The Cantor set is uncountable

February 13, 2009

Every $x \in [0,1]$ has at most two *ternary expansions* with a leading zero; that is, there are at most two sequences $(d_n)_{n\geq 1}$ taking values in $\{0,1,2\}$ with

$$x = 0.d_1d_2d_3\cdots \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} d_n \cdot 3^{-n}.$$

For example, $\frac{1}{3} = 0.10000\cdots = 0.022222\ldots$. Moreover, this is essentially the only way in which ambiguity can arise: if $0.d_1d_2d_3\ldots$ and $0.e_1e_2e_3\ldots$ are two different ternary expansions of x then writing $n = \min\{m \geq 1\colon d_m \neq e_m\}$, if $d_n < e_n$ then we necessarily have $e_n = d_n + 1$, $d_{n+1} = d_{n+2} = \cdots = 2$ and $e_{n+1} = e_{n+2} = \cdots = 0$. In particular, either e_n or d_n is equal to 1; so there is at most one ternary expansion of x which does not contain a 1.

The Cantor set C is defined by

$$\begin{aligned}
C_0 &= [0,1],\\
C_1 &= [0,\tfrac{1}{3}] \cup [\tfrac{2}{3},1],\\
C_2 &= [0,\tfrac{1}{9}] \cup [\tfrac{2}{9},\tfrac{1}{3}] \cup [\tfrac{2}{3},\tfrac{7}{9}] \cup [\tfrac{8}{9},1],\\
&\ \vdots\\
C &= \bigcap_{n=0}^{\infty} C_n.
\end{aligned}$$

Observe that

$$C_n = \{0.d_1d_2d_3\ldots : d_j \in \{0,2\} \text{ for } 1 \leq j \leq n\}$$

is the set of all numbers in $[0,1]$ which have a ternary expansion containing only the digits 0 or 2 in the first n places. So C is the set of all numbers in $[0,1]$ which have at least one ternary expansion containing only the digits 0 and 2. Moreover, from the remarks above it follows that each $x \in C$ has one and only one ternary expansion using only 0s and 2s.

Recall that a set X is countable if there is a sequence of elements of X which exhausts it; that is, if X can be written as

$$X = \{x_1, x_2, x_3, \ldots\}.$$

Otherwise, we say that X is uncountable.

Theorem. *The Cantor set is uncountable.*

Proof. We use a method of proof known as Cantor's diagonal argument. Suppose instead that C is countable, say $C = \{x^1, x^2, x^3, x^4, \ldots\}$. Write $x^i = 0.d_1^i d_2^i d_3^i d_4^i \ldots$ as a ternary expansion using only 0s and 2s. Then the elements of C all appear in the list:

$$\begin{aligned}
x^1 &= 0.d_1^1\, d_2^1\, d_3^1\, d_4^1\, \ldots \\
x^2 &= 0.d_1^2\, d_2^2\, d_3^2\, d_4^2\, \ldots \\
x^3 &= 0.d_1^3\, d_2^3\, d_3^3\, d_4^3\, \ldots \\
x^4 &= 0.d_1^4\, d_2^4\, d_3^4\, d_4^4\, \ldots \\
&\vdots
\end{aligned}$$

Let $(d_1, d_2, d_3, d_4, \ldots)$ be the sequence that differs from the diagonal sequence $(d_1^1, d_2^2, d_3^3, d_4^4, \ldots)$ in every entry, so that

$$d_j = \begin{cases} 0 & \text{if } d_j^j = 2, \\ 2 & \text{if } d_j^j = 0. \end{cases}$$

The ternary expansion $0.d_1\, d_2\, d_3\, d_4 \ldots$ does not appear in the list above since $d_j \neq d_j^j$. Now $x = 0.d_1\, d_2\, d_3\, d_4 \ldots$ is in C, but no element of C has two different ternary expansions using only 0s and 2s. So x does not appear in the list above, which is a contradiction. So C must be uncountable. $\square$