More on the Cantor set C.
Recall:

$$C = \left\{ \sum_{n=1}^{\infty} \frac{x_n}{3^n} : x_n \in \{0,2\} \ \forall n \in \mathbb{N} \right\}$$

$$= [0,1] \setminus \left(\bigcup_{n=1}^{\infty} S_n \right),$$

where $S_1 = (\frac{1}{3}, \frac{2}{3})$ and, for $n > 1$, S_n is the union of all intervals $(\frac{m}{3^n}, \frac{m+1}{3^n})$ that remain after removing $S_1, S_2, \ldots S_{n-1}$ from $[0,1]$.

__Definition__

A set $S \subseteq \mathbb{R}$ has __measure zero__ if, given $\varepsilon > 0$, $\exists$ a countable collection $\{I_n : n \in \mathbb{N}\}$ of intervals such that

$$S \subseteq \bigcup_{n=1}^{\infty} I_n \quad \text{and} \quad \sum_{n=1}^{\infty} \text{length}(I_n) < \varepsilon.$$

For example, any countable set $X = \{x_n : n \in \mathbb{N}\}$ of real numbers x_n has measure zero. (Proof: Let $I_n = [x_n, x_n]$ for each $n \in \mathbb{N}$.) But one shows that no positive-length interval has measure zero.

We do have:

__Theorem C1.__

C has measure zero.

__Proof.__ Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $(2/3)^N < \varepsilon$. Note that

$$C = [0,1] \setminus \left(\bigcup_{n=1}^{\infty} S_n \right) \subseteq [0,1] \setminus \left(\bigcup_{n=1}^{N} S_n \right).$$

Call the right hand side C_N. Since S_n is a disjoint union of 2^{n-1} open intervals, each of length 3^{-n}, and since the S_n's are disjoint from each other, we find that C_N is a union of (closed) intervals, of total length

$$1-\sum_{n=1}^{N}(\text{total length of } S_n) = \quad 1-\sum_{n=1}^{N}\frac{2^{n-1}}{3^n}$$

$$= 1-\left(1-\left(\tfrac{2}{3}\right)^N\right) = \left(\tfrac{2}{3}\right)^N < \varepsilon.$$

Since $C \subseteq C_N$, we see that C has measure zero. $\square$

We also have:

<u>Theorem C_2.</u> C is uncountable.

<u>Proof.</u>

Define $f: C \to [0,1]$ as follows. Let $c \in C$, so that c has a ternary expansion

$$c = \sum_{n=1}^{\infty}\frac{c_n}{3^n}$$

where $c_n = 0$ or 2 $\forall n \in \mathbb{N}$. Define

$$f(c) = \sum_{n=1}^{\infty}\frac{d_n}{2^n},$$

where $$d_n = c_n/2 = \begin{cases} 0 & \text{if } c_n = 0, \\ 1 & \text{if } c_n = 2. \end{cases}$$

Then f is a 1-to-1 correspondence of C onto $[0,1]$ (realized in binary). Since $[0,1]$ is uncountable, so is C. $\square$

Remarks.

So C is, in a sense, larger than $\mathbb{Q}$, because $\mathbb{Q}$ is countable but C is not.

But in a sense, C is smaller than $\mathbb{Q}$ because, given $\varepsilon > 0$, $\exists$ a finite set of intervals that cover C and have total length $< \varepsilon$ (see the proof of Theorem C_1 above). This is not the case for $\mathbb{Q}$, or even $\mathbb{Q} \cap (a,b)$ for any nonempty open interval (a,b)!