Chapter 3

Basic Topology of R

3.1 Discussion: The Cantor Set

What follows is a fascinating mathematical construction, due to Georg Cantor, which is extremely useful for extending the horizons of our intuition about the nature of subsets of the real line. Cantor's name has already appeared in the first chapter in our discussion of uncountable sets. Indeed, Cantor's proof that **R** is uncountable occupies another spot on the short list of the most significant contributions toward understanding the mathematical infinite. In the words of the mathematician David Hilbert, "No one shall expel us from the paradise that Cantor has created for us."

Let C_0 be the closed interval $[0, 1]$, and define C_1 to be the set that results when the open middle third is removed; that is,

$$C_1 = C_0 \backslash \left(\frac{1}{3}, \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Now, construct C_2 in a similar way by removing the open middle third of each of the two components of C_1:

$$C_2 = \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right]\right) \cup \left(\left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]\right).$$

If we continue this process inductively, then for each $n = 0, 1, 2, \ldots$ we get a set C_n consisting of 2^n closed intervals each having length $1/3^n$. Finally, we define the *Cantor set* C (Fig. 3.1) to be the intersection

$$C = \bigcap_{n=0}^{\infty} C_n.$$

S. Abbott, *Understanding Analysis*, Undergraduate Texts in Mathematics, DOI 10.1007/978-1-4939-2712-8_3

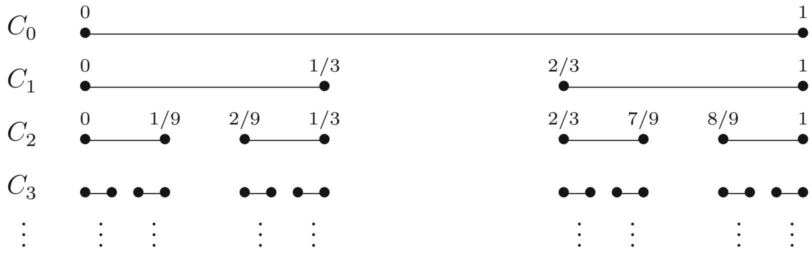


Figure 3.1: DEFINING THE CANTOR SET; $C = \bigcap_{n=0}^{\infty} C_n$.

It may be useful to understand C as the remainder of the interval $[0,1]$ after the iterative process of removing open middle thirds is taken to infinity:

$$C = [0,1] \backslash \left[\left(\frac{1}{3}, \frac{2}{3} \right) \cup \left(\frac{1}{9}, \frac{2}{9} \right) \cup \left(\frac{7}{9}, \frac{8}{9} \right) \cup \cdots \right].$$

There is some initial doubt whether anything remains at all, but notice that because we are always removing open middle thirds, then for every $n \in \mathbf{N}$, $0 \in C_n$ and hence $0 \in C$. The same argument shows $1 \in C$. In fact, if y is the endpoint of some closed interval of some particular set C_n, then it is also an endpoint of one of the intervals of C_{n+1}. Because, at each stage, endpoints are never removed, it follows that $y \in C_n$ for all n. Thus, C at least contains the endpoints of all of the intervals that make up each of the sets C_n.

Is there anything else? Is C countable? Does C contain any intervals? Any irrational numbers? These are difficult questions at the moment. All of the endpoints mentioned earlier are rational numbers (they have the form $m/3^n$), which means that if it is true that C consists of only these endpoints, then C would be a subset of $\mathbf{Q}$ and hence countable. We shall see about this. There is some strong evidence that not much is left in C if we consider the total length of the intervals removed. To form C_1, an open interval of length $1/3$ was taken out. In the second step, we removed two intervals of length $1/9$, and to construct C_n we removed 2^{n-1} middle thirds of length $1/3^n$. There is some logic, then, to defining the "length" of C to be 1 minus the total

$$\frac{1}{3} + 2\left(\frac{1}{9}\right) + 4\left(\frac{1}{27}\right) + \cdots + 2^{n-1}\left(\frac{1}{3^n}\right) + \cdots = \frac{\frac{1}{3}}{1-\frac{2}{3}} = 1.$$

The Cantor set has *zero length.*

To this point, the information we have collected suggests a mental picture of C as a relatively small, thin set. For these reasons, the set C is often referred to as Cantor "dust." But there are some strong counterarguments that imply a very different picture. First, C is actually *uncountable*, with cardinality equal to the cardinality of $\mathbf{R}$. One slightly intuitive but convincing way to see this is to create a 1–1 correspondence between C and sequences of the form $(a_n)_{n=1}^{\infty}$, where $a_n = 0$ or 1. For each $c \in C$, set $a_1 = 0$ if c falls in the left-hand component

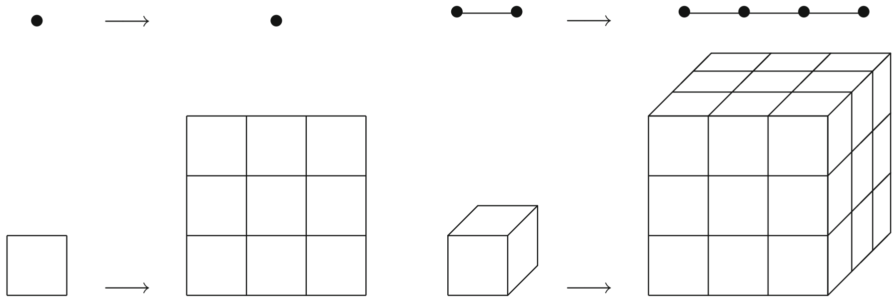


Figure 3.2: MAGNIFYING SETS BY A FACTOR OF 3.

of C_1 and set $a_1 = 1$ if c falls in the right-hand component. Having established where in C_1 the point c is located, there are now two possible components of C_2 that might contain c. This time, we set $a_2 = 0$ or 1 depending on whether c falls in the left or right half of these two components of C_2. Continuing in this way, we come to see that every element $c \in C$ yields a sequence $(a_1, a_2, a_3, \ldots)$ of zeros and ones that acts as a set of directions for how to locate c within C. Likewise, every such sequence corresponds to a point in the Cantor set. Because the set of sequences of zeros and ones is uncountable (Exercise 1.6.4), we must conclude that C is uncountable as well.

What does this imply? In the first place, because the endpoints of the approximating sets C_n form a countable set, we are forced to accept the fact that not only are there other points in C but there are uncountably many of them. From the point of view of *cardinality*, C is quite large—as large as $\mathbf{R}$, in fact. This should be contrasted with the fact that from the point of view of *length*, C measures the same size as a single point. We conclude this discussion with a demonstration that from the point of view of *dimension*, C strangely falls somewhere in between.

There is a sensible agreement that a point has dimension zero, a line segment has dimension one, a square has dimension two, and a cube has dimension three. Without attempting a formal definition of dimension (of which there are several), we can nevertheless get a sense of how one might be defined by observing how the dimension affects the result of magnifying each particular set by a factor of 3 (Fig. 3.2). (The reason for the choice of 3 will become clear when we turn our attention back to the Cantor set). A single point undergoes no change at all, whereas a line segment triples in length. For the square, magnifying each length by a factor of 3 results in a larger square that contains 9 copies of the original square. Finally, the magnified cube yields a cube that contains 27 copies of the original cube within its volume. Notice that, in each case, to compute the "size" of the new set, the dimension appears as the exponent of the magnification factor.

	dim	×3	new copies
point	0	$\to$	$1 = 3^0$
segment	1	$\to$	$3 = 3^1$
square	2	$\to$	$9 = 3^2$
cube	3	$\to$	$27 = 3^3$
C	x	$\to$	$2 = 3^x$

Figure 3.3: DIMENSION OF C; $2 = 3^x \Rightarrow x = \log 2/\log 3$.

Now, apply this transformation to the Cantor set. The set $C_0 = [0,1]$ becomes the interval $[0,3]$. Deleting the middle third leaves $[0,1] \cup [2,3]$, which is where we started in the original construction except that we now stand to produce an additional copy of C in the interval $[2,3]$. Magnifying the Cantor set by a factor of 3 yields *two* copies of the original set. Thus, if x is the dimension of C, then x should satisfy $2 = 3^x$, or $x = \log 2/\log 3 \approx .631$ (Fig. 3.3).

The notion of a noninteger or fractional dimension is the impetus behind the term "fractal," coined in 1975 by Benoit Mandlebrot to describe a class of sets whose intricate structures have much in common with the Cantor set. Cantor's construction, however, is over a hundred years old and for us represents an invaluable testing ground for the upcoming theorems and conjectures about the often elusive nature of subsets of the real line.

3.2 Open and Closed Sets

Given $a \in \mathbf{R}$ and $\epsilon > 0$, recall that the ϵ-*neighborhood* of a is the set

$$V_\epsilon(a) = \{x \in \mathbf{R} : |x - a| < \epsilon\}.$$

In other words, $V_\epsilon(a)$ is the open interval $(a - \epsilon, a + \epsilon)$, centered at a with radius ϵ.

Definition 3.2.1. A set $O \subseteq \mathbf{R}$ is *open* if for all points $a \in O$ there exists an ϵ-neighborhood $V_\epsilon(a) \subseteq O$.

Example 3.2.2. (i) Perhaps the simplest example of an open set is $\mathbf{R}$ itself. Given an arbitrary element $a \in \mathbf{R}$, we are free to pick any ϵ-neighborhood we like and it will always be true that $V_\epsilon(a) \subseteq \mathbf{R}$. It is also the case that the logical structure of Definition 3.2.1 requires us to classify the empty set $\emptyset$ as an open subset of the real line.

(ii) For a more useful collection of examples, consider the open interval

$$(c, d) = \{x \in \mathbf{R} : c < x < d\}.$$

To see that (c,d) is open in the sense just defined, let $x \in (c,d)$ be arbitrary. If we take $\epsilon = \min\{x-c, d-x\}$, then it follows that $V_\epsilon(x) \subseteq (c,d)$. It is important to see where this argument breaks down if the interval includes either one of its endpoints.

The union of open intervals is another example of an open set. This observation leads to the next result.

Theorem 3.2.3. (i) *The union of an arbitrary collection of open sets is open.*

(ii) *The intersection of a finite collection of open sets is open.*

Proof. To prove (i), we let $\{O_\lambda : \lambda \in \Lambda\}$ be a collection of open sets and let $O = \bigcup_{\lambda\in\Lambda} O_\lambda$. Let a be an arbitrary element of O. In order to show that O is open, Definition 3.2.1 insists that we produce an ϵ-neighborhood of a completely contained in O. But $a \in O$ implies that a is an element of at least one particular $O_{\lambda'}$. Because we are assuming $O_{\lambda'}$ is open, we can use Definition 3.2.1 to assert that there exists $V_\epsilon(a) \subseteq O_{\lambda'}$. The fact that $O_{\lambda'} \subseteq O$ allows us to conclude that $V_\epsilon(a) \subseteq O$. This completes the proof of (i).

For (ii), let $\{O_1, O_2, \ldots, O_N\}$ be a finite collection of open sets. Now, if $a \in \bigcap_{k=1}^N O_k$, then a is an element of each of the open sets. By the definition of an open set, we know that, for each $1 \leq k \leq N$, there exists $V_{\epsilon_k}(a) \subseteq O_k$. We are in search of a *single* ϵ-neighborhood of a that is contained in every O_k, so the trick is to take the smallest one. Letting $\epsilon = \min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_N\}$, it follows that $V_\epsilon(a) \subseteq V_{\epsilon_k}(a)$ for all k, and hence $V_\epsilon(a) \subseteq \bigcap_{k=1}^N O_k$, as desired. □

Closed Sets

Definition 3.2.4. A point x is a *limit point* of a set A if every ϵ-neighborhood $V_\epsilon(x)$ of x intersects the set A at some point other than x.

Limit points are also often referred to as "cluster points" or "accumulation points," but the phrase "x is a limit point of A" has the advantage of explicitly reminding us that x is quite literally the limit of a sequence in A.

Theorem 3.2.5. *A point x is a limit point of a set A if and only if $x = \lim a_n$ for some sequence (a_n) contained in A satisfying $a_n \neq x$ for all $n \in \mathbf{N}$.*

Proof. ($\Rightarrow$) Assume x is a limit point of A. In order to produce a sequence (a_n) converging to x, we are going to consider the particular ϵ-neighborhoods obtained using $\epsilon = 1/n$. By Definition 3.2.4, every neighborhood of x intersects A in some point other than x. This means that, for each $n \in \mathbf{N}$, we are justified in picking a point

$$a_n \in V_{1/n}(x) \cap A$$

with the stipulation that $a_n \neq x$. It should not be too difficult to see why $(a_n) \to x$. Given an arbitrary $\epsilon > 0$, choose N such that $1/N < \epsilon$. It follows that $|a_n - x| < \epsilon$ for all $n \geq N$.

($\Leftarrow$) For the reverse implication we assume $\lim a_n = x$ where $a_n \in A$ but $a_n \neq x$, and let $V_\epsilon(x)$ be an arbitrary ϵ-neighborhood. The definition of convergence assures us that there exists a term a_N in the sequence satisfying $a_N \in V_\epsilon(x)$, and the proof is complete. □

The restriction that $a_n \neq x$ in Theorem 3.2.5 deserves a comment. Given a point $a \in A$, it is always the case that a is the limit of a sequence in A if we are allowed to consider the constant sequence $(a, a, a, \ldots)$. There will be occasions where we will want to avoid this somewhat uninteresting situation, so it is important to have a vocabulary that can distinguish limit points of a set from *isolated points*.

Definition 3.2.6. A point $a \in A$ is an *isolated point of* A if it is not a limit point of A.

As a word of caution, we need to be a little careful about how we understand the relationship between these concepts. Whereas an isolated point is always an element of the relevant set A, it is quite possible for a limit point of A not to belong to A. As an example, consider the endpoint of an open interval. This situation is the subject of the next important definition.

Definition 3.2.7. A set $F \subseteq \mathbf{R}$ is *closed* if it contains its limit points.

The adjective "closed" appears in several other mathematical contexts and is usually employed to mean that an operation on the elements of a given set does not take us out of the set. In linear algebra, for example, a vector space is a set that is "closed" under addition and scalar multiplication. In analysis, the operation we are concerned with is the limiting operation. Topologically speaking, a closed set is one where convergent sequences within the set have limits that are also in the set.

Theorem 3.2.8. *A set $F \subseteq \mathbf{R}$ is closed if and only if every Cauchy sequence contained in F has a limit that is also an element of F.*

Proof. Exercise 3.2.5. □

Example 3.2.9. (i) Consider

$$A = \left\{ \frac{1}{n} : n \in \mathbf{N} \right\}.$$

Let's show that each point of A is isolated. Given $1/n \in A$, choose $\epsilon = 1/n - 1/(n+1)$. Then,

$$V_\epsilon(1/n) \cap A = \left\{ \frac{1}{n} \right\}.$$

It follows from Definition 3.2.4 that $1/n$ is not a limit point and so is isolated. Although all of the points of A are isolated, the set does have one limit point, namely 0. This is because every neighborhood centered at zero, no matter how small, is going to contain points of A. Because $0 \notin A$, A is not closed. The set $F = A \cup \{0\}$ is an example of a closed set and is called the *closure of* A. (The closure of a set is discussed in a moment.)

(ii) Let's prove that a closed interval

$$[c, d] = \{x \in \mathbf{R} : c \leq x \leq d\}$$

is a closed set using Definition 3.2.7. If x is a limit point of $[c, d]$, then by Theorem 3.2.5 there exists $(x_n) \subseteq [c, d]$ with $(x_n) \to x$. We need to prove that $x \in [c, d]$.

The key to this argument is contained in the Order Limit Theorem (Theorem 2.3.4), which summarizes the relationship between inequalities and the limiting process. Because $c \leq x_n \leq d$, it follows from Theorem 2.3.4 (iii) that $c \leq x \leq d$ as well. Thus, $[c, d]$ is closed.

(iii) Consider the set $\mathbf{Q} \subseteq \mathbf{R}$ of rational numbers. An extremely important property of $\mathbf{Q}$ is that its set of limit points is actually *all of* $\mathbf{R}$. To see why this is so, recall Theorem 1.4.3 from Chapter 1, which is referred to as the density property of $\mathbf{Q}$ in $\mathbf{R}$.

Let $y \in \mathbf{R}$ be arbitrary, and consider any neighborhood $V_\epsilon(y) = (y - \epsilon, y + \epsilon)$. Theorem 1.4.3 allows us to conclude that there exists a rational number $r \neq y$ that falls in this neighborhood. Thus, y is a limit point of $\mathbf{Q}$.

The density property of $\mathbf{Q}$ can now be reformulated in the following way.

Theorem 3.2.10 (Density of $\mathbf{Q}$ in $\mathbf{R}$). *For every $y \in \mathbf{R}$, there exists a sequence of rational numbers that converges to y.*

Proof. Combine the preceding discussion with Theorem 3.2.5. □

The same argument can also be used to show that every real number is the limit of a sequence of irrational numbers. Although interesting, part of the allure of the rational numbers is that, in addition to being dense in $\mathbf{R}$, they are countable. As we will see, this tangible aspect of $\mathbf{Q}$ makes it an extremely useful set, both for proving theorems and for producing interesting counterexamples.

Closure

Definition 3.2.11. Given a set $A \subseteq \mathbf{R}$, let L be the set of all limit points of A. The *closure* of A is defined to be $\overline{A} = A \cup L$.

In Example 3.2.9 (i), we saw that if $A = \{1/n : n \in \mathbf{N}\}$, then the closure of A is $\overline{A} = A \cup \{0\}$. Example 3.2.9 (iii) verifies that $\overline{\mathbf{Q}} = \mathbf{R}$. If A is an open interval (a, b), then $\overline{A} = [a, b]$. If A is a closed interval, then $\overline{A} = A$. It is not for lack of imagination that in each of these examples $\overline{A}$ is always a closed set.

Theorem 3.2.12. *For any $A \subseteq \mathbf{R}$, the closure $\overline{A}$ is a closed set and is the smallest closed set containing A.*

Proof. If L is the set of limit points of A, then it is immediately clear that $\overline{A}$ contains the limit points of A. There is still something more to prove, however, because taking the union of L with A could potentially produce some new limit points of $\overline{A}$. In Exercise 3.2.7, we outline the argument that this does not happen.

Now, *any* closed set containing A must contain L as well. This shows that $\overline{A} = A \cup L$ is the smallest closed set containing A. □

Complements

The mathematical notions of open and closed are not antonyms the way they are in standard English. If a set is not open, that does not imply it must be closed. Many sets such as the half-open interval $(c, d] = \{x \in \mathbf{R} : c < x \leq d\}$ are neither open nor closed. The sets $\mathbf{R}$ and $\emptyset$ are both simultaneously open and closed although, thankfully, these are the only ones with this disorienting property (Exercise 3.2.13). There is, however, an important relationship between open and closed sets. Recall that the *complement* of a set $A \subseteq \mathbf{R}$ is defined to be the set

$$A^c = \{x \in \mathbf{R} : x \notin A\}.$$

Theorem 3.2.13. *A set O is open if and only if O^c is closed. Likewise, a set F is closed if and only if F^c is open.*

Proof. Given an open set $O \subseteq \mathbf{R}$, let's first prove that O^c is a closed set. To prove O^c is closed, we need to show that it contains all of its limit points. If x is a limit point of O^c, then *every* neighborhood of x contains some point of O^c. But that is enough to conclude that x cannot be in the open set O because $x \in O$ would imply that there exists a neighborhood $V_\epsilon(x) \subseteq O$. Thus, $x \in O^c$, as desired.

For the converse statement, we assume O^c is closed and argue that O is open. Thus, given an arbitrary point $x \in O$, we must produce an ϵ-neighborhood $V_\epsilon(x) \subseteq O$. Because O^c is closed, we can be sure that x is *not* a limit point of O^c. Looking at the definition of limit point, we see that this implies that there must be some neighborhood $V_\epsilon(x)$ of x that does not intersect the set O^c. But this means $V_\epsilon(x) \subseteq O$, which is precisely what we needed to show.

The second statement in Theorem 3.2.13 follows quickly from the first using the observation that $(E^c)^c = E$ for any set $E \subseteq \mathbf{R}$. □

The last theorem of this section should be compared to Theorem 3.2.3.

Theorem 3.2.14. (i) *The union of a finite collection of closed sets is closed.*

(ii) *The intersection of an arbitrary collection of closed sets is closed.*

Proof. De Morgan's Laws state that for any collection of sets $\{E_\lambda : \lambda \in \Lambda\}$ it is true that

$$\left(\bigcup_{\lambda\in\Lambda} E_\lambda\right)^c = \bigcap_{\lambda\in\Lambda} E_\lambda^c \quad \text{and} \quad \left(\bigcap_{\lambda\in\Lambda} E_\lambda\right)^c = \bigcup_{\lambda\in\Lambda} E_\lambda^c.$$

The result follows directly from these statements and Theorem 3.2.3. The details are requested in Exercise 3.2.9. □

Exercises

Exercise 3.2.1. (a) Where in the proof of Theorem 3.2.3 part (ii) does the assumption that the collection of open sets be *finite* get used?

(b) Give an example of a countable collection of open sets $\{O_1, O_2, O_3, \ldots\}$ whose intersection $\bigcap_{n=1}^{\infty} O_n$ is closed, not empty and not all of $\mathbf{R}$.

Exercise 3.2.2. Let

$$A = \left\{(-1)^n + \frac{2}{n} : n = 1, 2, 3, \ldots\right\} \quad \text{and} \quad B = \{x \in \mathbf{Q} : 0 < x < 1\}.$$

Answer the following questions for each set:

(a) What are the limit points?

(b) Is the set open? Closed?

(c) Does the set contain any isolated points?

(d) Find the closure of the set.

Exercise 3.2.3. Decide whether the following sets are open, closed, or neither. If a set is not open, find a point in the set for which there is no ϵ-neighborhood contained in the set. If a set is not closed, find a limit point that is not contained in the set.

(a) $\mathbf{Q}$.

(b) $\mathbf{N}$.

(c) $\{x \in \mathbf{R} : x \neq 0\}$.

(d) $\{1 + 1/4 + 1/9 + \cdots + 1/n^2 : n \in \mathbf{N}\}$.

(e) $\{1 + 1/2 + 1/3 + \cdots + 1/n : n \in \mathbf{N}\}$.

Exercise 3.2.4. Let A be nonempty and bounded above so that $s = \sup A$ exists.

(a) Show that $s \in \overline{A}$.

(b) Can an open set contain its supremum?

Exercise 3.2.5. Prove Theorem 3.2.8.

Exercise 3.2.6. Decide whether the following statements are true or false. Provide counterexamples for those that are false, and supply proofs for those that are true.

(a) An open set that contains every rational number must necessarily be all of **R**.

(b) The Nested Interval Property remains true if the term "closed interval" is replaced by "closed set."

(c) Every nonempty open set contains a rational number.

(d) Every bounded infinite closed set contains a rational number.

(e) The Cantor set is closed.

Exercise 3.2.7. Given $A \subseteq \mathbf{R}$, let L be the set of all limit points of A.

(a) Show that the set L is closed.

(b) Argue that if x is a limit point of $A \cup L$, then x is a limit point of A. Use this observation to furnish a proof for Theorem 3.2.12.

Exercise 3.2.8. Assume A is an open set and B is a closed set. Determine if the following sets are definitely open, definitely closed, both, or neither.

(a) $\overline{A \cup B}$

(b) $A \backslash B = \{x \in A : x \notin B\}$

(c) $(A^c \cup B)^c$

(d) $(A \cap B) \cup (A^c \cap B)$

(e) $\overline{A}^c \cap \overline{A^c}$

Exercise 3.2.9 (De Morgan's Laws). A proof for De Morgan's Laws in the case of two sets is outlined in Exercise 1.2.5. The general argument is similar.

(a) Given a collection of sets $\{E_\lambda : \lambda \in \Lambda\}$, show that

$$\left(\bigcup_{\lambda\in\Lambda} E_\lambda\right)^c = \bigcap_{\lambda\in\Lambda} E_\lambda^c \qquad \text{and} \qquad \left(\bigcap_{\lambda\in\Lambda} E_\lambda\right)^c = \bigcup_{\lambda\in\Lambda} E_\lambda^c.$$

(b) Now, provide the details for the proof of Theorem 3.2.14.

Exercise 3.2.10. Only one of the following three descriptions can be realized. Provide an example that illustrates the viable description, and explain why the other two cannot exist.

(i) A countable set contained in $[0,1]$ with no limit points.

(ii) A countable set contained in $[0,1]$ with no isolated points.

(iii) A set with an uncountable number of isolated points.

Exercise 3.2.11. (a) Prove that $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

(b) Does this result about closures extend to infinite unions of sets?

Exercise 3.2.12. Let A be an uncountable set and let B be the set of real numbers that divides A into two uncountable sets; that is, $s \in B$ if both $\{x : x \in A \text{ and } x < s\}$ and $\{x : x \in A \text{ and } x > s\}$ are uncountable. Show B is nonempty and open.

Exercise 3.2.13. Prove that the only sets that are both open and closed are $\mathbf{R}$ and the empty set $\emptyset$.

Exercise 3.2.14. A dual notion to the closure of a set is the interior of a set. The *interior* of E is denoted E° and is defined as

$$E^\circ = \{x \in E : \text{there exists } V_\epsilon(x) \subseteq E\}.$$

Results about closures and interiors possess a useful symmetry.

(a) Show that E is closed if and only if $\overline{E} = E$. Show that E is open if and only if $E^\circ = E$.

(b) Show that $\overline{E}^c = (E^c)^\circ$, and similarly that $(E^\circ)^c = \overline{E^c}$.

Exercise 3.2.15. A set A is called an F_σ set if it can be written as the countable union of closed sets. A set B is called a G_δ set if it can be written as the countable intersection of open sets.

(a) Show that a closed interval $[a,b]$ is a G_δ set.

(b) Show that the half-open interval $(a,b]$ is both a G_δ and an F_σ set.

(c) Show that $\mathbf{Q}$ is an F_σ set, and the set of irrationals $\mathbf{I}$ forms a G_δ set. (We will see in Section 3.5 that $\mathbf{Q}$ is *not* a G_δ set, nor is $\mathbf{I}$ an F_σ set.)

3.3 Compact Sets

The central challenge in analysis is to exploit the power of the mathematical infinite—via limits, series, derivatives, integrals, etc.—without falling victim to erroneous logic or faulty intuition. A major tool for maintaining a rigorous footing in this endeavor is the concept of compact sets. In ways that will become clear, especially in our upcoming study of continuous functions, employing compact sets in a proof often has the effect of bringing a finite quality to the argument, thereby making it much more tractable.

Definition 3.3.1 (Compactness). A set $K \subseteq \mathbf{R}$ is *compact* if every sequence in K has a subsequence that converges to a limit that is also in K.

Example 3.3.2. The most basic example of a compact set is a closed interval. To see this, notice that if (a_n) is contained in an interval $[c, d]$, then the Bolzano–Weierstrass Theorem guarantees that we can find a convergent subsequence (a_{n_k}). Because a closed interval is a closed set (Example 3.2.9, (ii)), we know that the limit of this subsequence is also in $[c, d]$.

What are the properties of closed intervals that we used in the preceding argument? The Bolzano–Weierstrass Theorem requires boundedness, and we used the fact that closed sets contain their limit points. As we are about to see, these two properties completely characterize compact sets in $\mathbf{R}$. The term "bounded" has thus far only been used to describe sequences (Definition 2.3.1), but an analogous statement can easily be made about sets.

Definition 3.3.3. A set $A \subseteq \mathbf{R}$ is *bounded* if there exists $M > 0$ such that $|a| \leq M$ for all $a \in A$.

Theorem 3.3.4 (Characterization of Compactness in R). *A set $K \subseteq \mathbf{R}$ is compact if and only if it is closed and bounded.*

Proof. Let K be compact. We will first prove that K must be bounded, so assume, for contradiction, that K is not a bounded set. The idea is to produce a sequence in K that marches off to infinity in such a way that it cannot have a convergent subsequence as the definition of compact requires. To do this, notice that because K is not bounded there must exist an element $x_1 \in K$ satisfying $|x_1| > 1$. Likewise, there must exist $x_2 \in K$ with $|x_2| > 2$, and in general, given any $n \in \mathbf{N}$, we can produce $x_n \in K$ such that $|x_n| > n$.

Now, because K is assumed to be compact, (x_n) should have a convergent subsequence (x_{n_k}). But the elements of the subsequence must satisfy $|x_{n_k}| > n_k$, and consequently (x_{n_k}) is unbounded. Because convergent sequences are bounded (Theorem 2.3.2), we have a contradiction. Thus, K must at least be a bounded set.

Next, we will show that K is also closed. To see that K contains its limit points, we let $x = \lim x_n$, where (x_n) is contained in K and argue that x must be in K as well. By Definition 3.3.1, the sequence (x_n) has a convergent subsequence (x_{n_k}), and by Theorem 2.5.2, we know (x_{n_k}) converges to the same limit x. Finally, Definition 3.3.1 requires that $x \in K$. This proves that K is closed.

The proof of the converse statement is requested in Exercise 3.3.3. □

There may be a temptation to consider closed intervals as being a kind of standard archetype for compact sets, but this is misleading. The structure of compact sets can be much more intricate and interesting. For instance, one implication of Theorem 3.3.4 is that the Cantor set is compact. It is more useful to think of compact sets as generalizations of closed intervals. Whenever a fact involving closed intervals is true, it is often the case that the same result holds when we replace "closed interval" with "compact set." As an example, let's experiment with the Nested Interval Property proved in the first chapter.

Theorem 3.3.5 (Nested Compact Set Property). *If*

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq K_4 \supseteq \cdots$$

is a nested sequence of nonempty compact sets, then the intersection $\bigcap_{n=1}^{\infty} K_n$ *is not empty.*

Proof. In order to take advantage of the compactness of each K_n, we are going to produce a sequence that is eventually in each of these sets. Thus, for each $n \in \mathbf{N}$, pick a point $x_n \in K_n$. Because the compact sets are nested, it follows that the sequence (x_n) is contained in K_1. By Definition 3.3.1, (x_n) has a convergent subsequence (x_{n_k}) whose limit $x = \lim x_{n_k}$ is an element of K_1.

In fact, x is an element of *every* K_n for essentially the same reason. Given a particular $n_0 \in \mathbf{N}$, the terms in the sequence (x_n) are contained in K_{n_0} as long as $n \geq n_0$. Ignoring the finite number of terms for which $n_k < n_0$, the same subsequence (x_{n_k}) is then also contained in K_{n_0}. The conclusion is that $x = \lim x_{n_k}$ is an element of K_{n_0}. Because n_0 was arbitrary, $x \in \bigcap_{n=1}^{\infty} K_n$ and the proof is complete. □

Open Covers

Defining compactness for sets in $\mathbf{R}$ is reminiscent of the situation we encountered with completeness in that there are a number of equivalent ways to describe this phenomenon. We demonstrated the equivalence of two such characterizations in Theorem 3.3.4. What this theorem implies is that we could have decided to *define* compact sets to be sets that are closed and bounded, and then *proved* that sequences contained in compact sets have convergent subsequences with limits in the set. There are some larger issues involved in deciding what the definition should be, but what is important at this moment is that we be versatile enough to use whatever description of compactness is most appropriate for a given situation.

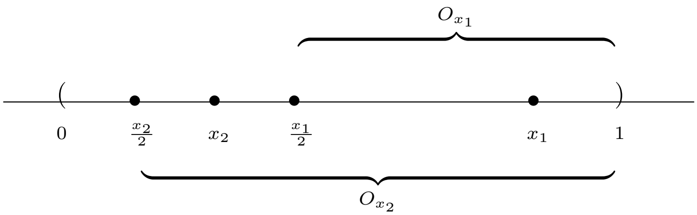
Although Theorem 3.3.4 is sufficient for most of our purposes, there is a third important characterization of compactness, equivalent to the two others, which is described in terms of *open covers* and *finite subcovers*.

Definition 3.3.6. Let $A \subseteq \mathbf{R}$. An *open cover* for A is a (possibly infinite) collection of open sets $\{O_\lambda : \lambda \in \Lambda\}$ whose union contains the set A; that is, $A \subseteq \bigcup_{\lambda\in\Lambda} O_\lambda$. Given an open cover for A, a *finite subcover* is a finite subcollection of open sets from the original open cover whose union still manages to completely contain A.

Example 3.3.7. Consider the open interval $(0,1)$. For each point $x \in (0,1)$, let O_x be the open interval $(x/2, 1)$. Taken together, the infinite collection $\{O_x : x \in (0,1)\}$ forms an open cover for the open interval $(0,1)$. Notice, however, that it is impossible to find a finite subcover. Given any proposed finite subcollection

$$\{O_{x_1}, O_{x_2}, \ldots, O_{x_n}\},$$

set $x' = \min\{x_1, x_2, \ldots, x_n\}$ and observe that any real number y satisfying $0 < y \le x'/2$ is not contained in the union $\bigcup_{i=1}^n O_{x_i}$.



Now, consider a similar cover for the closed interval $[0,1]$. For $x \in (0,1)$, the sets $O_x = (x/2, 1)$ do a fine job covering $(0,1)$, but in order to have an open cover of the closed interval $[0,1]$, we must also cover the endpoints. To remedy this, we could fix $\epsilon > 0$, and let $O_0 = (-\epsilon, \epsilon)$ and $O_1 = (1-\epsilon, 1+\epsilon)$. Then, the collection

$$\{O_0, O_1, O_x : x \in (0,1)\}$$

is an open cover for $[0,1]$. But this time, notice there is a finite subcover. Because of the addition of the set O_0, we can choose x' so that $x'/2 < \epsilon$. It follows that $\{O_0, O_{x'}, O_1\}$ is a finite subcover for the closed interval $[0,1]$.

Theorem 3.3.8 (Heine–Borel Theorem). *Let K be a subset of $\mathbf{R}$. All of the following statements are equivalent in the sense that any one of them implies the two others:*

(i) *K is compact.*

(ii) *K is closed and bounded.*

(iii) *Every open cover for K has a finite subcover.*

Proof. The equivalence of (i) and (ii) is the content of Theorem 3.3.4. What remains is to show that (iii) is equivalent to (i) and (ii). Let's first assume (iii), and prove that it implies (ii) (and thus (i) as well).

To show that K is bounded, we construct an open cover for K by defining O_x to be an open interval of radius 1 around each point $x \in K$. In the language of neighborhoods, $O_x = V_1(x)$. The open cover $\{O_x : x \in K\}$ then must have a finite subcover $\{O_{x_1}, O_{x_2}, \ldots, O_{x_n}\}$. Because K is contained in a finite union of bounded sets, K must itself be bounded.

The proof that K is closed is more delicate, and we argue it by contradiction. Let (y_n) be a Cauchy sequence contained in K with $\lim y_n = y$. To show that K is closed, we must demonstrate that $y \in K$, so assume for contradiction that this is not the case. If $y \notin K$, then every $x \in K$ is some positive distance away from y. We now construct an open cover by taking O_x to be an interval of radius $|x-y|/2$ around each point x in K. Because we are assuming (iii), the resulting open cover $\{O_x : x \in K\}$ must have a finite subcover $\{O_{x_1}, O_{x_2}, \ldots, O_{x_n}\}$. The contradiction arises when we realize that, in the spirit of Example 3.3.7, this finite subcover cannot contain all of the elements of the sequence (y_n). To make this explicit, set

$$\epsilon_0 = \min\left\{\frac{|x_i - y|}{2} : 1 \leq i \leq n\right\}.$$

Because $(y_n) \to y$, we can certainly find a term y_N satisfying $|y_N - y| < \epsilon_0$. But such a y_N must necessarily be excluded from each O_{x_i}, meaning that

$$y_N \notin \bigcup_{i=1}^{n} O_{x_i}.$$

Thus our supposed subcover does not actually cover all of K. This contradiction implies that $y \in K$, and hence K is closed and bounded.

The proof that (ii) implies (iii) is outlined in Exercise 3.3.9. To be historically accurate, it is this particular implication that is most appropriately referred to as the Heine–Borel Theorem. □

Exercises

Exercise 3.3.1. Show that if K is compact and nonempty, then $\sup K$ and $\inf K$ both exist and are elements of K.

Exercise 3.3.2. Decide which of the following sets are compact. For those that are not compact, show how Definition 3.3.1 breaks down. In other words, give an example of a sequence contained in the given set that does not possess a subsequence converging to a limit in the set.

(a) $\mathbf{N}$.

(b) $\mathbf{Q} \cap [0,1]$.

(c) The Cantor set.

(d) $\{1 + 1/2^2 + 1/3^2 + \cdots + 1/n^2 : n \in N\}$.

(e) $\{1, 1/2, 2/3, 3/4, 4/5, \ldots\}$.

Exercise 3.3.3. Prove the converse of Theorem 3.3.4 by showing that if a set $K \subseteq \mathbf{R}$ is closed and bounded, then it is compact.

Exercise 3.3.4. Assume K is compact and F is closed. Decide if the following sets are definitely compact, definitely closed, both, or neither.

(a) $K \cap F$

(b) $\overline{F^c \cup K^c}$

(c) $K \backslash F = \{x \in K : x \notin F\}$

(d) $\overline{K \cap F^c}$

Exercise 3.3.5. Decide whether the following propositions are true or false. If the claim is valid, supply a short proof, and if the claim is false, provide a counterexample.

(a) The arbitrary intersection of compact sets is compact.

(b) The arbitrary union of compact sets is compact.

(c) Let A be arbitrary, and let K be compact. Then, the intersection $A \cap K$ is compact.

(d) If $F_1 \supseteq F_2 \supseteq F_3 \supseteq F_4 \supseteq \cdots$ is a nested sequence of nonempty closed sets, then the intersection $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$.

Exercise 3.3.6. This exercise is meant to illustrate the point made in the opening paragraph to Section 3.3. Verify that the following three statements are true if every blank is filled in with the word "finite." Which are true if every blank is filled in with the word "compact." Which are true if every blank is filled in with the word "closed."

(a) Every ______ set has a maximum.

(b) If A and B are ______, then $A+B = \{a+b : a \in A, b \in B\}$ is also ______.

(c) If $\{A_n : n \in \mathbf{N}\}$ is a collection of ______ sets with the property that every finite subcollection has a nonempty intersection, then $\bigcap_{n=1}^{\infty} A_n$ is nonempty as well.

Exercise 3.3.7. As some more evidence of the surprising nature of the Cantor set, follow these steps to show that the sum $C+C = \{x+y : x, y \in C\}$ is equal to the closed interval $[0,2]$. (Keep in mind that C has zero length and contains no intervals.)

Because $C \subseteq [0,1]$, $C + C \subseteq [0,2]$, so we only need to prove the reverse inclusion $[0,2] \subseteq \{x+y : x, y \in C\}$. Thus, given $s \in [0,2]$, we must find two elements $x, y \in C$ satisfying $x + y = s$.

(a) Show that there exist $x_1, y_1 \in C_1$ for which $x_1 + y_1 = s$. Show in general that, for an arbitrary $n \in \mathbf{N}$, we can always find $x_n, y_n \in C_n$ for which $x_n + y_n = s$.

(b) Keeping in mind that the sequences (x_n) and (y_n) do not necessarily converge, show how they can nevertheless be used to produce the desired x and y in C satisfying $x + y = s$.

Exercise 3.3.8. Let K and L be nonempty compact sets, and define

$$d = \inf\{|x - y| : x \in K \text{ and } y \in L\}.$$

This turns out to be a reasonable definition for the *distance* between K and L.

(a) If K and L are disjoint, show $d > 0$ and that $d = |x_0 - y_0|$ for some $x_0 \in K$ and $y_0 \in L$.

(b) Show that it's possible to have $d = 0$ if we assume only that the disjoint sets K and L are closed.

Exercise 3.3.9. Follow these steps to prove the final implication in Theorem 3.3.8.

Assume K satisfies (i) and (ii), and let $\{O_\lambda : \lambda \in \Lambda\}$ be an open cover for K. For contradiction, let's assume that no finite subcover exists. Let I_0 be a closed interval containing K.

(a) Show that there exists a nested sequence of closed intervals $I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots$ with the property that, for each n, $I_n \cap K$ cannot be finitely covered and $\lim |I_n| = 0$.

(b) Argue that there exists an $x \in K$ such that $x \in I_n$ for all n.

(c) Because $x \in K$, there must exist an open set O_{λ_0} from the original collection that contains x as an element. Explain how this leads to the desired contradiction.

Exercise 3.3.10. Here is an alternate proof to the one given in Exercise 3.3.9 for the final implication in the Heine–Borel Theorem.

Consider the special case where K is a closed interval. Let $\{O_\lambda : \lambda \in \Lambda\}$ be an open cover for $[a, b]$ and define S to be the set of all $x \in [a, b]$ such that $[a, x]$ has a finite subcover from $\{O_\lambda : \lambda \in \Lambda\}$.

(a) Argue that S is nonempty and bounded, and thus $s = \sup S$ exists.

(b) Now show $s = b$, which implies $[a, b]$ has a finite subcover.

(c) Finally, prove the theorem for an arbitrary closed and bounded set K.

Exercise 3.3.11. Consider each of the sets listed in Exercise 3.3.2. For each one that is not compact, find an open cover for which there is no finite subcover.

Exercise 3.3.12. Using the concept of open covers (and explicitly avoiding the Bolzano–Weierstrass Theorem), prove that every bounded infinite set has a limit point.

Exercise 3.3.13. Let's call a set *clompact* if it has the property that every *closed* cover (i.e., a cover consisting of closed sets) admits a finite subcover. Describe all of the clompact subsets of $\mathbf{R}$.

3.4 Perfect Sets and Connected Sets

One of the underlying goals of topology is to strip away all of the extraneous information that comes with our intuitive picture of the real numbers and isolate just those properties that are responsible for the phenomenon we are studying. For example, we were quick to observe that any closed interval is a compact set. The content of Theorem 3.3.4, however, is that the compactness of a closed interval has nothing to do with the fact that the set is an *interval* but is a consequence of the set being bounded and closed. In Chapter 1, we argued that the set of real numbers between 0 and 1 is an uncountable set. This turns out to be the case for any nonempty closed set that does not contain isolated points.

Perfect Sets

Definition 3.4.1. A set $P \subseteq \mathbf{R}$ is *perfect* if it is closed and contains no isolated points.

Closed intervals (other than the singleton sets $[a, a]$) serve as the most obvious class of perfect sets, but there are more interesting examples.

Example 3.4.2 (Cantor Set). It is not too hard to see that the Cantor set is perfect. In Section 3.1, we defined the Cantor set as the intersection

$$C = \bigcap_{n=0}^{\infty} C_n,$$

where each C_n is a finite union of closed intervals. By Theorem 3.2.14, each C_n is closed, and by the same theorem, C is closed as well. It remains to show that no point in C is isolated.

Let $x \in C$ be arbitrary. To convince ourselves that x is not isolated, we must construct a sequence (x_n) of points in C, different from x, that converges to x. From our earlier discussion, we know that C at least contains the endpoints of the intervals that make up each C_n. In Exercise 3.4.3, we sketch the argument that these are all that is needed to construct (x_n).

One argument for the uncountability of the Cantor set was presented in Section 3.1. Another, perhaps more satisfying, argument for the same conclusion can be obtained from the next theorem.

Theorem 3.4.3. *A nonempty perfect set is uncountable.*

Proof. If P is perfect and nonempty, then it must be infinite because otherwise it would consist only of isolated points. Let's assume, for contradiction, that P is countable. Thus, we can write

$$P = \{x_1, x_2, x_3, \ldots\},$$

where *every* element of P appears on this list. The idea is to construct a sequence of nested compact sets K_n, all contained in P, with the property that

$x_1 \notin K_2$, $x_2 \notin K_3$, $x_3 \notin K_4$, ... Some care must be taken to ensure that each K_n is nonempty, for then we can use Theorem 3.3.5 to produce an

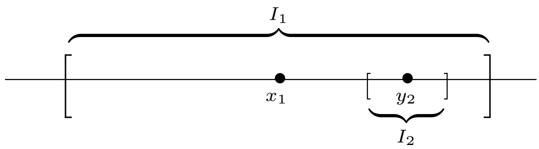
$$x \in \bigcap_{n=1}^{\infty} K_n \subseteq P$$

that *cannot* be on the list $\{x_1, x_2, x_3, \ldots\}$.

Let I_1 be a closed interval that contains x_1 in its interior (i.e., x_1 is not an endpoint of I_1). Now, x_1 is not isolated, so there exists some other point $y_2 \in P$ that is also in the interior of I_1. Construct a closed interval I_2, centered on y_2, so that $I_2 \subseteq I_1$ but $x_1 \notin I_2$. More explicitly, if $I_1 = [a, b]$, let

$$\epsilon = \min\{y_2 - a, b - y_2, |x_1 - y_2|\}.$$

Then, the interval $I_2 = [y_2 - \epsilon/2, y_2 + \epsilon/2]$ has the desired properties.



This process can be continued. Because $y_2 \in P$ is not isolated, there must exist another point $y_3 \in P$ in the interior of I_2, and we may insist that $y_3 \neq x_2$. Now, construct I_3 centered on y_3 and small enough so that $x_2 \notin I_3$ and $I_3 \subseteq I_2$. Observe that $I_3 \cap P \neq \emptyset$ because this intersection contains at least y_3.

If we carry out this construction inductively, the result is a sequence of closed intervals I_n satisfying

(i) $I_{n+1} \subseteq I_n$,

(ii) $x_n \notin I_{n+1}$, and

(iii) $I_n \cap P \neq \emptyset$.

To finish the proof, we let $K_n = I_n \cap P$. For each $n \in \mathbf{N}$, we have that K_n is closed because it is the intersection of closed sets, and bounded because it is contained in the bounded set I_n. Hence, K_n is compact. By construction, K_n is not empty and $K_{n+1} \subseteq K_n$. Thus, we can employ the Nested Compact Set Property (Theorem 3.3.5) to conclude that the intersection

$$\bigcap_{n=1}^{\infty} K_n \neq \emptyset.$$

But each K_n is a subset of P, and the fact that $x_n \notin I_{n+1}$ leads to the conclusion that $\bigcap_{n=1}^{\infty} K_n = \emptyset$, which is the sought-after contradiction. $\square$

Connected Sets

Although the two open intervals $(1,2)$ and $(2,5)$ have the limit point $x=2$ in common, there is still some space between them in the sense that no limit point of one of these intervals is actually contained in the other. Said another way, the closure of $(1,2)$ (see Definition 3.2.11) is disjoint from $(2,5)$, and the closure of $(2,5)$ does not intersect $(1,2)$. Notice that this same observation cannot be made about $(1,2]$ and $(2,5)$, even though these latter sets are disjoint.

Definition 3.4.4. Two nonempty sets $A, B \subseteq \mathbf{R}$ are *separated* if $\overline{A} \cap B$ and $A \cap \overline{B}$ are both empty. A set $E \subseteq \mathbf{R}$ is *disconnected* if it can be written as $E = A \cup B$, where A and B are nonempty separated sets.

A set that is not disconnected is called a *connected* set.

Example 3.4.5. (i) If we let $A=(1,2)$ and $B=(2,5)$, then it is not difficult to verify that $E=(1,2)\cup(2,5)$ is disconnected. Notice that the sets $C=(1,2]$ and $D=(2,5)$ are not separated because $C \cap \overline{D} = \{2\}$ is not empty. This should be comforting. The union $C \cup D$ is equal to the interval $(1,5)$, which better not qualify as a disconnected set. We will prove in a moment that every interval is a connected subset of **R** and vice versa.

(ii) Let's show that the set of rational numbers is disconnected. If we let

$$A = \mathbf{Q} \cap (-\infty, \sqrt{2}) \quad \text{and} \quad B = \mathbf{Q} \cap (\sqrt{2}, \infty),$$

then we certainly have $\mathbf{Q} = A \cup B$. The fact that $A \subseteq (-\infty, \sqrt{2})$ implies (by the Order Limit Theorem) that any limit point of A will necessarily fall in $(-\infty, \sqrt{2}]$. Because this is disjoint from B, we get $\overline{A} \cap B = \emptyset$. We can similarly show that $A \cap \overline{B} = \emptyset$, which implies that A and B are separated.

The definition of connected is stated as the negation of disconnected, but a little care with the logical negation of the quantifiers in Definition 3.4.4 results in a positive characterization of connectedness. Essentially, a set E is connected if, no matter how it is partitioned into two nonempty disjoint sets, it is always possible to show that at least one of the sets contains a limit point of the other.

Theorem 3.4.6. *A set $E \subseteq \mathbf{R}$ is connected if and only if, for all nonempty disjoint sets A and B satisfying $E = A \cup B$, there always exists a convergent sequence $(x_n) \to x$ with (x_n) contained in one of A or B, and x an element of the other.*

Proof. Exercise 3.4.6. □

The concept of connectedness is more relevant when working with subsets of the plane and other higher-dimensional spaces. This is because, in **R**, the connected sets coincide precisely with the collection of intervals (with the understanding that unbounded intervals such as $(-\infty, 3)$ and $[0,\infty)$ are included).

Theorem 3.4.7. *A set $E \subseteq \mathbf{R}$ is connected if and only if whenever $a < c < b$ with $a, b \in E$, it follows that $c \in E$ as well.*

Proof. Assume E is connected, and let $a, b \in E$ and $a < c < b$. Set

$$A = (-\infty, c) \cap E \quad \text{and} \quad B = (c, \infty) \cap E.$$

Because $a \in A$ and $b \in B$, neither set is empty and, just as in Example 3.4.5 (ii), neither set contains a limit point of the other. If $E = A \cup B$, then we would have that E is disconnected, which it is not. It must then be that $A \cup B$ is missing some element of E, and c is the only possibility. Thus, $c \in E$.

Conversely, assume that E is an interval in the sense that whenever $a, b \in E$ satisfy $a < c < b$ for some c, then $c \in E$. Our intent is to use the characterization of connected sets in Theorem 3.4.6, so let $E = A \cup B$, where A and B are nonempty and disjoint. We need to show that one of these sets contains a limit point of the other. Pick $a_0 \in A$ and $b_0 \in B$, and, for the sake of the argument, assume $a_0 < b_0$. Because E is itself an interval, the interval $I_0 = [a_0, b_0]$ is contained in E. Now, bisect I_0 into two equal halves. The midpoint of I_0 must either be in A or B, and so choose $I_1 = [a_1, b_1]$ to be the half that allows us to have $a_1 \in A$ and $b_1 \in B$. Continuing this process yields a sequence of nested intervals $I_n = [a_n, b_n]$, where $a_n \in A$, $b_n \in B$, and the length $(b_n - a_n) \to 0$. The remainder of this argument should feel familiar. By the Nested Interval Property, there exists an

$$x \in \bigcap_{n=0}^{\infty} I_n,$$

and it is straightforward to show that the sequences of endpoints each satisfy $\lim a_n = x$ and $\lim b_n = x$. But now $x \in E$ must belong to either A or B, thus making it a limit point of the other. This completes the argument. $\square$

Exercises

Exercise 3.4.1. If P is a perfect set and K is compact, is the intersection $P \cap K$ always compact? Always perfect?

Exercise 3.4.2. Does there exist a perfect set consisting of only rational numbers?

Exercise 3.4.3. Review the portion of the proof given for Theorem 3.4.2 and follow these steps to complete the argument.

(a) Because $x \in C_1$, argue that there exists an $x_1 \in C \cap C_1$ with $x_1 \neq x$ satisfying $|x - x_1| \leq 1/3$.

(b) Finish the proof by showing that for each $n \in \mathbf{N}$, there exists $x_n \in C \cap C_n$, different from x, satisfying $|x - x_n| \leq 1/3^n$.

Exercise 3.4.4. Repeat the Cantor construction from Section 3.1 starting with the interval $[0, 1]$. This time, however, remove the open middle *fourth* from each component.

(a) Is the resulting set compact? Perfect?

(b) Using the algorithms from Section 3.1, compute the length and dimension of this Cantor-like set.

Exercise 3.4.5. Let A and B be nonempty subsets of $\mathbf{R}$. Show that if there exist disjoint open sets U and V with $A \subseteq U$ and $B \subseteq V$, then A and B are separated.

Exercise 3.4.6. Prove Theorem 3.4.6.

Exercise 3.4.7. A set E is *totally disconnected* if, given any two distinct points $x, y \in E$, there exist separated sets A and B with $x \in A$, $y \in B$, and $E = A \cup B$.

(a) Show that $\mathbf{Q}$ is totally disconnected.

(b) Is the set of irrational numbers totally disconnected?

Exercise 3.4.8. Follow these steps to show that the Cantor set is totally disconnected in the sense described in Exercise 3.4.7.

Let $C = \bigcap_{n=0}^{\infty} C_n$, as defined in Section 3.1.

(a) Given $x, y \in C$, with $x < y$, set $\epsilon = y - x$. For each $n = 0, 1, 2, \ldots$, the set C_n consists of a finite number of closed intervals. Explain why there must exist an N large enough so that it is impossible for x and y both to belong to the same closed interval of C_N.

(b) Show that C is totally disconnected.

Exercise 3.4.9. Let $\{r_1, r_2, r_3, \ldots\}$ be an enumeration of the rational numbers, and for each $n \in \mathbf{N}$ set $\epsilon_n = 1/2^n$. Define $O = \bigcup_{n=1}^{\infty} V_{\epsilon_n}(r_n)$, and let $F = O^c$.

(a) Argue that F is a closed, nonempty set consisting only of irrational numbers.

(b) Does F contain any nonempty open intervals? Is F totally disconnected? (See Exercise 3.4.7 for the definition.)

(c) Is it possible to know whether F is perfect? If not, can we modify this construction to produce a nonempty perfect set of irrational numbers?

3.5 Baire's Theorem

The nature of the real line can be deceptively elusive. The closer we look, the more intricate and enigmatic $\mathbf{R}$ becomes, and the more we are reminded to proceed carefully (i.e., axiomatically) with all of our conclusions about properties of subsets of $\mathbf{R}$. The structure of open sets is fairly straightforward. Every open set is either a finite or countable union of open intervals. Standing in opposition to this tidy description of all open sets is the Cantor set. The Cantor set is a closed, uncountable set that contains no intervals of any kind. Thus, no such characterization of closed sets should be anticipated.

Recall that the arbitrary union of open sets is always an open set. Likewise, the arbitrary intersection of closed sets is closed. By taking unions of closed sets or intersections of open sets, however, it is possible to obtain a new selection of subsets of $\mathbf{R}$.

Definition 3.5.1. A set $A \subseteq \mathbf{R}$ is called an F_σ *set* if it can be written as the countable union of closed sets. A set $B \subseteq \mathbf{R}$ is called a G_δ *set* if it can be written as the countable intersection of open sets.

Exercise 3.5.1. Argue that a set A is a G_δ set if and only if its complement is an F_σ set.

Exercise 3.5.2. Replace each ______ with the word *finite* or *countable*, depending on which is more appropriate.

(a) The ______ union of F_σ sets is an F_σ set.

(b) The ______ intersection of F_σ sets is an F_σ set.

(c) The ______ union of G_δ sets is a G_δ set.

(d) The ______ intersection of G_δ sets is a G_δ set.

Exercise 3.5.3. (This exercise has already appeared as Exercise 3.2.15.)

(a) Show that a closed interval $[a, b]$ is a G_δ set.

(b) Show that the half-open interval $(a, b]$ is both a G_δ and an F_σ set.

(c) Show that $\mathbf{Q}$ is an F_σ set, and the set of irrationals $\mathbf{I}$ forms a G_δ set.

It is not readily obvious that the class F_σ does not include every subset of $\mathbf{R}$, but we are now ready to argue that $\mathbf{I}$ is not an F_σ set (and consequently $\mathbf{Q}$ is not a G_δ set). This will follow from a theorem due to René Louis Baire (1874–1932).

Recall that a set $G \subseteq \mathbf{R}$ is *dense* in $\mathbf{R}$ if, given any two real numbers $a < b$, it is possible to find a point $x \in G$ with $a < x < b$.

Theorem 3.5.2. *If $\{G_1, G_2, G_3, \ldots\}$ is a countable collection of dense, open sets, then the intersection $\bigcap_{n=1}^{\infty} G_n$ is not empty.*

Proof. Before embarking on the proof, notice that we have seen a conclusion like this before. Theorem 3.3.5 asserts that a nested sequence of compact sets has a nontrivial intersection. In this theorem, we are dealing with dense, open sets, but as it turns out, we are going to use Theorem 3.3.5—and actually, just the Nested Interval Property—as the crucial step in the argument.

Exercise 3.5.4. Starting with $n = 1$, inductively construct a nested sequence of *closed* intervals $I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots$ satisfying $I_n \subseteq G_n$. Give special attention to the issue of the endpoints of each I_n. Show how this leads to a proof of the theorem. □

Exercise 3.5.5. Show that it is impossible to write

$$\mathbf{R} = \bigcup_{n=1}^{\infty} F_n,$$

where for each $n \in \mathbf{N}$, F_n is a closed set containing no nonempty open intervals.

Exercise 3.5.6. Show how the previous exercise implies that the set $\mathbf{I}$ of irrationals cannot be an F_σ set, and $\mathbf{Q}$ cannot be a G_δ set.

Exercise 3.5.7. Using Exercise 3.5.6 and versions of the statements in Exercise 3.5.2, construct a set that is neither in F_σ nor in G_δ.

Nowhere-Dense Sets

We have encountered several equivalent ways to assert that a particular set G is dense in $\mathbf{R}$. In Section 3.2, we observed that G is dense in $\mathbf{R}$ if and only if every point of $\mathbf{R}$ is a limit point of G. Because the closure of any set is obtained by taking the union of the set and its limit points, we have that

$$G \text{ is dense in } \mathbf{R} \text{ if and only if } \overline{G} = \mathbf{R}.$$

The set $\mathbf{Q}$ is dense in $\mathbf{R}$; the set $\mathbf{Z}$ is clearly not. In fact, in the jargon of analysis, $\mathbf{Z}$ is nowhere-dense in $\mathbf{R}$.

Definition 3.5.3. A set E is *nowhere-dense* if $\overline{E}$ contains no nonempty open intervals.

Exercise 3.5.8. Show that a set E is nowhere-dense in $\mathbf{R}$ if and only if the complement of $\overline{E}$ is dense in $\mathbf{R}$.

Exercise 3.5.9. Decide whether the following sets are dense in $\mathbf{R}$, nowhere-dense in $\mathbf{R}$, or somewhere in between.

(a) $A = \mathbf{Q} \cap [0, 5]$.

(b) $B = \{1/n : n \in \mathbf{N}\}$.

(c) the set of irrationals.

(d) the Cantor set.

We can now restate Theorem 3.5.2 in a slightly more general form.

Theorem 3.5.4 (Baire's Theorem). *The set of real numbers* $\mathbf{R}$ *cannot be written as the countable union of nowhere-dense sets.*

Proof. For contradiction, assume that $E_1, E_2, E_3, \ldots$ are each nowhere-dense and satisfy $\mathbf{R} = \bigcup_{n=1}^{\infty} E_n$.

Exercise 3.5.10. Finish the proof by finding a contradiction to the results in this section. □

3.6 Epilogue

Baire's Theorem is yet another statement about the size of **R**. We have already encountered several ways to describe the sizes of infinite sets. In terms of cardinality, countable sets are relatively small whereas uncountable sets are large. We also briefly discussed the concept of "length," or "measure," in Section 3.1. Baire's Theorem offers a third perspective. From this point of view, nowhere-dense sets are considered to be "thin" sets. Any set that is the countable union—i.e., a not very large union—of these small sets is called a "meager" set or a set of "first category." A set that is not of first category is of "second category." Intuitively, sets of the second category are the "fat" subsets. The Baire Category Theorem, as it is often called, states that **R** is of second category.

There is a significance to the Baire Category Theorem that is difficult to appreciate at the moment because we are only seeing a special case of this result. The real numbers are an example of a *complete metric space*. Metric spaces are discussed in some detail in Section 8.2, but here is the basic idea. Given a set of mathematical objects such as real numbers, points in the plane or continuous functions defined on [0,1], a "metric" is a rule that assigns a "distance" between two elements in the set. In **R**, we have been using $|x-y|$ as the distance between the real numbers x and y. The point is that if we can create a satisfactory notion of "distance" on these other spaces (we will need the triangle inequality to hold, for instance), then the concepts of convergence, Cauchy sequences, and open sets, for example, can be naturally transferred over. A complete metric space is any set with a suitably defined metric in which Cauchy sequences have limits. We have spent a good deal of time discussing the fact that **R** is a complete metric space whereas **Q** is not.

The Baire Category Theorem in its more general form states that *any* complete metric space must be too large to be the countable union of nowhere-dense subsets. One particularly interesting example of a complete metric space is the set of continuous functions defined on the interval $[0, 1]$. (The distance between two functions f and g in this space is defined to be $\sup|f(x) - g(x)|$, where $x \in [0, 1]$.) Now, in this space we will see that the collection of continuous functions that are differentiable at even one point *can* be written as the countable union of nowhere-dense sets. Thus, a fascinating consequence of Baire's Theorem in this setting is that *most continuous functions do not have derivatives at any point.* Chapter 5 concludes with a construction of one such function. This odd situation mirrors the roles of **Q** and **I** as subsets of **R**. Just as the familiar rational numbers constitute a minute proportion of the real line, the differentiable functions of calculus are exceedingly atypical of continuous functions in general.