta$ implies $|f(x) - 0| < \epsilon$.

For a point $c \in A$ different from zero, we need to estimate $|\sqrt{x} - \sqrt{c}|$. This time, write

$$|\sqrt{x} - \sqrt{c}| = |\sqrt{x} - \sqrt{c}| \left(\frac{\sqrt{x} + \sqrt{c}}{\sqrt{x} + \sqrt{c}} \right) = \frac{|x - c|}{\sqrt{x} + \sqrt{c}} \leq \frac{|x - c|}{\sqrt{c}}.$$

In order to make this quantity less than ϵ, it suffices to pick $\delta = \epsilon\sqrt{c}$. Then, $|x - c| < \delta$ implies

$$|\sqrt{x} - \sqrt{c}| < \frac{\epsilon\sqrt{c}}{\sqrt{c}} = \epsilon,$$

as desired.

Although we have now shown that both polynomials and the square root function are continuous, the Algebraic Continuity Theorem does not provide the justification needed to conclude that a function such as $h(x) = \sqrt{3x^2 + 5}$ is continuous. For this, we must prove that *compositions* of continuous functions are continuous.

Theorem 4.3.9 (Composition of Continuous Functions). *Given $f : A \to \mathbf{R}$ and $g : B \to \mathbf{R}$, assume that the range $f(A) = \{f(x) : x \in A\}$ is contained in the domain B so that the composition $g \circ f(x) = g(f(x))$ is defined on A.*

If f is continuous at $c \in A$, and if g is continuous at $f(c) \in B$, then $g \circ f$ is continuous at c.

Proof. Exercise 4.3.3. □

Exercises

Exercise 4.3.1. Let $g(x) = \sqrt[3]{x}$.

(a) Prove that g is continuous at $c = 0$.

(b) Prove that g is continuous at a point $c \neq 0$. (The identity $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$ will be helpful.)

Exercise 4.3.2. To gain a deeper understanding of the relationship between ϵ and δ in the definition of continuity, let's explore some modest variations of Definition 4.3.1. In all of these, let f be a function defined on all of $\mathbf{R}$.

(a) Let's say f is *onetinuous* at c if for all $\epsilon > 0$ we can choose $\delta = 1$ and it follows that $|f(x) - f(c)| < \epsilon$ whenever $|x - c| < \delta$. Find an example of a function that is onetinuous on all of $\mathbf{R}$.

(b) Let's say f is *equaltinuous* at c if for all $\epsilon > 0$ we can choose $\delta = \epsilon$ and it follows that $|f(x) - f(c)| < \epsilon$ whenever $|x - c| < \delta$. Find an example of a function that is equaltinuous on $\mathbf{R}$ that is nowhere onetinuous, or explain why there is no such function.

(c) Let's say f is *lesstinuous* at c if for all $\epsilon > 0$ we can choose $0 < \delta < \epsilon$ and it follows that $|f(x) - f(c)| < \epsilon$ whenever $|x - c| < \delta$. Find an example of a function that is lesstinuous on $\mathbf{R}$ that is nowhere equaltinuous, or explain why there is no such function.

(d) Is every lesstinuous function continuous? Is every continuous function lesstinuous? Explain.

Exercise 4.3.3. (a) Supply a proof for Theorem 4.3.9 using the ϵ–δ characterization of continuity.

(b) Give another proof of this theorem using the sequential characterization of continuity (from Theorem 4.3.2 (iii)).

Exercise 4.3.4. Assume f and g are defined on all of $\mathbf{R}$ and that $\lim_{x\to p} f(x) = q$ and $\lim_{x\to q} g(x) = r$.

(a) Give an example to show that it may not be true that

$$\lim_{x\to p} g(f(x)) = r.$$

(b) Show that the result in (a) does follow if we assume f and g are continuous.

(c) Does the result in (a) hold if we only assume f is continuous? How about if we only assume that g is continuous?

Exercise 4.3.5. Show using Definition 4.3.1 that if c is an isolated point of $A \subseteq \mathbf{R}$, then $f : A \to \mathbf{R}$ is continuous at c.

Exercise 4.3.6. Provide an example of each or explain why the request is impossible.

(a) Two functions f and g, neither of which is continuous at 0 but such that $f(x)g(x)$ and $f(x) + g(x)$ are continuous at 0.

(b) A function $f(x)$ continuous at 0 and $g(x)$ not continuous at 0 such that $f(x) + g(x)$ is continuous at 0.

(c) A function $f(x)$ continuous at 0 and $g(x)$ not continuous at 0 such that $f(x)g(x)$ is continuous at 0.

(d) A function $f(x)$ not continuous at 0 such that $f(x) + \frac{1}{f(x)}$ is continuous at 0.

(e) A function $f(x)$ not continuous at 0 such that $[f(x)]^3$ is continuous at 0.

Exercise 4.3.7. (a) Referring to the proper theorems, give a formal argument that Dirichlet's function from Section 4.1 is nowhere-continuous on $\mathbf{R}$.

(b) Review the definition of Thomae's function in Section 4.1 and demonstrate that it fails to be continuous at every rational point.

(c) Use the characterization of continuity in Theorem 4.3.2 (iii) to show that Thomae's function is continuous at every irrational point in $\mathbf{R}$. (Given $\epsilon > 0$, consider the set of points $\{x \in \mathbf{R} : t(x) \geq \epsilon\}$.)

Exercise 4.3.8. Decide if the following claims are true or false, providing either a short proof or counterexample to justify each conclusion. Assume throughout that g is defined and continuous on all of $\mathbf{R}$.

(a) If $g(x) \geq 0$ for all $x < 1$, then $g(1) \geq 0$ as well.

(b) If $g(r) = 0$ for all $r \in \mathbf{Q}$, then $g(x) = 0$ for all $x \in \mathbf{R}$.

(c) If $g(x_0) > 0$ for a single point $x_0 \in \mathbf{R}$, then $g(x)$ is in fact strictly positive for uncountably many points.

Exercise 4.3.9. Assume $h : \mathbf{R} \to \mathbf{R}$ is continuous on $\mathbf{R}$ and let $K = \{x : h(x) = 0\}$. Show that K is a closed set.

Exercise 4.3.10. Observe that if a and b are real numbers, then

$$\max\{a, b\} = \frac{1}{2}[(a + b) + |a - b|].$$

(a) Show that if $f_1, f_2, \ldots, f_n$ are continuous functions, then

$$g(x) = \max\{f_1(x), f_2(x), \ldots, f_n(x)\}$$

is a continuous function.

(b) Let's explore whether the result in (a) extends to the infinite case. For each $n \in \mathbf{N}$, define f_n on $\mathbf{R}$ by

$$f_n(x) = \begin{cases} 1 & \text{if } |x| \geq 1/n \\ n|x| & \text{if } |x| < 1/n. \end{cases}$$

Now explicitly compute $h(x) = \sup\{f_1(x), f_2(x), f_3(x), \ldots\}$.

Exercise 4.3.11 (Contraction Mapping Theorem). Let f be a function defined on all of $\mathbf{R}$, and assume there is a constant c such that $0 < c < 1$ and

$$|f(x) - f(y)| \leq c|x - y|$$

for all $x, y \in \mathbf{R}$.

(a) Show that f is continuous on $\mathbf{R}$.

(b) Pick some point $y_1 \in \mathbf{R}$ and construct the sequence

$$(y_1, f(y_1), f(f(y_1)), \ldots).$$

In general, if $y_{n+1} = f(y_n)$, show that the resulting sequence (y_n) is a Cauchy sequence. Hence we may let $y = \lim y_n$.

(c) Prove that y is a fixed point of f (i.e., $f(y) = y$) and that it is unique in this regard.

(d) Finally, prove that if x is *any* arbitrary point in $\mathbf{R}$, then the sequence $(x, f(x), f(f(x)), \ldots)$ converges to y defined in (b).

Exercise 4.3.12. Let $F \subseteq \mathbf{R}$ be a nonempty closed set and define $g(x) = \inf\{|x - a| : a \in F\}$. Show that g is continuous on all of $\mathbf{R}$ and $g(x) \neq 0$ for all $x \notin F$.

Exercise 4.3.13. Let f be a function defined on all of $\mathbf{R}$ that satisfies the additive condition $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbf{R}$.

(a) Show that $f(0) = 0$ and that $f(-x) = -f(x)$ for all $x \in \mathbf{R}$.

(b) Let $k = f(1)$. Show that $f(n) = kn$ for all $n \in \mathbf{N}$, and then prove that $f(z) = kz$ for all $z \in \mathbf{Z}$. Now, prove that $f(r) = kr$ for any rational number r.

(c) Show that if f is continuous at $x = 0$, then f is continuous at every point in $\mathbf{R}$ and conclude that $f(x) = kx$ for all $x \in \mathbf{R}$. Thus, any additive function that is continuous at $x = 0$ must necessarily be a linear function through the origin.

Exercise 4.3.14. (a) Let F be a closed set. Construct a function $f : \mathbf{R} \to \mathbf{R}$ such that the set of points where f fails to be continuous is precisely F. (The concept of the interior of a set, discussed in Exercise 3.2.14, may be useful.)

(b) Now consider an open set O. Construct a function $g : \mathbf{R} \to \mathbf{R}$ whose set of discontinuous points is precisely O. (For this problem, the function in Exercise 4.3.12 may be useful.)

4.4 Continuous Functions on Compact Sets

Given a function $f : A \to \mathbf{R}$ and a subset $B \subseteq A$, the notation $f(B)$ refers to the range of f over the set B; that is,

$$f(B) = \{f(x) : x \in B\}.$$

The adjectives open, closed, bounded, compact, perfect, and connected are all used to describe subsets of the real line. An interesting question is to sort out which, if any, of these properties are preserved when a particular set B is mapped to $f(B)$ via a continuous function. For instance, if B is open and f is continuous, is $f(B)$ necessarily open? The answer to this question is no. If $f(x) = x^2$ and B is the open interval $(-1, 1)$, then $f(B)$ is the interval $[0, 1)$, which is not open.

The corresponding conjecture for closed sets also turns out to be false, although constructing a counterexample requires a little more thought. Consider the function

$$g(x) = \frac{1}{1 + x^2}$$

and the closed set $B = [0, \infty) = \{x : x \geq 0\}$. Because $g(B) = (0, 1]$ is not closed, we must conclude that continuous functions do not, in general, map closed sets to closed sets. Notice, however, that our particular counterexample required using an *unbounded* closed set B. This is not incidental. Sets that are closed and bounded—that is, compact sets—always get mapped to closed and bounded subsets by continuous functions.

Theorem 4.4.1 (Preservation of Compact Sets). *Let $f : A \to \mathbf{R}$ be continuous on A. If $K \subseteq A$ is compact, then $f(K)$ is compact as well.*

Proof. Let (y_n) be an arbitrary sequence contained in the range set $f(K)$. To prove this result, we must find a subsequence (y_{n_k}), which converges to a limit also in $f(K)$. The strategy is to take advantage of the assumption that the domain set K is compact by translating the sequence (y_n)—which is in the range of f—back to a sequence in the domain K.

To assert that $(y_n) \subseteq f(K)$ means that, for each $n \in \mathbf{N}$, we can find (at least one) $x_n \in K$ with $f(x_n) = y_n$. This yields a sequence $(x_n) \subseteq K$. Because K is compact, there exists a convergent subsequence (x_{n_k}) whose limit $x = \lim x_{n_k}$ is also in K. Finally, we make use of the fact that f is assumed to be continuous on A and so is continuous at x in particular. Given that $(x_{n_k}) \to x$, we conclude that $(y_{n_k}) \to f(x)$. Because $x \in K$, we have that $f(x) \in f(K)$, and hence $f(K)$ is compact. □

An extremely important corollary is obtained by combining this result with the observation that compact sets are bounded and contain their supremums and infimums.

Theorem 4.4.2 (Extreme Value Theorem). *If $f : K \to \mathbf{R}$ is continuous on a compact set $K \subseteq \mathbf{R}$, then f attains a maximum and minimum value. In other words, there exist $x_0, x_1 \in K$ such that $f(x_0) \le f(x) \le f(x_1)$ for all $x \in K$.*

Proof. Because $f(K)$ is compact, we can set $\alpha = \sup f(K)$ and know $\alpha \in f(K)$ (Exercise 3.3.1). It follows that there exist $x_1 \in K$ with $\alpha = f(x_1)$. The argument for the minimum value is similar. □

Uniform Continuity

Although we have proved that polynomials are always continuous on $\mathbf{R}$, there is an important lesson to be learned by constructing direct proofs that the functions $f(x) = 3x + 1$ and $g(x) = x^2$ (previously studied in Example 4.2.2) are everywhere continuous.

Example 4.4.3. (i) To show directly that $f(x) = 3x + 1$ is continuous at an arbitrary point $c \in \mathbf{R}$, we must argue that $|f(x) - f(c)|$ can be made arbitrarily small for values of x near c. Now,

$$|f(x) - f(c)| = |(3x + 1) - (3c + 1)| = 3|x - c|,$$

so, given $\epsilon > 0$, we choose $\delta = \epsilon/3$. Then, $|x - c| < \delta$ implies

$$|f(x) - f(c)| = 3|x - c| < 3\left(\frac{\epsilon}{3}\right) = \epsilon.$$

Of particular importance for this discussion is the fact that the choice of δ is the same regardless of which point $c \in \mathbf{R}$ we are considering.

(ii) Let's contrast this with what happens when we prove $g(x) = x^2$ is continuous on $\mathbf{R}$. Given $c \in \mathbf{R}$, we have

$$|g(x) - g(c)| = |x^2 - c^2| = |x - c||x + c|.$$

As discussed in Example 4.2.2, we need an upper bound on $|x+c|$, which is obtained by insisting that our choice of δ not exceed 1. This guarantees that all values of x under consideration will necessarily fall in the interval $(c-1, c+1)$. It follows that

$$|x+c| \leq |x|+|c| \leq (|c|+1)+|c| = 2|c|+1.$$

Now, let $\epsilon > 0$. If we choose $\delta = \min\{1, \epsilon/(2|c|+1)\}$, then $|x-c| < \delta$ implies

$$|f(x)-f(c)| = |x-c||x+c| < \left(\frac{\epsilon}{2|c|+1}\right)(2|c|+1) = \epsilon.$$

Now, there is nothing deficient about this argument, but it is important to notice that, in the second proof, the algorithm for choosing the response δ depends on the value of c. The statement

$$\delta = \frac{\epsilon}{2|c|+1}$$

means that larger values of c are going to require smaller values of δ, a fact that should be evident from a consideration of the graph of $g(x) = x^2$ (Fig. 4.7). Given, say, $\epsilon = 1$, a response of $\delta = 1/3$ is sufficient for $c = 1$ because $2/3 < x < 4/3$ certainly implies $0 < x^2 < 2$. However, if $c = 10$, then the steepness of the graph of $g(x)$ means that a much smaller δ is required—$\delta = 1/21$ by our rule—to force $99 < x^2 < 101$.

The next definition is meant to distinguish between these two examples.

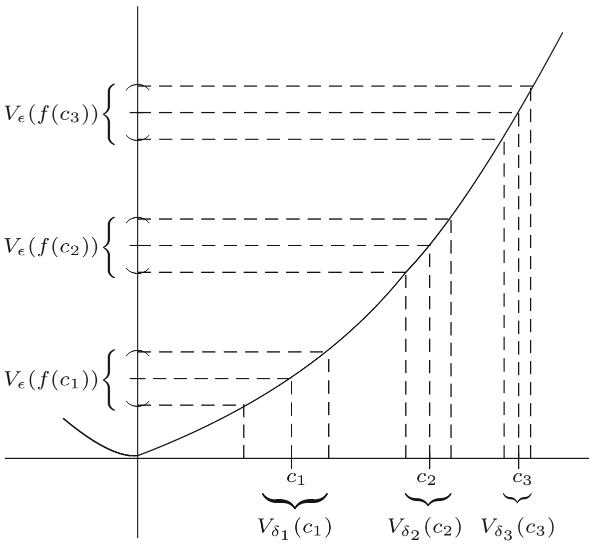


Figure 4.7: $g(x) = x^2$; A LARGER c REQUIRES A SMALLER δ.

Definition 4.4.4 (Uniform Continuity). A function $f : A \to \mathbf{R}$ is *uniformly continuous on* A if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x, y \in A$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.

Recall that to say that "f is continuous on A" means that f is continuous at each individual point $c \in A$. In other words, given $\epsilon > 0$ and $c \in A$, we can find a $\delta > 0$ *perhaps depending on* c such that if $|x - c| < \delta$, then $|f(x) - f(c)| < \epsilon$. Uniform continuity is a strictly stronger property. The key distinction between asserting that f is "uniformly continuous on A" versus simply "continuous on A" is that, given an $\epsilon > 0$, a single $\delta > 0$ can be chosen that works simultaneously for all points c in A. To say that a function is *not* uniformly continuous on a set A, then, does not necessarily mean it is not continuous at some point. Rather, it means that there is some $\epsilon_0 > 0$ for which no single $\delta > 0$ is a suitable response for all $c \in A$.

Theorem 4.4.5 (Sequential Criterion for Absence of Uniform Continuity). *A function* $f : A \to \mathbf{R}$ *fails to be uniformly continuous on* A *if and only if there exists a particular* $\epsilon_0 > 0$ *and two sequences* (x_n) *and* (y_n) *in* A *satisfying*

$$|x_n - y_n| \to 0 \quad \textit{but} \quad |f(x_n) - f(y_n)| \geq \epsilon_0.$$

Proof. The negation of Definition 4.4.4 states that f is not uniformly continuous on A if and only if there exists $\epsilon_0 > 0$ such that for all $\delta > 0$ we can find two points x and y satisfying $|x - y| < \delta$ but with $|f(x) - f(y)| \geq \epsilon_0$. Thus, if we set $\delta_1 = 1$, then there exist two points x_1 and y_1 where $|x_1 - y_1| < 1$ but $|f(x_1) - f(y_1)| \geq \epsilon_0$.

In a similar way, if we set $\delta_n = 1/n$ where $n \in \mathbf{N}$, it follows that there exist points x_n and y_n with $|x_n - y_n| < 1/n$ but where $|f(x_n) - f(y_n)| \geq \epsilon_0$. The resulting sequences (x_n) and (y_n) satisfy the requirements described in the theorem.

Conversely, if ϵ_0, (x_n) and (y_n) exist as described, it is straightforward to see that no $\delta > 0$ is a suitable response for ϵ_0. □

Example 4.4.6. The function $h(x) = \sin(1/x)$ (Fig. 4.5) is continuous at every point in the open interval $(0, 1)$ but is not uniformly continuous on this interval. The problem arises near zero, where the increasingly rapid oscillations take domain values that are quite close together to range values a distance 2 apart. To illustrate Theorem 4.4.5, take $\epsilon_0 = 2$ and set

$$x_n = \frac{1}{\pi/2 + 2n\pi} \quad \text{and} \quad y_n = \frac{1}{3\pi/2 + 2n\pi}.$$

Because each of these sequences tends to zero, we have $|x_n - y_n| \to 0$, and a short calculation reveals $|h(x_n) - h(y_n)| = 2$ for all $n \in \mathbf{N}$.

Whereas continuity is defined at a single point, uniform continuity is always discussed in reference to a particular domain. In Example 4.4.3, we were not able to prove that $g(x) = x^2$ is uniformly continuous on $\mathbf{R}$ because larger values of x require smaller and smaller values of δ. (As another illustration of Theorem 4.4.5, take $x_n = n$ and $y_n = n + 1/n$.) It is true, however, that $g(x)$ *is* uniformly continuous on the bounded set $[-10, 10]$. Returning to the argument set forth in Example 4.4.3 (ii), notice that if we restrict our attention to the domain $[-10, 10]$, then $|x + y| \leq 20$ for all x and y. Given $\epsilon > 0$, we can now choose $\delta = \epsilon/20$, and verify that if $x, y \in [-10, 10]$ satisfy $|x - y| < \delta$, then

$$|f(x) - f(y)| = |x^2 - y^2| = |x - y||x + y| < \left(\frac{\epsilon}{20}\right) 20 = \epsilon.$$

In fact, it is not difficult to see how to modify this argument to show that $g(x)$ is uniformly continuous on any bounded set A in $\mathbf{R}$.

Now, Example 4.4.6 is included to keep us from jumping to the erroneous conclusion that functions that are continuous on bounded domains are necessarily uniformly continuous. A general result does follow, however, if we assume that the domain is compact.

Theorem 4.4.7 (Uniform Continuity on Compact Sets). *A function that is continuous on a compact set K is uniformly continuous on K.*

Proof. Assume $f : K \to \mathbf{R}$ is continuous at every point of a compact set $K \subseteq \mathbf{R}$. To prove that f is uniformly continuous on K we argue by contradiction.

By the criterion in Theorem 4.4.5, if f is not uniformly continuous on K, then there exist two sequences (x_n) and (y_n) in K such that

$$\lim |x_n - y_n| = 0 \quad \text{while} \quad |f(x_n) - f(y_n)| \geq \epsilon_0$$

for some particular $\epsilon_0 > 0$. Because K is compact, the sequence (x_n) has a convergent subsequence (x_{n_k}) with $x = \lim x_{n_k}$ also in K.

We could use the compactness of K again to produce a convergent subsequence of (y_n), but notice what happens when we consider the particular subsequence (y_{n_k}) consisting of those terms in (y_n) that correspond to the terms in the convergent subsequence (x_{n_k}). By the Algebraic Limit Theorem,

$$\lim(y_{n_k}) = \lim((y_{n_k} - x_{n_k}) + x_{n_k}) = 0 + x.$$

The conclusion is that both (x_{n_k}) and (y_{n_k}) converge to $x \in K$. Because f is assumed to be continuous at x, we have $\lim f(x_{n_k}) = f(x)$ and $\lim f(y_{n_k}) = f(x)$, which implies

$$\lim(f(x_{n_k}) - f(y_{n_k})) = 0.$$

A contradiction arises when we recall that (x_n) and (y_n) were chosen to satisfy

$$|f(x_n) - f(y_n)| \geq \epsilon_0$$

for all $n \in \mathbf{N}$. We conclude, then, that f is indeed uniformly continuous on K. □

Exercises

Exercise 4.4.1. (a) Show that $f(x) = x^3$ is continuous on all of $\mathbf{R}$.

(b) Argue, using Theorem 4.4.5, that f is not uniformly continuous on $\mathbf{R}$.

(c) Show that f is uniformly continuous on any bounded subset of $\mathbf{R}$.

Exercise 4.4.2. (a) Is $f(x) = 1/x$ uniformly continuous on $(0, 1)$?

(b) Is $g(x) = \sqrt{x^2 + 1}$ uniformly continuous on $(0, 1)$?

(c) Is $h(x) = x \sin(1/x)$ uniformly continuous on $(0, 1)$?

Exercise 4.4.3. Show that $f(x) = 1/x^2$ is uniformly continuous on the set $[1, \infty)$ but not on the set $(0, 1]$.

Exercise 4.4.4. Decide whether each of the following statements is true or false, justifying each conclusion.

(a) If f is continuous on $[a, b]$ with $f(x) > 0$ for all $a \leq x \leq b$, then $1/f$ is bounded on $[a, b]$ (meaning $1/f$ has bounded range).

(b) If f is uniformly continuous on a bounded set A, then $f(A)$ is bounded.

(c) If f is defined on $\mathbf{R}$ and $f(K)$ is compact whenever K is compact, then f is continuous on $\mathbf{R}$.

Exercise 4.4.5. Assume that g is defined on an open interval (a, c) and it is known to be uniformly continuous on $(a, b]$ and $[b, c)$, where $a < b < c$. Prove that g is uniformly continuous on (a, c).

Exercise 4.4.6. Give an example of each of the following, or state that such a request is impossible. For any that are impossible, supply a short explanation for why this is the case.

(a) A continuous function $f : (0, 1) \to \mathbf{R}$ and a Cauchy sequence (x_n) such that $f(x_n)$ is not a Cauchy sequence;

(b) A uniformly continuous function $f : (0, 1) \to \mathbf{R}$ and a Cauchy sequence (x_n) such that $f(x_n)$ is not a Cauchy sequence;

(c) A continuous function $f : [0, \infty) \to \mathbf{R}$ and a Cauchy sequence (x_n) such that $f(x_n)$ is not a Cauchy sequence;

Exercise 4.4.7. Prove that $f(x) = \sqrt{x}$ is uniformly continuous on $[0, \infty)$.

Exercise 4.4.8. Give an example of each of the following, or provide a short argument for why the request is impossible.

(a) A continuous function defined on $[0, 1]$ with range $(0, 1)$.

(b) A continuous function defined on $(0, 1)$ with range $[0, 1]$.

(c) A continuous function defined on $(0, 1]$ with range $(0, 1)$.

Exercise 4.4.9 (Lipschitz Functions). A function $f : A \to \mathbf{R}$ is called *Lipschitz* if there exists a bound $M > 0$ such that

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq M$$

for all $x \neq y \in A$. Geometrically speaking, a function f is Lipschitz if there is a uniform bound on the magnitude of the slopes of lines drawn through any two points on the graph of f.

(a) Show that if $f : A \to \mathbf{R}$ is Lipschitz, then it is uniformly continuous on A.

(b) Is the converse statement true? Are all uniformly continuous functions necessarily Lipschitz?

Exercise 4.4.10. Assume that f and g are uniformly continuous functions defined on a common domain A. Which of the following combinations are necessarily uniformly continuous on A:

$$f(x) + g(x), \quad f(x)g(x), \quad \frac{f(x)}{g(x)}, \quad f(g(x))\,?$$

(Assume that the quotient and the composition are properly defined and thus at least continuous.)

Exercise 4.4.11 (Topological Characterization of Continuity). Let g be defined on all of $\mathbf{R}$. If B is a subset of $\mathbf{R}$, define the set $g^{-1}(B)$ by

$$g^{-1}(B) = \{x \in \mathbf{R} : g(x) \in B\}.$$

Show that g is continuous if and only if $g^{-1}(O)$ is open whenever $O \subseteq \mathbf{R}$ is an open set.

Exercise 4.4.12. Review Exercise 4.4.11, and then determine which of the following statements is true about a continuous function defined on $\mathbf{R}$:

(a) $f^{-1}(B)$ is finite whenever B is finite.

(b) $f^{-1}(K)$ is compact whenever K is compact.

(c) $f^{-1}(A)$ is bounded whenever A is bounded.

(d) $f^{-1}(F)$ is closed whenever F is closed.

Exercise 4.4.13 (Continuous Extension Theorem). (a) Show that a uniformly continuous function preserves Cauchy sequences; that is, if $f : A \to \mathbf{R}$ is uniformly continuous and $(x_n) \subseteq A$ is a Cauchy sequence, then show $f(x_n)$ is a Cauchy sequence.

(b) Let g be a continuous function on the open interval (a, b). Prove that g is uniformly continuous on (a, b) if and only if it is possible to define values $g(a)$ and $g(b)$ at the endpoints so that the extended function g is continuous on $[a, b]$. (In the forward direction, first produce candidates for $g(a)$ and $g(b)$, and then show the extended g is continuous.)

Exercise 4.4.14. Construct an alternate proof of Theorem 4.4.7 using the open cover characterization of compactness from the Heine–Borel Theorem (Theorem 3.3.8 (iii)).

4.5 The Intermediate Value Theorem

The Intermediate Value Theorem (IVT) is the name given to the very intuitive observation that a continuous function f on a closed interval $[a, b]$ attains every value that falls between the range values $f(a)$ and $f(b)$ (Fig. 4.8).

Here is this observation in the language of analysis.

Theorem 4.5.1 (Intermediate Value Theorem). *Let $f : [a, b] \to \mathbf{R}$ be continuous. If L is a real number satisfying $f(a) < L < f(b)$ or $f(a) > L > f(b)$, then there exists a point $c \in (a, b)$ where $f(c) = L$.*

This theorem was freely used by mathematicians of the 18th century (including Euler and Gauss) without any consideration of its validity. In fact, the first analytical proof was not offered until 1817 by Bolzano in a paper that also contains the first appearance of a somewhat modern definition of continuity. This emphasizes the significance of this result. As discussed in Section 4.1, Bolzano and his contemporaries had arrived at a point in the evolution of mathematics where it was becoming increasingly important to firm up the foundations of the subject. Doing so, however, was not simply a matter of going back and supplying the missing proofs. The real battle lay in first obtaining a thorough and mutually agreed-upon understanding of the relevant concepts. The importance of the Intermediate Value Theorem for us is similar in that our understanding of continuity and the nature of the real line is now mature enough for a proof *to be possible.* Indeed, there are several satisfying arguments for this simple result, each one isolating, in a slightly different way, the interplay between continuity and completeness.

Preservation of Connected Sets

The most potentially useful way to understand the Intermediate Value Theorem (IVT) is as a special case of the fact that continuous functions map connected sets to connected sets. In Theorem 4.4.1, we saw that if f is a continuous function on a compact set K, then the range set $f(K)$ is also compact. The analogous observation holds for connected sets.

Theorem 4.5.2 (Preservation of Connected Sets). *Let $f : G \to \mathbf{R}$ be continuous. If $E \subseteq G$ is connected, then $f(E)$ is connected as well.*

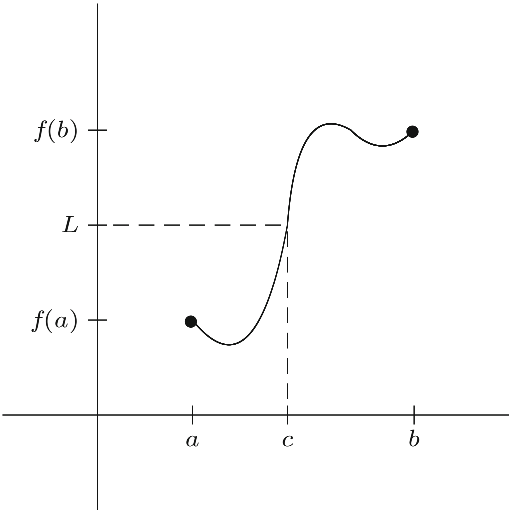


Figure 4.8: INTERMEDIATE VALUE THEOREM.

Proof. Intending to use the characterization of connected sets in Theorem 3.4.6, let $f(E) = A \cup B$ where A and B are disjoint and nonempty. Our goal is to produce a sequence contained in one of these sets that converges to a limit in the other.

Let

$$C = \{x \in E : f(x) \in A\} \quad \text{and} \quad D = \{x \in E : f(x) \in B\}.$$

The sets C and D are called the *preimages* of A and B, respectively. Using the properties of A and B, it is straightforward to check that C and D are nonempty and disjoint and satisfy $E = C \cup D$. Now, we are assuming E is a connected set, so by Theorem 3.4.6, there exists a sequence (x_n) contained in one of C or D with $x = \lim x_n$ contained in the other. Finally, because f is continuous at x, we get $f(x) = \lim f(x_n)$. Thus, it follows that $f(x_n)$ is a convergent sequence contained in either A or B while the limit $f(x)$ is an element of the other. With another nod to Theorem 3.4.6, the proof is complete. □

In **R**, a set is connected if and only if it is a (possibly unbounded) interval. This fact, together with Theorem 4.5.2, leads to a short proof of the Intermediate Value Theorem (Exercise 4.5.1). We should point out that the proof of Theorem 4.5.2 does not make use of the equivalence between connected sets and intervals in **R** but relies only on the general definitions. The previous comment that this is the most *useful* way to approach IVT stems from the fact that, although it is not discussed here, the definitions of continuity and connectedness can be easily adapted to higher-dimensional settings. Theorem 4.5.2, then, remains a valid conclusion in higher dimensions, whereas the Intermediate Value Theorem is essentially a one-dimensional result.

Completeness

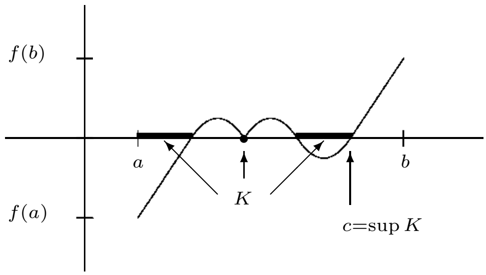
A typical way the Intermediate Value Theorem is applied is to prove the existence of roots. Given $f(x) = x^2 - 2$, for instance, we see that $f(1) = -1$ and $f(2) = 2$. Therefore, there exists a point $c \in (1, 2)$ where $f(c) = 0$.

In this case, we can easily compute $c = \sqrt{2}$, meaning that we really did not need IVT to show that f has a root. We spent a good deal of time in Chapter 1 proving that $\sqrt{2}$ exists, which was only possible once we insisted on the Axiom of Completeness as part of our assumptions about the real numbers. The fact that the Intermediate Value Theorem has just asserted that $\sqrt{2}$ exists suggests that another way to understand this result is in terms of the relationship between the continuity of f and the completeness of $\mathbf{R}$.

The Axiom of Completeness (AoC) from the first chapter states that "Nonempty sets that are bounded above have least upper bounds." Later, we saw that the Nested Interval Property (NIP) is an equivalent way to assert that the real numbers have no "gaps." Either of these characterizations of completeness can be used as the cornerstone for an alternate proof of Theorem 4.5.1.

Proof. **I.** (*First approach using AoC.*) To simplify matters a bit, let's consider the special case where f is a continuous function satisfying $f(a) < 0 < f(b)$ and show that $f(c) = 0$ for some $c \in (a, b)$. First let

$$K = \{x \in [a, b] : f(x) \leq 0\}.$$



Notice that K is bounded above by b, and $a \in K$ so K is not empty. Thus we may appeal to the Axiom of Completeness to assert that $c = \sup K$ exists.

There are three cases to consider:

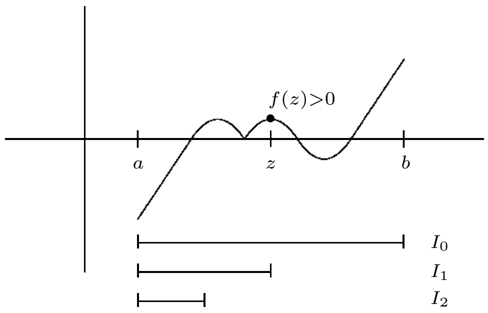
$$f(c) > 0, \ f(c) < 0, \text{ and } f(c) = 0.$$

The fact that c is the least upper bound of K can be used to rule out the first two cases, resulting in the desired conclusion that $f(c) = 0$. The details are requested in Exercise 4.5.5(a).

II. (*Second approach using NIP.*) Again, consider the special case where $L = 0$ and $f(a) < 0 < f(b)$. Let $I_0 = [a, b]$, and consider the midpoint

$$z = (a + b)/2.$$

If $f(z) \geq 0$, then set $a_1 = a$ and $b_1 = z$. If $f(z) < 0$, then set $a_1 = z$ and $b_1 = b$. In either case, the interval $I_1 = [a_1, b_1]$ has the property that f is negative at the left endpoint and nonnegative at the right.



This procedure can be inductively repeated, setting the stage for an application of the Nested Interval Property. The remainder of the argument is left as Exercise 4.5.5(b). □

The Intermediate Value Property

Does the Intermediate Value Theorem have a converse?

Definition 4.5.3. A function f has the *intermediate value property* on an interval $[a, b]$ if for all $x < y$ in $[a, b]$ and all L between $f(x)$ and $f(y)$, it is always possible to find a point $c \in (x, y)$ where $f(c) = L$.

Another way to summarize the Intermediate Value Theorem is to say that every continuous function on $[a, b]$ has the intermediate value property. There is an understandable temptation to suspect that any function that has the intermediate value property must necessarily be continuous, but that is not the case. We have seen that

$$g(x) = \begin{cases} \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is not continuous at zero (Example 4.2.6), but it does have the intermediate value property on $[0, 1]$.

The intermediate value property *does* imply continuity if we insist that our function is monotone (Exercise 4.5.3).

Exercises

Exercise 4.5.1. Show how the Intermediate Value Theorem follows as a corollary to Theorem 4.5.2.

Exercise 4.5.2. Provide an example of each of the following, or explain why the request is impossible

(a) A continuous function defined on an open interval with range equal to a closed interval.

(b) A continuous function defined on a closed interval with range equal to an open interval.

(c) A continuous function defined on an open interval with range equal to an unbounded closed set different from $\mathbf{R}$.

(d) A continuous function defined on all of $\mathbf{R}$ with range equal to $\mathbf{Q}$.

Exercise 4.5.3. A function f is *increasing* on A if $f(x) \leq f(y)$ for all $x < y$ in A. Show that if f is increasing on $[a, b]$ and satisfies the intermediate value property (Definition 4.5.3), then f is continuous on $[a, b]$.

Exercise 4.5.4. Let g be continuous on an interval A and let F be the set of points where g fails to be one-to-one; that is,

$$F = \{x \in A : f(x) = f(y) \text{ for some } y \neq x \text{ and } y \in A\}.$$

Show F is either empty or uncountable.

Exercise 4.5.5. (a) Finish the proof of the Intermediate Value Theorem using the Axiom of Completeness started previously.

(b) Finish the proof of the Intermediate Value Theorem using the Nested Interval Property started previously.

Exercise 4.5.6. Let $f : [0, 1] \to \mathbf{R}$ be continuous with $f(0) = f(1)$.

(a) Show that there must exist $x, y \in [0, 1]$ satisfying $|x - y| = 1/2$ and $f(x) = f(y)$.

(b) Show that for each $n \in \mathbf{N}$ there exist $x_n, y_n \in [0, 1]$ with $|x_n - y_n| = 1/n$ and $f(x_n) = f(y_n)$.

(c) If $h \in (0, 1/2)$ is not of the form $1/n$, there does not necessarily exist $|x - y| = h$ satisfying $f(x) = f(y)$. Provide an example that illustrates this using $h = 2/5$.

Exercise 4.5.7. Let f be a continuous function on the closed interval $[0, 1]$ with range also contained in $[0, 1]$. Prove that f must have a fixed point; that is, show $f(x) = x$ for at least one value of $x \in [0, 1]$.

Exercise 4.5.8 (Inverse functions). If a function $f : A \to \mathbf{R}$ is one-to-one, then we can define the inverse function f^{-1} on the range of f in the natural way: $f^{-1}(y) = x$ where $y = f(x)$.

Show that if f is continuous on an interval $[a, b]$ and one-to-one, then f^{-1} is also continuous.

4.6 Sets of Discontinuity

Given a function $f : \mathbf{R} \to \mathbf{R}$, define $D_f \subseteq \mathbf{R}$ to be the set of points where the function f fails to be continuous. In Section 4.1, we saw that Dirichlet's function $g(x)$ had $D_g = \mathbf{R}$. The modification $h(x)$ of Dirichlet's function had $D_h = \mathbf{R}\backslash\{0\}$, zero being the only point of continuity. Finally, for Thomae's function $t(x)$, we saw that $D_t = \mathbf{Q}$.

Exercise 4.6.1. Using modifications of these functions, construct a function $f : \mathbf{R} \to \mathbf{R}$ so that

(a) $D_f = \mathbf{Z}^c$.

(b) $D_f = \{x : 0 < x \leq 1\}$.

Exercise 4.6.2. Given a countable set $A = \{a_1, a_2, a_3, \ldots\}$, define $f(a_n) = 1/n$ and $f(x) = 0$ for all $x \notin A$. Find D_f.

We concluded the introduction with a question about whether D_f could take the form of *any* arbitrary subset of the real line. As it turns out, this is not the case. The set of discontinuities of a real-valued function on $\mathbf{R}$ has a specific topological structure that is not possessed by every subset of $\mathbf{R}$. Specifically, D_f, no matter how f is chosen, can always be written as the countable union of closed sets. In the case where f is *monotone*, these closed sets can be taken to be single points.

Monotone Functions

Classifying D_f for an arbitrary f is somewhat involved, so it is interesting that describing D_f is fairly straightforward for the class of monotone functions.

Definition 4.6.1. A function $f : A \to \mathbf{R}$ is *increasing on* A if $f(x) \leq f(y)$ whenever $x < y$ and *decreasing* if $f(x) \geq f(y)$ whenever $x < y$ in A. A *monotone* function is one that is either increasing or decreasing.

Continuity of f at a point c means that $\lim_{x\to c} f(x) = f(c)$. One particular way for a discontinuity to occur is if the limit from the right at c is different from the limit from the left at c. As always with new terminology, we need to be precise about what we mean by "from the left" and "from the right."

Definition 4.6.2. Given a limit point c of a set A and a function $f : A \to \mathbf{R}$, we write

$$\lim_{x\to c^+} f(x) = L$$

if for all $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(x)-L| < \epsilon$ whenever $0 < x-c < \delta$.

Equivalently, in terms of sequences, $\lim_{x\to c^+} f(x) = L$ if $\lim f(x_n) = L$ for all sequences (x_n) satisfying $x_n > c$ and $\lim(x_n) = c$.

Exercise 4.6.3. State a similar definition for the left-hand limit

$$\lim_{x \to c^-} f(x) = L.$$

Theorem 4.6.3. *Given $f : A \to \mathbf{R}$ and a limit point c of A, $\lim_{x\to c} f(x) = L$ if and only if*

$$\lim_{x \to c^-} f(x) = L \quad \textit{and} \quad \lim_{x \to c^+} f(x) = L.$$

Exercise 4.6.4. Supply a proof for this proposition.

Generally speaking, discontinuities can be divided into three categories:

(i) If $\lim_{x\to c} f(x)$ exists but has a value different from $f(c)$, the discontinuity at c is called *removable*.

(ii) If $\lim_{x\to c^+} f(x) \neq \lim_{x\to c^-} f(x)$, then f has a *jump* discontinuity at c.

(iii) If $\lim_{x\to c} f(x)$ does not exist for some other reason, then the discontinuity at c is called an *essential* discontinuity.

We are now equipped to characterize the set D_f for an arbitrary monotone function f.

Exercise 4.6.5. Prove that the only type of discontinuity a monotone function can have is a jump discontinuity.

Exercise 4.6.6. Construct a bijection between the set of jump discontinuities of a monotone function f and a subset of $\mathbf{Q}$. Conclude that D_f for a monotone function f must either be finite or countable, but not uncountable.

D_f for an Arbitrary Function

Recall that the intersection of an infinite collection of closed sets is closed, but for unions we must restrict ourselves to *finite* collections of closed sets in order to ensure the union is closed. For open sets the situation is reversed. The arbitrary union of open sets is open, but only finite intersections of open sets are necessarily open.

Definition 4.6.4. A set that can be written as the countable union of closed sets is in the class F_σ. (This definition also appeared in Section 3.5.)

In Section 4.1 we constructed functions where the set of discontinuity was $\mathbf{R}$ (Dirichlet's function), $\mathbf{R}\backslash\{0\}$ (modified Dirichlet function), and $\mathbf{Q}$ (Thomae's function).

Exercise 4.6.7. (a) Show that in each of the above cases we get an F_σ set as the set where the function is discontinuous.

(b) Show that the two sets of discontinuity in Exercise 4.6.1 are F_σ sets.

The upcoming argument depends on a concept called α-continuity.

Definition 4.6.5. Let f be defined on $\mathbf{R}$, and let $\alpha > 0$. The function f is *α-continuous at $x \in \mathbf{R}$* if there exists a $\delta > 0$ such that for all $y, z \in (x-\delta, x+\delta)$ it follows that $|f(y) - f(z)| < \alpha$.

The most important thing to note about this definition is that there is no "for all" in front of the $\alpha > 0$. As we will investigate, adding this quantifier would make this definition equivalent to our definition of continuity. In a sense, α-continuity is a measure of the variation of the function in the neighborhood of a particular point. A function is α-continuous at a point c if there is some interval centered at c in which the variation of the function never exceeds the value $\alpha > 0$.

Given a function f on $\mathbf{R}$, define D_f^α to be the set of points where the function f fails to be α-continuous. In other words,

$$D_f^\alpha = \{x \in \mathbf{R} : f \text{ is not } \alpha\text{-continuous at } x\}.$$

Exercise 4.6.8. Prove that, for a fixed $\alpha > 0$, the set D_f^α is closed.

The stage is set. It is time to characterize the set of discontinuity for an arbitrary function f on $\mathbf{R}$.

Theorem 4.6.6. *Let $f : \mathbf{R} \to \mathbf{R}$ be an arbitrary function. Then, D_f is an F_σ set.*

Proof. Recall that

$$D_f = \{x \in \mathbf{R} : f \text{ is not continuous at } x\}.$$

Exercise 4.6.9. If $\alpha < \alpha'$, show that $D_f^{\alpha'} \subseteq D_f^\alpha$.

Exercise 4.6.10. Let $\alpha > 0$ be given. Show that if f is continuous at x, then it is α-continuous at x as well. Explain how it follows that $D_f^\alpha \subseteq D_f$.

Exercise 4.6.11. Show that if f is not continuous at x, then f is not α-continuous for some $\alpha > 0$. Now explain why this guarantees that

$$D_f = \bigcup_{n=1}^{\infty} D_f^{\alpha_n},$$

where $\alpha_n = 1/n$.

Because each $D_f^{\alpha_n}$ is closed, the proof is complete. $\square$

4.7 Epilogue

Theorem 4.6.6 is only interesting if we can demonstrate that not every subset of **R** is in an F_σ set. This takes some effort and was included as an exercise in Section 3.5 on the Baire Category Theorem. Baire's Theorem states that if **R** is written as the countable union of closed sets, then at least one of these sets must contain a nonempty open interval. Now **Q** is the countable union of singleton points, and we can view each point as a closed set that obviously contains no intervals. If the set of irrationals **I** were a countable union of closed sets, it would have to be that none of these closed sets contained any open intervals or else they would then contain some rational numbers. But this leads to a contradiction to Baire's Theorem. Thus, **I** is not the countable union of closed sets, and consequently it is not an F_σ set. We may therefore conclude that there is no function f that is continuous at every rational point and discontinuous at every irrational point. This should be compared with Thomae's function discussed earlier.

The converse question is interesting as well. Given an arbitrary F_σ set, W.H. Young showed in 1903 that it is always possible to construct a function that has discontinuities precisely on this set. Exercise 4.3.14 gives some clues for how to do this in the simpler case of an arbitrary closed set, and Exercise 4.6.2 handles the case of an arbitrary countable set. Combining the techniques in these two exercises with the Dirichlet-type definitions we have seen leads to a proof of Young's result. (Try it!) A function demonstrating the converse for the monotone case described in Exercise 4.6.6 is also not too difficult to describe. Let

$$D = \{x_1, x_2, x_3, x_4, \ldots\}$$

be an arbitrary countable set of real numbers. In order to construct a monotone function that has discontinuities precisely on D, we first consider a particular $x_n \in D$ and define the step function

$$u_n(x) = \begin{cases} 1/2^n & \text{for } x > x_n \\ 0 & \text{for } x \leq x_n. \end{cases}$$

Observing that each $u_n(x)$ is monotone and everywhere continuous except for a single discontinuity at x_n, we now set

$$f(x) = \sum_{n=1}^{\infty} u_n(x).$$

The convergence of the series $\sum 1/2^n$ guarantees that our function f is defined on all of **R**, and intuition certainly suggests that f is monotone with jump discontinuities precisely on D. Providing a rigorous proof for this conclusion is one of the many pleasures that awaits in Chapter 6, where we take up the study of infinite series of functions.