Math 4111
October 28, 2020 Lecture

Steven G. Krantz



Figure: This is your instructor.

4.4 The Cantor Set

In this section we describe the construction of a remarkable subset of $\mathbb{R}$ with many pathological properties. It only begins to suggest the richness of the structure of the real number system.

We begin with the unit interval $S_0 = [0, 1]$. We extract from S_0 its open middle third; thus $S_1 = S_0 \setminus (1/3, 2/3)$. Observe that S_1 consists of two closed intervals of equal length $1/3$. See the figure:

Figure: Construction of the Cantor set.

Now we construct S_2 from S_1 by extracting from each of its two intervals the middle third:
$S_2 = [0, 1/9] \cup [2/9, 1/3] \cup [6/9, 7/9] \cup [8/9, 1]$. The figure below shows S_2.



Figure: Second step in the construction of the Cantor set.

Continuing in this fashion, we construct S_{j+1} from S_j by extracting the middle third from each of its component subintervals. We define the Cantor set C to be

$$C = \bigcap_{j=1}^{\infty} S_j \, .$$

Notice that each of the sets S_j is closed and bounded, hence compact. By Proposition 4.39 of the last section, C is therefore not empty (one can also note that the endpoints of the removed intervals are all in the Cantor set). The set C is closed and bounded, hence compact.

Proposition

The Cantor set C has zero length, in the sense that the complementary set $[0,1] \setminus C$ has length 1.

Proof: In the construction of S_1, we removed from the unit interval one interval of length 3^{-1}. In constructing S_2, we further removed two intervals of length 3^{-2}. In constructing S_j, we removed 2^{j-1} intervals of length 3^{-j}. Thus the total length of the intervals removed from the unit interval is

$$\sum_{j=1}^{\infty} 2^{j-1} \cdot 3^{-j} = \frac{1}{3} \sum_{j=0}^{\infty} \left(\frac{2}{3}\right)^j = \frac{1}{3} \left(\frac{1}{1-2/3}\right) = 1.$$

Thus the Cantor set has length zero because its complement in the unit interval has length one. □

Proposition

The Cantor set is uncountable.

Proof: We assign to each element of the Cantor set a "label" consisting of a sequence of 0s and 1s that identifies its location in the set.

Fix an element x in the Cantor set. Then x is in S_1. If x is in the left half of S_1, then the first digit in the "label" of x is 0; otherwise it is 1. See the figure below.



Figure: The first digit of the label of a point in the Cantor set.

Likewise $x \in S_2$. By the first part of this argument, it is either in the left half S_2^0 of S_2 (when the first digit in the label is 0) or the right half S_2^1 of S_2 (when the first digit of the label is 1). Whichever of these is correct, that half will consist of two intervals of length 3^{-2}. If x is in the leftmost of these two intervals then the second digit of the "label" of x is 0. Otherwise the second digit is 1. Refer to the figure below.

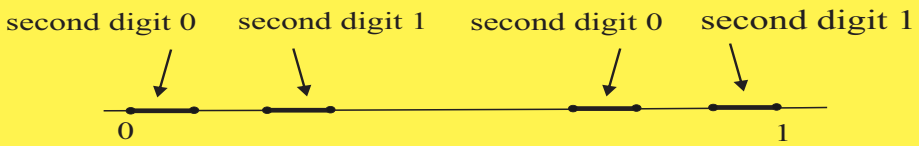


Figure: The second digit of the label of a point in the Cantor set.

Continuing in this fashion, we may assign to x an infinite sequence of 0s and 1s.

Conversely, if $a, b, c, \ldots$ is a sequence of 0s and 1s, then we may locate a unique corresponding element y of the Cantor set. If the first digit is a zero then y is in the left half of S_1; otherwise y is in the right half of S_1. Likewise the second digit locates y within S_2, and so forth.

Thus we have a one-to-one correspondence between the Cantor set and the collection of all infinite sequences of zeroes and ones. [Notice that we are in effect thinking of the point assigned to a sequence $c_1c_2c_3\ldots$ of 0s and 1s as the limit of the points assigned to $c_1, c_1c_2, c_1c_2c_3, \ldots$ Thus we are using the fact that C is closed.]

However, as we can learn in Appendix II at the end of the book, the set of all infinite sequences of zeroes and ones is uncountable. Thus we see that the Cantor set is uncountable. □

The Cantor set is quite thin (it has zero length) but it is large in the sense that it has uncountably many elements. Also it is compact. The next result reveals a surprising, and not generally well known, property of this "thin" set.

Theorem

Let C be the Cantor set and define

$$S = \{x + y : x \in C, y \in C\}.$$

Then $S = [0, 2]$.

Proof: We sketch the proof here and treat the details in the exercises.

Since $C \subseteq [0,1]$ it is clear that $S \subseteq [0,2]$. For the reverse inclusion, fix an element $t \in [0,2]$. Our job is to find two elements c and d in C such that $c + d = t$.
First observe that $\{x + y : x \in S_1, y \in S_1\} = [0,2]$. Therefore there exist $x_1 \in S_1$ and $y_1 \in S_1$ such that $x_1 + y_1 = t$.
Similarly, $\{x + y : x \in S_2, y \in S_2\} = [0,2]$. Therefore there exist $x_2 \in S_2$ and $y_2 \in S_2$ such that $x_2 + y_2 = t$.

Continuing in this fashion we may find for each j numbers x_j and y_j such that $x_j, y_j \in S_j$ and $x_j + y_j = t$. Of course $\{x_j\} \subseteq C$ and $\{y_j\} \subseteq C$ hence there are subsequences $\{x_{j_k}\}$ and $\{y_{j_k}\}$ which converge to real numbers c and d respectively. Since C is compact, we can be sure that $c \in C$ and $d \in C$. But the operation of addition respects limits, thus we may pass to the limit as $k \to \infty$ in the equation

$$x_{j_k} + y_{j_k} = t$$

to obtain

$$c + d = t\,.$$

Therefore $[0, 2] \subseteq \{x + y : x \in C\}$. This completes the proof. $\square$

In the exercises at the end of the section we shall explore constructions of other Cantor-like sets, some of which have zero length and some of which have positive length.

The Cantor set that we have discussed in detail in the present section is sometimes distinguished with the name "the Cantor ternary set." We shall also consider in the exercises other ways to construct the Cantor ternary set.

Observe that, whereas any open set in $\mathbb{R}$ is the countable or finite disjoint union of open intervals, the existence of the Cantor set shows us that there is no such structure theorem for closed sets. That is to say, we cannot hope to write an arbitrary closed set as the disjoint union of closed intervals.

However, de Morgan's Law shows that an arbitrary closed set can be written as the countable intersection of sets, each of which is the union of disjoint closed intervals.

In fact, closed intervals are atypically simple when considered as examples of closed sets.