The Cantor Set

The Cantor set is a famous set first constructed by Georg Cantor in 1883. It is simply a subset of the interval $[0, 1]$, but the set has some very interesting properties. We will first describe how to construct this set, and then prove some interesting properties of the set.

Let $I = [0, 1]$.

Remove the open third segment $\left(\frac{1}{3}, \frac{2}{3}\right)$ and let

$$A_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Now remove the open third segments in each part. Let

$$A_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

Continue in this way always removing the middle third of each segment to get $A_3, A_4, \ldots$.

Note that $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. And for each $k \in \mathbb{N}$, A_k is the union of 2^k closed intervals, each of length 3^{-k}.

Let $C = \cap_{i=1}^{\infty} A_i$. Then C is the Cantor set.

Now we will prove some interesting properties of C.

1. C is compact.

   Proof: Each A_k is a finite union of closed sets, so A_k is closed for all k by Corollary 1(b). Then $C = \cap A_k$ is also closed by Corollary 1(a). Also, C is bounded since $C \subseteq [0, 1]$. So by the Heine-Borel Theorem, C is compact.

   □

2. Let $x = 0.a_1 a_2 a_3 \ldots$ be the base 3 expansion of a number $x \in [0, 1]$. Then $x \in C$ iff $a_n \in \{0, 2\}$ for all $n \in \mathbb{N}$.

   Proof: For a review of converting numbers to a different base, see http://www.mathpath.org/concepts/Num/frac.htm. The fast explanation of base 3 is that the decimal records which "third" the number is in. For example 0.120 is in the second third in A_1, and then a third third in A_2, and then a first third of A_3.

   Let $x \in [0, 1]$ and let $0.a_1 a_2 a_3 \ldots$ be its base 3 expansion. Assume there is some $k \in \mathbb{N}$ such that $a_k = 1$ in the expansion. Then $0.a_1 a_2 \ldots a_{k-1} \in A_{k-1}$ but $a_k = 1 \Longrightarrow x \notin A_k \Longrightarrow x \notin C$.

   On the other hand, by the definition of base 3 expansion, if $a_n \in \{0, 2\}$, for all $n \in \mathbb{N}$, then $x \in C$.

   □

3. C is uncountable.

   Proof: This is the same diagonalizing proof that we did for showing $\mathbb{R}$ is uncountable. Suppose C is countable, and list it's elements as $C = \{x_1, x_2, x_3, \ldots\}$. Now look at the base 3 expansion of each of those numbers. We can write

$$\begin{aligned} x_1 &= 0.a_{11}a_{12}a_{13}\ldots \\ x_2 &= 0.a_{21}a_{22}a_{23}\ldots \\ &\vdots \\ x_k &= 0.a_{k1}a_{k2}a_{k3}\ldots \end{aligned}$$

   Where $a_{ij} = 0$ or 2 for all i, j.

Let $y = 0.b_1b_2b_3\ldots$ where

$$b_i = \begin{cases} 0 & \text{if } a_{ii} = 2 \\ 2 & \text{if } a_{ii} = 0 \end{cases}$$

Then $y \neq x_1$ since $b_1 \neq a_{11}$, $y \neq x_2$ since $b_2 \neq a_{22}$, and so on. This implies that $y \notin C$, but this is a contradiction since $b_i \in \{0, 2\}$ for each i, and by the previous problem, $y \in C$. Therefore C is uncountable.

□

4. C contains no intervals.

   Proof: Let $(a, b) \subseteq [0, 1]$, and assume $a < b$. Let $M = \{n \in \mathbb{N} : -\log_3(b - a) < n\}$. Notice that $(a, b) \subseteq [0, 1]$, so $b - a < 1$ which implies that $-\log_3(b - a) > 0$, and $-\log_3(b - a) \in \mathbb{R}$, so by the Archimedean property, there is some $m \in M$ such that $m \leq k$ for all $k \in M$. So we have $-\log_3(b - a) < m$ which implies that $3^{-(-\log_3(b-a))} > 3^{-m}$. The left hand side of that inequality can be simplified as follows: $3^{-(-\log_3(b-a))} = 3^{\log_3(b-a)} = b - a = |b - a| < 3^{-m}$. But A_m is the union of subsets of $[0, 1]$ of length 3^{-m} which implies that $(b, a) \nsubseteq A_m$. Therefore $(b, a) \nsubseteq C$.

□

5. $\dfrac{1}{4} \in C$, but $\dfrac{1}{4}$ is not an endpoint of any of the intervals in any of the sets A_k for $k \in \mathbb{N}$.

   Proof: The base 3 decimal expansion of $\frac{1}{4}$ is $0.\overline{02}$. Thus by part (b), $\frac{1}{4} \in C$. Notice that $x \in A_k$ is an endpoint if $x = 0, x = 1$, or if $x = 3^{-k}$ for some $k \in \mathbb{N}$. Clearly $\frac{1}{4} \neq 0, 1$, and for all $k \in \mathbb{N}$, $\frac{1}{4} \neq 3^{-k}$. Therefore $\frac{1}{4}$ is not an endpoint.

□