A BRIEF INTRODUCTION TO THE CANTOR SET

JAMIE PEARCE

This write-up is intended to introduce the reader to the Cantor set. The Cantor set is defined as a subspace of $\mathbb{R}$. We begin with its definition.
Intuitively, the Cantor set is constructed by first taking the interval $I = [0, 1]$ inside $\mathbb{R}$ and then removing the open middle third; then removing the open middle thirds of each of the two remaining intervals; and continuing this process indefinitely. In other words, at the nth step, we remove the open middle thirds from each of the n intervals obtained from the $n-1$st step.

Definition 0.1. We will denote the Cantor set $C = \bigcap_{n=1}^{\infty} C_n$ where C_n are the subsets $C_1 = [0,1], C_2 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], C_3 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1], \cdots, C_n = \frac{C_{n-1}}{3} \cup (\frac{2}{3} + \frac{C_{n-1}}{3}), \cdots$

Here the notation $\frac{S}{n}$ for a subset S of $\mathbb{R}$ denotes the set $\{x \in \mathbb{R} : \exists s \in S \text{ such that } x = \frac{s}{n}$.

The Cantor set has many surprising properties as a topological space. It is, among other things, uncountable, compact, metrizable and totally disconnected. Because of these properties, the space serves as an interesting counterexample to many naive conjectures in topology. For instance, it is a counterexample to the following natural conjecture, whose converse is obviously true: if a topological space X is totally disconnected, then it has the discrete topology.

Definition 0.2. A space is totally disconnected if its only connected subspaces are one-point sets.

Theorem 0.3. *The Cantor set is totally disconnected, and it does not have the discrete topology.*

Proof. First, we notice that the subsets C_n are not connected; for each n, the set C_n is the union of disjoint closed intervals. The intervals that make up C_n each have length $\frac{1}{3^{n-1}}$. For $x, y \in C$ there exists an $N \in \mathbb{N}$ such that $\frac{1}{3^{N-1}} <| x - y |$. This means x and y must be in different closed intervals inside C_N. Therefore the only connected subspaces are one point sets. This shows that the Cantor set is totally disconnected.
To prove that the Cantor set does not have the discrete topology, we show that it has a limit point. Elements in C are of the form $\frac{2a_1}{3} + \frac{2a_2}{3^2} + \frac{2a_3}{3^3} + \frac{2a_4}{3^4} + ...$ where the a_i's are either 1 or 0. There exists a sequence of endpoints in C which converges to $\frac{1}{4}$: $\{x_n = \sum_{k=1}^{n} \frac{2}{3^{2k}}\}$. We have that $x_n = \sum_{k=1}^{n} \frac{2}{3^{2k}} = \frac{1 - \frac{1}{9}^n}{4}$ and $\frac{1}{9}^n \to 0$, so $\frac{1-\frac{1}{9}^n}{4} \to \frac{1}{4}$. Therefore $\frac{1}{4}$ s a limit point in C. We conclude that the Cantor set is totally disconnected and does not have the discrete topology. □

The Cantor set is also useful in its own right, aside from its role as a frequent counterexample. The set can be used to prove some surprising results, such as the existence of a space-filling curve in $[0,1]^2$. One would not anticipate that a curve, a clearly one dimensional object, could fill a higher dimensional space. By curve in a space X we mean a continuous function $[0,1] \to X$.

Definition 0.4. A space-filling curve is a curve whose range contains the entire 2-dimensional unit square.

Theorem 0.5. *There exists a surjective, continuous function* $F : [0,1] \to [0,1]^2$.

Proof. Let $h : C \to [0,1]$ be the surjective continuous function defined by $h(x) = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$ where $x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n} \in C$ for $a_n \in \{0,1\}$. From h, we get a continuous function H from $C \times C$ onto $[0,1] \times [0,1] = [0,1]^2$ defined as $H(x,y) = (h(x), h(y))$. Since the Cantor set is homeomorphic to the product $C \times C$, there is a continuous bijection g from the Cantor set onto $C \times C$. The composition f of H and g is a continuous function that maps the Cantor set onto the entire unit square.
From there we can extend f linearly to a continuous function F whose domain is $[0,1]$ by defining the extension part of F on (a,b) to be the line segment within $[0,1]^2$ connecting $f(a)$ and $f(b)$ where each (a,b) is a deleted open interval of the Cantor set. This proves the existence of a surjective, continuous function from the unit interval to the unit square. □