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

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

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:

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

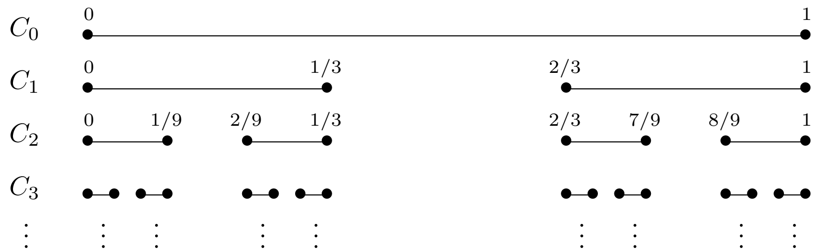

Figure 3.1: Defining the Cantor set; $C = \bigcap_{n=0}^{\infty} C_n$.

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 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 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

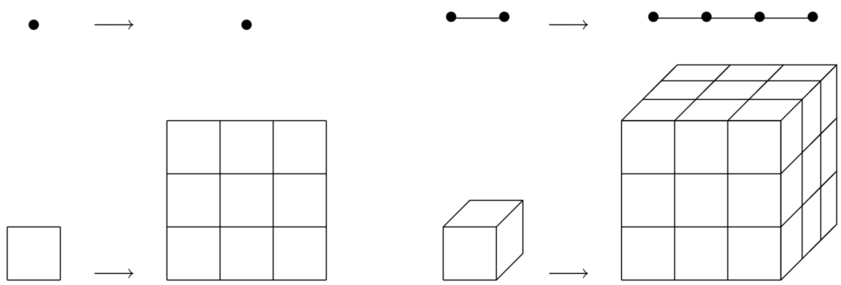

Figure 3.2: Magnifying sets by a factor of 3.

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 1.5.4), 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 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.

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 $x = \ln 2 / \ln 3 \approx .631$ (Fig. 3.3).

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

Figure 3.3: Dimension of C; $2 = 3^x \Rightarrow x = \ln 2/\ln 3$.

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.