The Cantor Set

Peter Borg

Consider the following sets, where C_1 is the real interval $[0, 1]$ without the middle $\frac{1}{3}$ of the interval, and C_k is constructed by removing $\frac{1}{3}$ of each real interval in the union of intervals in C_{k-1}.

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

$$C_2 = \left[0, \frac{1}{3^2}\right] \cup \left[\frac{2}{3^2}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{2}{3} + \frac{1}{3^2}\right] \cup \left[2\left(\frac{1}{3} + \frac{1}{3^2}\right), 1\right]$$

$$C_3 = \left[0, \frac{1}{3^3}\right] \cup \left[\frac{2}{3^3}, \frac{1}{3^2}\right] \cup \left[\frac{2}{3^2}, \frac{2}{3^2} + \frac{1}{3^3}\right] \cup \left[2\left(\frac{1}{3^2} + \frac{1}{3^3}\right), \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{2}{3} + \frac{1}{3^3}\right]$$

$$\cup \left[2\left(\frac{1}{3} + \frac{1}{3^3}\right), \frac{2}{3} + \frac{1}{3^2}\right] \cup \left[2\left(\frac{1}{3} + \frac{1}{3^2}\right), 2\left(\frac{1}{3} + \frac{1}{3^2}\right) + \frac{1}{3^3}\right] \cup \left[2\left(\frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3}\right), 1\right]$$

... etc.

It can be proved by induction on n that

$$C_n = \bigcup_{j=0}^{2^n - 1} I_j$$

where

$$I_j = \left[2\sum_{i=0}^{n-1} a_i \left(\frac{1}{3}\right)^i, 2\sum_{i=0}^{n-1} a_i \left(\frac{1}{3}\right)^i + \frac{1}{3^n}\right]$$

and

$$a_i = \begin{cases} 0 \text{ if } i \bmod 2^j = 0 \\ 1 \text{ if } i \bmod 2^j = 1 \end{cases}$$

The Cantor Set is:

$$C = \lim_{n \to \infty} C_n$$

Hence, in the limit, the intervals I_j become points of the form

$$2\sum_{i=0}^{\infty} a_i \left(\frac{1}{3}\right)^i$$

where a_i is 0 or 1.

$$\text{Hence } x \in C \iff x = 2\sum_{i=0}^{\infty} a_i \left(\frac{1}{3}\right)^i, \text{ where } a_i = 0 \text{ or } 1$$

Having established which points are in the Cantor set, we can now show that these points form an uncountable set. But first we shall show that C has measure 0, and we shall do this by considering the lengths (Lesbesgue measure) of all the disjoint intervals removed from $[0,1], C_1, C_2, \ldots$ and C_{k-1} to obtain C_k, and then let $k \to \infty$. To obtain C_1 an interval of length $\frac{1}{3}$ was removed, for C_2, $2(\frac{1}{3})^2$ was removed, for C_3, $2^2(\frac{1}{3})^3$ was removed and for C_k, $2^{k-1}(\frac{1}{3})^k$ was removed. The sum of all the lengths removed is

$$2^{-1}\sum_{i=1}^{k}\left(\frac{2}{3}\right)^i = 1-\left(\frac{2}{3}\right)^k \to 1 \text{ as } k \to \infty$$

Hence having removed a total length of 1 from $[0,1]$ we are left with a measure of 0 for C.

The binary representation for any real number in the interval $[0,1]$ is of the form

$$y = \sum_{i=1}^{\infty} a_i \left(\frac{1}{2}\right)^i$$

and moreover, since the real numbers in the interval in $[0,1]$ form an uncountable set and each have a binary representation, then the set B of such binary representations is uncountable.

Now if we construct the function $f : C \to B$ defined by $f(x) = y$, i.e.

$$f\left(2\sum_{i=0}^{\infty} a_i \left(\frac{1}{3}\right)^i\right) = \sum_{i=0}^{\infty} a_i \left(\frac{1}{2}\right)^i$$

we get a one-to-one and onto mapping. Therefore one can say that there are as many points in C as there are in B, which implies that the set C is uncountable.

Note:[1] The idea of defining measures using covers of sets was introduced by Carathéodory (1914). Hausdorff (1919) used this method to define the measures that now bear his name, and showed that the middle third Cantor set has positive and finite measure of dimension $\frac{\log 2}{\log 3}$.

[1]Thanks to Cettina Gauci Pulo for this information