The Construction of Cantor Sets

By Ng Tze Beng

Cantor sets play an important role in real analysis, particularly in furnishing counter examples and exotic or pathological functions. Most frequently, we meet Cantor set of zero measure but the construction is canonical enough to apply to give Cantor set of positive measure. These are subsets of the closed interval [0,1] having measure greater or equal to 0 but less than 1. For the use of Cantor sets see *Composition and Riemann Integrability*, *Lebesgue Integration and Composition* and *Change of Variable or Subsitution in Riemann Integration.*

The first Cantor set we shall construct is the Cantor set of measure zero. This is the set left over after repeatedly deleting the middle third open intervals.

The Cantor set C_0

We shall start from the closed unit interval [0,1]. At the first stage, we delete the middle open interval (1/3, 2/3) from [0, 1]. We shall enumerate the open intervals to be deleted. We denote (1/3, 2/3) by $I(1,1)$. That is $I(1,1) = (1/3,2/3)$. The first construction gives us the remaining 2 closed intervals

$$[0, 1] - I(1,1) = [0,1] - (1/3,2/3) = [0, 1/3] \cup [2/3, 1].$$

Then at the second stage we delete the middle third open interval from each of the closed intervals. Thus, there are 2 open intervals to be deleted and they are

$I(2,1) = 1/3\ I(1, 1) =(1/9, 2/9)$ and

$I(2, 2) = I(2, 1) + 2/3$, the translation of $I(2,1)$ by 2/3.

Hence, we are left with $4 = 2^2$ closed intervals, two in [0, 1/3] and two in [2/3, 1]. By rescaling we see that the middle open third intervals of each of these 4 closed intervals are given by

$I(3,1) = 1/3\ I(2,1)$, $I(3,2) =1/3\ I(2,2)$ and

$I(3, 3) = I(3,1) + 2/3$, $I(3, 4) = I(3,2) +2/3$.

Repeating this construction, at the k-th stage, there are 2^{k-1} open intervals, each of length $1/3^k$, to be deleted from [0, 1] and for $k \geq 3$, they are given by

$I(k, j) = 1/3\ I(k-1, j)$, $j = 1, 2, \ldots, 2^{k-2}$;

$I(k, 2^{k-2}+j) = I(k, j) + 2/3$, $j = 1, 2, \ldots, 2^{k-2}$.

Let G be the union of these countable collection of disjoint open intervals, i.e.,

$$G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j).$$

So G is open. The Cantor set C_0 is defined to be the complement of G in [0, 1]. That is $C_0 = [0, 1] - G$ and is therefore a closed subset of [0, 1]. Note that at the k-th stage the total length of the intervals to be deleted is $2^{k-1}(1/3^k)$. Therefore, the measure of G, since the intervals in G are all disjoint, is given by

$$\mu(G) = \sum_{k=1}^{\infty} 2^{k-1}\frac{1}{3^k} = \frac{1}{3}\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^{k-1} = \frac{1}{3}\lim_{k\to\infty}\left(\frac{1-\left(\frac{2}{3}\right)^k}{1-\left(\frac{2}{3}\right)}\right) = 1.$$

Thus the measure of C_0 is equal to $\mu([0, 1]) - \mu(G) = 1-1 = 0$.

In order to define the Cantor function, we shall consider another representation of the Cantor set C_0 . This will also show that C_0 has the same cardinality as [0, 1] and so it is uncountable. We shall make use of the ternary representation of numbers in [0, 1]. There is a unique way of representing real numbers in (0, 1] by non-terminating ternary expansion.

We shall redefine these real numbers as the supremum of a sequence of rational numbers. Take any $1 \geq x > 0$ in [0,1]. Choose the greatest integer a_1 such that $a_1/3 < x \leq a_1/3 + 1/3$. Consider the rational numbers 0/3, 1/3 , 2/3. Then take the largest of these which is $< x$. That is we choose a_1 in {0,1,2} such that $\frac{a_1}{3} < x \leq \frac{a_1}{3} + \frac{1}{3}$. Thus $a_1 = 0 \Leftrightarrow 0/3 < x \leq 1/3$, $a_1 = 1 \Leftrightarrow 1/3 < x \leq 2/3$ and $a_1 = 2 \Leftrightarrow 2/3 < x \leq 3/3=1$. Let $d_1 = 0 + \frac{a_1}{3}$. Then $d_1 < x \leq d_1 + \frac{1}{3}$. Now choose integer a_2 in {0,1,2} such that $d_1 + \frac{a_2}{3^2} < x \leq d_1 + \frac{a_2}{3^2} + \frac{1}{3^2}$. Let now $d_2 = d_1 + \frac{a_2}{3^2} = 0 + \frac{a_1}{3} + \frac{a_2}{3^2}$. We have $d_2 < x \leq d_2 + \frac{1}{3^2}$. Continuing like this we shall obtain a sequence a_0 , a_1 , a_2 , …, of integers, $0 \leq a_i \leq 2$, such that if $d_n = d_{n-1} + \frac{a_n}{3^n} = 0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_n}{3^n}$ then $d_n < x \leq d_n + \frac{1}{3^n}$. The symbol $0 \cdot a_1\ a_2\ a_3 \ldots$ is called the *non-terminating ternary expansion* of x . So the set $\{d_1 , d_2 , d_3 , d_4 \ldots\}$ is bounded above by x. Then the least upper bound or supremum of this set is x. We deduce this as follows. If $x \neq \sup\{d_1 , d_2 , d_3 , d_4 \ldots\} = M$, then $x > M$. Then by the Archimedean property of the set of real numbers, there exists a counting number m such that $\frac{1}{m} < x - M$ and so $M + \frac{1}{m} < x$. Now take a non-terminating ternary expansion $0 \cdot c_1\ c_2\ c_3 \ldots$ for $\frac{1}{m}$. Let L be the first integer such that $c_L \neq 0$. Then $\frac{1}{3^L} < \frac{1}{m}$. Therefore, $M + \frac{1}{3^L} < M + \frac{1}{m} < x$. Hence since $M = \sup\{d_1 , d_2 , d_3 , d_4 \ldots\}$, $d_L + \frac{1}{3^L} \leq M + \frac{1}{3^L} < x$ and this contradicts $x \leq d_L + \frac{1}{3^L}$. Hence $x = \sup\{d_1 , d_2 , d_3 , d_4 \ldots\}$.

This representation of the real number x in (0, 1] is unique. Suppose two non-terminating ternary expressions $0\cdot a_1\ a_2\ a_3 \ldots$ and $0\cdot b_1\ b_2\ b_3 \ldots$ are such that for some integer $j \geq 0$, $a_j \neq b_j$ and $a_i = b_i$ for $i \leq j-1$. If $0\cdot a_1\ a_2\ a_3 \ldots$ represents x and $0\cdot b_1\ b_2\ b_3 \ldots$ represents y we shall show that $x \neq y$. Suppose that $a_j < b_j$. By the hypothesis we have

$$0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_{j-1}}{3^{j-1}} + \frac{a_j}{3^j} < x \leq 0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_{j-1}}{3^{j-1}} + \frac{a_j + 1}{3^j} .$$

But $0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_{j-1}}{3^{j-1}} + \frac{a_j + 1}{3^j} = 0 + \frac{b_1}{3} + \frac{b_2}{3^2} + \ldots + \frac{b_{j-1}}{3^{j-1}} + \frac{a_j + 1}{3^j}$

$$\leq 0 + \frac{b_1}{3} + \frac{b_2}{3^2} + \ldots + \frac{b_{j-1}}{3^{j-1}} + \frac{b_j}{3^j} < y.$$

Therefore, $x < y$ and so $x \neq y$. Similarly if $a_j > b_j$, we can show that $x \neq y$. Thus, this way of representing any real number in (0, 1] is unique. Let 0 be represented by the terminating ternary expansion 0.00. Notice that for $x = \frac{1}{3}$, the corresponding non-terminating expansion is the expression 0.022222…… which is a non-terminating expression with recurring '2'. $\frac{2}{3}$ is represented by 0.122222....... a non-terminating expression with recurring '2'. 1 is represented by 0.22222........ a non-terminating expression with recurring '2'. We are going to change our representation a little. We are going to use some terminating expansion in our representation as follows. If the non-terminating expansion of x consists of exactly one '1' in its expansion, say $0\cdot a_1\ a_2\ a_3$ .... has precisely $a_j = 1$ and that $a_k = 2$ for $k > j$, then we replace it by the terminating expansion $0\cdot b_1\ b_2\ b_3 \ldots b_{j-1}\ 2$ with $b_k = a_k$ *for* $k \leq j-1$, $b_j = 2$

and $b_k = 0$ for $k > j$, hence this representation has no '1' in it. Thus we are going to use the following convention: if a number x in [0, 1] has two ternary expansion, one with no 1's and one with at least one '1', then it is the one that has no 1's that is to be used. We refer to this as our system of expansion. We are going to use this representation to describe the set G in [0, 1].

$x \in I(1,1) = (1/3, 2/3) \Leftrightarrow$ the non terminating expansion of x has $a_1 = 1$ and a_2 and subsequent a_j are not all equal to 2 $\Leftrightarrow$ the system of expansion used for x has $a_1 = 1$. (Thi is because the terminating expansion excluded here has no '1', namely 2/3 which has terminating expansion 0.2 and recurring expansion $0.1\underline{2}$, where the underscore denotes repeating infinitely many times the number underscored.)

Now dividing by 3 has the effect of shifting the expansion by one place to the right and introducing a zero, i.e., it has the effect of moving the ternary point to the left. Thus,

$x \in I(2,1) = 1/3\ I(1,1) = (1/9, 2/9)$

$\Leftrightarrow$ the system of expansion used for x has $a_1 = 0$ and $a_2 = 1$.

Therefore, $x \in I(2,2) = I(2,1) + 2/3$

$\Leftrightarrow$ the system of expansion used for x has $a_1 = 2$ and $a_2 = 1$.

Hence $x \in I(2,1) \cup I(2,2) \Leftrightarrow$ the system of expansion used for x has *the first* '1' occurring in the second ternary place.

Then since $I(3,1) = 1/3\ I(2,1)$ and $I(3,2) = 1/3\ I(2,2)$, $x \in I(3,1) \cup I(3,2) \Leftrightarrow$ the system of expansion used for x has $a_1 = 0$ and has *the first* '1' occurring in the third ternary place. Since $I(3, 3) = I(3,1) + 2/3$ and $I(3, 4) = I(3,2) + 2/3$, $x \in I(3, 3) \cup I(3, 4) \Leftrightarrow$ the system of expansion used for x has $a_1 = 2$ and has *the first* '1' occurring in the third ternary place. Therefore, $x \in I(3,1) \cup I(3,2) \cup I(3, 3) \cup I(3, 4) \Leftrightarrow$ the system of expansion used for x has the *first* '1' occurring in the third ternary place. Thus by induction we see that $x \in \bigcup_{j=1}^{2^{k-1}} I(k,j) \Leftrightarrow$ the system of expansion used for x has the *first* '1' occurring in the k-th ternary place. Therefore, $x \in G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j) \Leftrightarrow$ the system of expansion used for x has at least one '1'. Hence $x \in C_0 = [0, 1] - G \Leftrightarrow$ the system of expansion used for x has no 1's. We have thus proved the following:

Lemma 1. C_0 consists of those numbers in [0, 1] whose representation in the above system of ternary expansion has no 1's. Therefore, we can write for any x in C_0

$x = \sum_{k=1}^{\infty} \frac{2b_k}{3^k}$, where $b_k = 0$ or 1.

Lemma 2. C_0 is uncountable.

Proof. Define $g : C_0 \to [0, 1]$ by $g(x) = \sum_{k=1}^{\infty} \frac{b_k}{2^k}$, where $x = \sum_{k=1}^{\infty} \frac{2b_k}{3^k}$. Then since every real number y in [0, 1] has a non-terminating binary representation of the form

$\sum_{k=1}^{\infty}\frac{b_k}{2^k}$ or $y = 0$, we can take x in C_0 to be $\sum_{k=1}^{\infty}\frac{2b_k}{3^k}$ or 0 if $y = 0$ and so $g(x) = y$. Hence, g is surjective. Since [0, 1] is uncountable, C_0 is uncountable.

Next we would like to extend the function to all of [0, 1]. Let $I(i, j)$ be denoted by the open interval $(a(i, j), b(i, j))$. Then $\{a(i, j), b(i, j)\} \subseteq C_0$. We shall show that $g(a(i, j)) = g(b(i, j))$. It is easily seen that in our system of ternary representation

$$a(i,j) = \sum_{k=1}^{i-1}\frac{2b_k}{3^k} + \frac{0}{3^i} + \sum_{k=i+1}^{\infty}\frac{2}{3^k} \text{ and}$$

$$b(i,j) = \sum_{k=1}^{i-1}\frac{2b_k}{3^k} + \frac{2}{3^i} \text{ , for some } b_k = 0 \text{ or } 1,\ k = 1,2,\ldots i-1.$$

Hence $g(a(i,j)) = \sum_{k=1}^{i-1}\frac{b_k}{2} + \frac{0}{2^i} + \sum_{k=i+1}^{\infty}\frac{1}{2^k} = \sum_{k=1}^{i-1}\frac{b_k}{2} + \frac{1}{2^i} = g(b(i,j))$. We define for x in $I(i, j)$, $g(x) = g(b(i, j)) = g(a(i, j))$. We shall next show that g is a non decreasing function, that is, g is a monotonic increasing function (not necessarily strictly increasing).

Lemma 3. Represent the real numbers in [0, 1] by non-terminating ternary expansions except for 0. In this representation suppose $x = \sum_{k=1}^{\infty}\frac{a_k}{3^k}$ and $y = \sum_{k=1}^{\infty}\frac{b_k}{3^k}$. Then

1. $x = y \Leftrightarrow a_i = b_i$ for all $i \geq 1$.

2. $x < y \Leftrightarrow a_1 < b_1$ or there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k < b_k$.

Proof. Part 1 follows from the uniqueness of the representation of the numbers either by the non terminating ternary expansion or by the system of representation used above.

Suppose $a_1 < b_1$. Then $x \leq \frac{a_1}{3} + \frac{1}{3} = \frac{a_1+1}{3} \leq \frac{b_1}{3} < y$. Similarly, if there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k < b_k$, then

$$x \leq \sum_{i=1}^{k-1}\frac{a_i}{3^i} + \frac{a_k}{3^k} + \frac{1}{3^k} = \sum_{i=1}^{k-1}\frac{b_i}{3^i} + \frac{a_k+1}{3^k} \leq \sum_{i=1}^{k-1}\frac{b_i}{3^i} + \frac{b_k}{3^k} < y.$$

Now if $x < y$, then $y - x > 0$. Then by the Archimedean property of the real number system, there exists an integer $k \geq 1$ such that $\frac{1}{3^k} < y - x$. Then

$$x \leq \sum_{i=1}^{k}\frac{a_i}{3^i} + \frac{1}{3^k} < \sum_{i=1}^{k}\frac{a_i}{3^i} + y - x < y \leq \sum_{i=1}^{k}\frac{b_i}{3^i} + \frac{1}{3^k}.$$

Therefore, $\sum_{i=1}^{k}\frac{a_i}{3^i} < \sum_{i=1}^{k}\frac{b_i}{3^i}$. Then $a_1 \leq b_1$. For if $a_1 > b_1$, then $\frac{a_1}{3} > \frac{b_1+1}{3} > \sum_{i=1}^{k}\frac{b_i}{3^i}$ and so $\sum_{i=1}^{k}\frac{a_i}{3^i} > \sum_{i=1}^{k}\frac{b_i}{3^i}$ contradicting $\sum_{i=1}^{k}\frac{a_i}{3^i} < \sum_{i=1}^{k}\frac{b_i}{3^i}$. Hence either $a_1 < b_1$ or $a_1 = b_1$. If

$a_1 = b_1$, then $\sum_{i=2}^{k} \frac{a_i}{3^i} < \sum_{i=2}^{k} \frac{b_i}{3^i}$. Therefore, we conclude again that $a_2 \le b_2$. If $a_2 < b_2$, then we get the conclusion of the Lemma. Because $\sum_{i=1}^{k} \frac{a_i}{3^i} < \sum_{i=1}^{k} \frac{b_i}{3^i}$, $a_i = b_i$ cannot hold for all $1 \le i \le k$. Thus for some j with $1 \le j \le k$ $a_i = b_i$ *for* $1 \le i < j$ and $a_j < b_j$. This completes the proof.

Next we shall see how our system of representation also gives the same conclusion. Observe that the convention used in our system of representation deviates from a non-terminating expansion only if the number is zero or has two possible ternary expansions with one of them involving no 1's. Thus if y has a terminating expansion either $y = 0$ or the terminating expansions has no 1's and ends in the number '2'. Thus if $x < y$, and y has a terminating ternary expansion,

$$y = \sum_{i=1}^{l-1} \frac{b_i}{3^i} + \frac{2}{3^l} = \sum_{i=1}^{l} \frac{b_i'}{3^i}, \text{ where } b_i' \text{ is equal to 0 or 2 and } b_l' = 2.$$

Then we can write $y = \sum_{i=1}^{l-1} \frac{b_i}{3^i} + \frac{1}{3^l} + \sum_{i=l+1}^{\infty} \frac{2}{3^i} = \sum_{i=1}^{\infty} \frac{c_i}{3^i}$. Thus if $x = \sum_{i=1}^{\infty} \frac{a_i}{3^i}$, $x < y$ implies that there exists integer $k \ge 1$ such that $a_i = c_i$ for all $1 \le i \le k-1$ and $a_k < c_k$ or $a_1 < c_1$, if $k = 1$. If $k \le l-1$, then we have $a_i = c_i = b_i'$ for all $1 \le i \le k-1$ and $a_k < b_k'$. If $k = l$, then we have $a_i = b_i'$ for all $1 \le i \le l-1$ and $a_l < c_l = 1 < 2 = b_l'$. If $k > l$, then we have $a_i = b_i'$ for all $1 \le i \le l-1$ and $a_l = c_l = 1 < 2 = b_l'$. Hence, in all cases, we obtain that $x < y$ implies that there exists integer $l \ge k > 1$ such that $a_i = b_i'$ for all $1 \le i \le k-1$ and $a_k < b_k'$ or $a_1 < b_1'$.

Now if x has the terminating ternary expansion $x = \sum_{i=1}^{p-1} \frac{a_i}{3^i} + \frac{2}{3^p} = \sum_{i=1}^{p} \frac{a_i'}{3^i}$, we can write it as $x = \sum_{i=1}^{p-1} \frac{a_i}{3^i} + \frac{1}{3^p} + \sum_{i=p+1}^{\infty} \frac{2}{3^i} = \sum_{i=1}^{\infty} \frac{d_i}{3^i}$. Hence by what we have just proved, there exists integer $l \ge k > 1$ such that $d_i = b_i'$ for all $1 \le i \le k-1$ and $d_k < b_k'$ or $d_1 < b_1'$. If $p = 1$, then $d_1 = 1$ and if $d_1 < b_1'$, $b_1' = 2$ and l must be greater than 1, otherwise $x = y$. Thus if $p = 1$, then $a_1' = 2 = b_1'$ and there exists an integer $l \ge k > 1$ such that $a_i = b_i'$ for all $1 \le i \le k-1$ and $a_k' < b_k'$. We now assume that $p > 1$. If $k < p$, then $a_i' = d_i = b_i'$ for all $1 \le i \le k-1$ and $a_k' = d_k < b_k'$. If If $k = p$, then $a_i' = d_i = b_i'$ for all $1 \le i \le k-1$ and $d_k = d_p = 1 < b_p'$. Hence $b_p' = 2$ and so $a_p' = 2 = b_p'$. Then it follows that $l > p$ and that there exists q such that $p < q \le l$ and $a_i' = b_i'$ for all $1 \le i \le q-1$ and $a_q < b_q'$. Now we shall see that $k \le p$. if $k > p$, then we have $a_i' = d_i = b_i'$ for all $1 \le i \le k-1$ and $d_k < b_k'$. Therefore, $a_i' = d_i = b_i'$ for $1 \le i \le p-1$, $1 = d_p = b_p'$, hence contradicting that b_p' is even. Hence we have shown that the same conclusion for Lemma 3 is true also for our system of representation of real numbers in [0, 1] by ternary expansion. The converse of the statement is obvious.

We summarise what we have proved in the following lemma.

Lemma 4. Represent the real numbers in [0, 1] by the system of ternary expansion as described above. In this representation suppose $x = \sum_{k=1}^{\infty} \frac{a'_k}{3^k}$ and $y = \sum_{k=1}^{\infty} \frac{b'_k}{3^k}$. Then

1. $x = y \Leftrightarrow a_i' = b_i'$ for all $i \geq 1$.
2. $x < y \Leftrightarrow a_1' < b_1'$ or there exists integer $k \geq 2$ such that $a_i' = b_i'$ for all $1 \leq i \leq k-1$ and $a_k' < b_k'$.

Now we are ready to investigate the monotonicity of the function g:[0, 1] →[0, 1]. We have shown that g is a surjective map.

Proposition 5. The map, g:[0, 1] →[0, 1], defined previously, is a bounded surjective monotonic increasing map. That is to say $x < y \Rightarrow g(x) \leq g(y)$.

Proof. We have already seen that g is surjective. It is obviously bounded. Now suppose x and y are in C_0 and that $x < y$. Then x and y have the representation, $x = \sum_{i=1}^{\infty} \frac{2a_i}{3^i}$ and $y = \sum_{i=1}^{\infty} \frac{2b_i}{3^i}$, where the a_i's and b_i's are either 0 or 1. By Lemma 4 either $a_1 = 0$ and $b_1 = 1$ or there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k = 0 < b_k = 1$. Note that $g(x) = \sum_{i=1}^{\infty} \frac{a_i}{2^i}$ and $g(y) = \sum_{i=1}^{\infty} \frac{b_i}{2^i}$. Hence if $a_1 = 0$ and $b_1 = 1$, then $g(x) = \frac{0}{2} + \sum_{i=2}^{\infty} \frac{a_i}{2^i} \leq \frac{1}{2} + \sum_{i=2}^{\infty} \frac{b_i}{2^i} = g(y)$. If there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k = 0 < b_k = 1$, then

$$g(x) = \sum_{i=1}^{k-1} \frac{a_i}{2^i} + \frac{0}{2^k} + \sum_{i=k+1}^{\infty} \frac{a_i}{2^i} \leq \sum_{i=1}^{k-1} \frac{a_i}{2^i} + \frac{1}{2^k} + \sum_{i=k+1}^{\infty} \frac{b_i}{2^i} = g(y).$$

Hence, if $x < y$ and x and y are in C_0, then $g(x) \leq g(y)$. Suppose now that $x < y$, $x \notin C_0$ and y is in C_0. Then $x \in I(i, j) = (a(i, j), b(i, j))$ for some (i, j). Obviously $x < y$ implies that $a(i, j) < y$ and so since both $a(i, j)$ and y are in C_0 by what we have just proved, $g(a(i, j)) \leq g(y)$. Therefore, by definition of g, $g(x) = g(a(i, j)) \leq g(y)$.

Similarly, if $x < y$, $y \notin C_0$ and x is in C_0, then $y \in I(i, j) = (a(i, j), b(i, j))$ for some (i, j) and $x < b(i, j)$. Again since both x and $b(i, j)$ are in C_0, we have $g(x) \leq g(b(i, j)) = g(y)$.

Suppose now $x < y$, $x, y \notin C_0$. Then for some (i, j) and (i', j'), $x \in I(i, j)$, $y \in I(i', j')$. Then $a(i, j) < x < y < b(i', j')$. Therefore, since both $a(i, j)$ and $b(i', j')$ are in C_0, $g(a(i, j)) \leq g(b(i', j'))$ and so $g(x) = g(a(i, j)) \leq g(b(i', j')) = g(y)$. Hence g is monotonically increasing. This completes the proof.

Next we shall state a general result concerning bounded monotone function whose range is an interval.

Theorem 6. Suppose f: $[a, b] \to \mathbf{R}$ is an increasing function. Then f is continuous, if and only if, the range of f, $J = f([a, b])$ is an interval.

Proof. If f is continuous, then by the Intermediate Value Theorem, the range $f([a, b])$ is an interval. Now if f is increasing, then by Theorem 2 of *Monotone Function, Function of Bounded Variation, Fundamental Theorem of Calculus,* the discontinuity of f can only be jump discontinuity. Suppose the range J is an interval. Suppose f is discontinuous at $x = k$ in (a, b). Then we have

$$\lim_{x\to k^-} f(x) = f(k^-) \le f(k) \le f(k^+) = \lim_{x\to k^+} f(x) \text{ and } f(k^-) < f(k^+).$$

Note that for any $y < k$, because f is increasing $f(y) \le f(k^-)$. Also for any $z > k$, $f(k^+) < f(z)$. Therefore, $(f([a, k)\cup(k, b]))\cap(f(k^-), f(k^+)) = \varnothing$. But by assumption the range J is an interval and so $(f(k^-), f(k^+)) \subseteq J$ and $(f([a, k)\cup(k, b]))\cap(f(k^-), f(k^+)) = (J - \{f(k)\})\cap(f(k^-), f(k^+)) \ne \varnothing$. This contradicts

$(f([a, k)\cup(k, b]))\cap(f(k^-), f(k^+)) = \varnothing$ and so we have $f(k^-) = f(k^+) = f(k)$ and so $\lim_{x\to k} f(x) = f(k)$ and that means f is continuous at $x = k$. We can similarly derive a contradiction when $k = a$ *or* b. If f is not continuous at $k = a$, then $f(a) < f(a^+)$ and so $(f(a), f(a^+))\cap J = \varnothing$, contradicting $(f(a), f(a^+)) \subseteq J$. Thus f must be continuous at $x = a$. If f is not continuous at $k = b$, then $f(b^-) < f(b)$ and so $(f(b^-), f(b))\cap J = \varnothing$, contradicting $(f(b^-), f(b)) \subseteq J$. Thus, f must be continuous at $x = b$. Therefore, f cannot have any discontinuity and so it is continuous. This completes the proof.

Proposition 7. The function g:[0, 1] $\to$[0, 1] is continuous.

Proof. By Proposition 5, the map g is onto and increasing. Since the range [0, 1] is an interval, by Theorem 6, g is continuous.

Next we shall reveal some interesting facts concerning the Cantor set C_0.

Theorem 8. The Cantor set C_0 is

(1) compact,

(2) no where dense, i.e., it contains no open intervals,

(3) its own boundary points,

(4) perfect, i.e.,it is its own set of accumulation point,

(5) totally disconnected and

(6) between any two points in C_0, there is an open interval not contained in C_0.

Proof. (1) Since C_0 is closed and bounded, it is compact by the Heine-Borel Theorem.

(2) Let $G_p = \bigcup_{k=1}^{p} \bigcup_{j=1}^{2^{k-1}} I(k,j)$ denote the union of the collection of disjoint open intervals deleted after the p-th stage in the construction. Suppose C_0 contains an open interval, then there is a point x in C_0 and a $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq C_0$. Hence $(x - \delta, x + \delta)\cap G = (x - \delta, x + \delta)\cap([0, 1] - C_0) = \varnothing$. Since $(x - \delta, x + \delta)$ and G are disjoint and are contained in [0,1], $\mu(G) + \mu((x - \delta, x + \delta)) \le 1$. Hence $1 + 2\delta \le 1$

implies $\delta \le 0$, contradicting $\delta > 0$. Therefore, C_0 cannot contain any open interval and so is nowhere dense. (We can also use $\mu((x-\delta, x+\delta)) \le \mu(C_0) = 0$ and so $2\delta \le 0$ contradicting $\delta > 0$.)

(3). C_0 is closed, by a characterization of closed set, because $C_0 = [0, 1] - G = \bigcap_{p=1}^{\infty}([0,1]-G_p)$ and each $[0, 1] - G_p$ is closed. Hence the closure of C_0 is C_0 and so its set of cluster points C_0' is contained in C_0. Let x be in $[0, 1]$. Then for any open set J containing x, say $J = U \cap [0, 1]$, where U is an open interval, J is non empty and contains more than one point and $J \cap G \ne \varnothing$ because $J \not\subset [0, 1] - G = C_0$ by part (2). This is because if $J \subseteq C_0$, then the interior of J, which is a non empty open interval is contained in C_0 contradicting part (2). Thus the closure of G, $\bar{G} = [0, 1]$. Therefore, the boundary of C_0, $\partial C_0 = \bar{C}_0 \cap \bar{G} = C_0 \cap [0,1] = C_0$.

(4) Since C_0 is closed the set of accumulation points of C_0, C_0' is a subset of C_0. We claim that $C_0' = C_0$. Let $x \in C_0$. Suppose $x \notin C_0'$. Then there exists $\delta > 0$, such that $(x-\delta, x+\delta) \cap C_0 = \{x\}$. Thus $(x-\delta, x+\delta) \cap ([0, 1] - C_0) = (x-\delta, x) \cup (x, x+\delta)$. That means $(x-\delta, x) \subseteq ([0, 1] - C_0) = G = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{2^{k-1}} I(k,j)$ and so

$$(x-\delta,\ x) = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{2^{k-1}} I(k,j) \cap (x-\delta, x).$$

Since the collection $\{I(k,j) : k = 1, \cdots, \infty; j = 1, 2, \cdots, 2^{k-1}\}$ is a collection of disjoint open intervals, and $(x-\delta, x)$ is connected, $(x-\delta, x) \subseteq I(i,j)$ for some i and j. This is because otherwise $(x-\delta, x)$ would be a disjoint union of open intervals contradicting that it is connected. Since $x \notin I(i,j)$, $x = \sup I(i,j) = b(i,j)$. Similarly,

$$(x,\ x+\delta) = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{2^{k-1}} I(k,j) \cap (x, x+\delta).$$

We deduce as before then that for some p and q, $(x, x+\delta) \subseteq I(p, q)$. Thus because $x \notin I(p, q)$, $x = \inf I(p, q) = a(p, q)$. Therefore, $a(p, q) = x = b(i, j)$, contradicting that $a(p, q) \ne b(i, j)$ by virtue of the definition of $I(i, j)$'s . Therefore $x \in C_0'$. Thus $C_0' = C_0$.

(5) Since by (2) C_0 does not contain any open interval and since the connected subsets of **R** are either the singleton sets or the intervals, the only components of C_0 are the singletons $\{x\}$, $x \in C_0$. Therefore, C_0 is totally disconnected.

(6) Suppose $x < y$ and x, y are in C_0. Then $(x, y) \cap ([0, 1] - C_0) \ne \varnothing$ and is a disjoint union of open intervals and so (x, y) contains at least one open interval.

The Cantor Set C_k

Now we turn our attention towards constructing a Cantor set of positive measure in [0, 1]. Indeed the construction also applies to give a different Cantor set of measure zero. The procedure of deleting the middle portion of each of the remaining closed sets is followed here but using a different specified length. (This idea can be generalised by not requiring the intervals to be deleted to be the middle portion, to give a *generalised Cantor set*.) Hence the construction is canonical. Let k be any real number with $0 \le k < 1$. Let $\delta = (1 - k) > 0$. Start with the closed unit interval [0, 1]. Let $I(1,1)$ be the middle open interval of length $\delta/2$. Then if we write $I(1,1) = (a(1,1), b(1,1))$, then $a(1,1) = 1/2 - \delta/4$. Then let $G_1 = I(1,1)$ and $F_1 = [0,1] - G_1$. The first stage is to form F_1. Then F_1 is the disjoint union of 2 closed intervals each of length equal to $a(1,1) = 1/2 - \delta/4$. Let $I(2,1)$ and $I(2,2)$ be respectively the middle open intervals, each of length $\delta/2^3$, of the 2 closed intervals of F_1. The open intervals are ordered from the left to the right by the second indices. Let $G_2 = I(1,1) \cup I(2,1) \cup I(2,2)$. Suppose $I(2,1) = (a(2,1), b(2,1))$. Then $a(2,1) = 1/2\, a(1,1) - \delta/2^4 = 1/2^2 - \delta(1/2^2 - 1/2^4)$. The second stage is to form $F_2 = [0,1] - G_2 = [0,1] - G_1 - I(2,1) \cup I(2,2)$. Then F_2 is the disjoint union of 2^2 closed intervals each of length equaling $a(2,1) = 1/2^2 - \delta(1/2^2 - 1/2^4)$. Then let $I(3, j), j = 1,..,2^2$ be respectively the middle portion of length $\delta/2^5$ of the 2^2 closed intervals again ordered by the second indices, from left to right in the sense that $I(3, j) < I(3, k)$, if and only if, $j < k$, if and only if,,there exist x in $I(3, j)$ and y in $I(3, k)$ such that $x < y$. Stage three is to remove from F_2 these 2^2 open intervals, that is to form $F_3 = F_2 - \cup\{ I(3, j), j = 1,..,2^2 \} = [0, 1] - G_3$, where $G_3 = \bigcup_{k=1}^{3} \bigcup_{j=1}^{2^{k-1}} I(k,j)$. Thus F_3 consists of 2^3 disjoint closed intervals, each of length equaling $a(3,1)$, where $I(3,1) = (a(3,1), b(3,1))$. It is easily seen that $a(3,1) = 1/2^3 - \delta(1/2^3 - 1/2^6)$. Let $H_4 = \bigcup_{j=1}^{2^{4-1}} I(4,j)$, where $I(4, j), j = 1,..,2^3$ are the middle open intervals each of length $\delta/2^7$, of each of the disjoint closed intervals. Continuing in this way, at the n-th stage we have $H_n = \bigcup_{j=1}^{2^{n-1}} I(n,j)$, where $I(n, j) = (a(n, j), b(n, j))$ is an open interval of length $\delta/2^{2n-1}$ and the $I(n, j)$'s are ordered from left to right by the second indices according to the natural ordering of elements in the intervals. Note that $a(n,1) = 1/2\, a(n-1,1) - \delta(1/2^{2n}) = 1/2^n - \delta(1/2^n - 1/2^{2n})$. We obtain F_n by deleting from F_{n-1} the uion of open intervals H_n. Hence $F_n = F_{n-1} - H_n = [0,1] - G_n$, where $G_n = \bigcup_{k=1}^{n} \bigcup_{j=1}^{2^{k-1}} I(k,j)$ and F_n is a disjoint union of 2^n closed intervals each of length equaling $a(n,1) = 1/2^n - \delta(1/2^n - 1/2^{2n})$. To obtained F_{n+1}, we shall delete from F_n open intervals each of length equaling $\delta/2^{2n+1}$ from the 2^n closed intervals. That is, if we let $H_{n+1} = \bigcup_{j=1}^{2^n} I(n+1,j)$, where $I(n+1, j), j = 1,..,2^n$ are these open intervals to be deleted, then $F_{n+1} = F_n - H_{n+1} = [0, 1] - G_{n+1}$, where $G_{n+1} = \bigcup_{k=1}^{n+1} \bigcup_{j=1}^{2^{k-1}} I(k,j) = \bigcup_{k=1}^{n+1} H_k$. Obviously F_{n+1} is a collection of disjoint closed intervals, each of length equal to $a(n+1,1)$. Note also that for each $n \ge 1$, $F_{n+1} \subseteq F_n$.

Let $G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j) = \bigcup_{k=1}^{\infty} H_k$. Then G is a disjoint union of open intervals. The length or measure of H_k is given by $2^{k-1}\frac{\delta}{2^{2k-1}} = \frac{\delta}{2^k}$. Thus the measure or length of G,

$$\mu(G) = \sum_{k=1}^{\infty} \mu(H_k) = \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \frac{\delta}{2} \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = \frac{\delta}{2} \lim_{k \to \infty} \left(\frac{1-\left(\frac{1}{2}\right)^k}{1-\left(\frac{1}{2}\right)} \right) = \delta.$$

The Cantor set C_k is defined to be the complement of G in [0, 1]. That is, $C_k = [0,1] - G = [0,1] - \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j) = [0,1] - \bigcup_{k=1}^{\infty} G_k = \bigcap_{k=1}^{\infty} ([0,1] - G_k) = \bigcap_{k=1}^{\infty} F_k$. Thus the measure of C_k is equal to $\mu([0,1]) - \mu(G) = 1 - \delta = k$.

Then C_k satisfies the properties stated in Theorem 8.

Theorem 9. The Cantor set C_k defined above for $0 \le k < 1$ is uncountable and

(1) compact,

(2) no where dense, i.e., it contains no open intervals,

(3) its own boundary points,

(4) perfect, i.e., is its own set of accumulation point,

(5) totally disconneceted and

(6) between any two points in C_k, there is an open interval not contained in C_k.

Proof. Part (1) and part (3) to part(6) are proved in exactly the same way as in Theorem 8. Part (2) needs a different approach to prove. Suppose C_k contains an open interval say (c, d). Then $(c,d) \subseteq C_k = \bigcap_{k=1}^{\infty} F_k$ implies that $(c, d) \subseteq F_k$ for each $k \ge 1$. Now since $\lim_{k\to\infty} (\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}})) = 0$, there exists an integer N such that $k > N$ implies that $\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}}) < d - c$. Now we fix a $k > N$. Note first that $(c, d) \cap F_k = (c, d)$. Since the non-trivial interval (c, d) is connected and F_k is a disjoint union of closed intervals, (c, d) must be contained in one of these closed intervals, each of length $\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}})$. Hence, $d - c \le \frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}})$ contradicting $\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}}) < d - c$. Therefore, C_k does not contain an open interval. (The main thrust of the argument is that the F_k consists of disjoint closed intervals, whose maximum length tends to zero as k tends to infinity.) This proves part (2). That C_k is uncountable is a consequence of the following proposition.

Proposition 10. There is a continuous strictly increasing bijective map $f : [0, 1] \to [0,1]$ mapping the Cantor set C_k onto the the Cantor set C_0 defined earlier using the "deleted middle third intervals" construction.
Proof. This is given by Lemma 1 of *Composition and Riemann Integrability.* Note that the proof given there applies also to the case $k = 0$.

Since C_0 constructed using the "deleted middle third intervals" construction is uncountable and the function f given by Proposition 10 maps C_k bijectively onto C_0, C_k is also uncountable.

The Construction of Cantor Sets

By Ng Tze Beng

Cantor sets play an important role in real analysis, particularly in furnishing counter examples and exotic or pathological functions. Most frequently, we meet Cantor set of zero measure but the construction is canonical enough to apply to give Cantor set of positive measure. These are subsets of the closed interval [0,1] having measure greater or equal to 0 but less than 1. For the use of Cantor sets see *Composition and Riemann Integrability*, *Lebesgue Integration and Composition* and *Change of Variable or Subsitution in Riemann Integration.*

The first Cantor set we shall construct is the Cantor set of measure zero. This is the set left over after repeatedly deleting the middle third open intervals.

The Cantor set C_0

We shall start from the closed unit interval [0,1]. At the first stage, we delete the middle open interval (1/3, 2/3) from [0, 1]. We shall enumerate the open intervals to be deleted. We denote (1/3, 2/3) by $I(1,1)$. That is $I(1,1) = (1/3,2/3)$. The first construction gives us the remaining 2 closed intervals

$$[0, 1] - I(1,1) = [0,1] - (1/3,2/3) = [0, 1/3] \cup [2/3, 1].$$

Then at the second stage we delete the middle third open interval from each of the closed intervals. Thus, there are 2 open intervals to be deleted and they are

$$I(2,1) = 1/3\ I(1, 1) = (1/9, 2/9) \text{ and}$$
$$I(2, 2) = I(2, 1) + 2/3 \text{ , the translation of } I(2,1) \text{ by } 2/3.$$

Hence, we are left with $4 = 2^2$ closed intervals, two in [0, 1/3] and two in [2/3, 1]. By rescaling we see that the middle open third intervals of each of these 4 closed intervals are given by

$$I(3,1) = 1/3\ I(2,1),\ \ I(3,2) = 1/3\ I(2,2) \text{ and}$$
$$I(3, 3) = I(3,1) + 2/3,\ \ I(3, 4) = I(3,2) + 2/3.$$

Repeating this construction, at the k-th stage, there are 2^{k-1} open intervals, each of length $1/3^k$, to be deleted from [0, 1] and for $k \geq 3$, they are given by

$$I(k, j) = 1/3\ I(k-1, j),\ \ j = 1, 2, \ldots, 2^{k-2};$$
$$I(k, 2^{k-2}+j) = I(k, j) + 2/3,\ \ j = 1, 2, \ldots, 2^{k-2}.$$

Let G be the union of these countable collection of disjoint open intervals, i.e.,

$$G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j).$$

So G is open. The Cantor set C_0 is defined to be the complement of G in [0, 1]. That is $C_0 = [0, 1] - G$ and is therefore a closed subset of [0, 1]. Note that at the k-th stage the total length of the intervals to be deleted is $2^{k-1}(1/3^k)$. Therefore, the measure of G, since the intervals in G are all disjoint, is given by

$$\mu(G) = \sum_{k=1}^{\infty} 2^{k-1}\frac{1}{3^k} = \frac{1}{3}\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^{k-1} = \frac{1}{3}\lim_{k\to\infty}\left(\frac{1-\left(\frac{2}{3}\right)^k}{1-\left(\frac{2}{3}\right)}\right) = 1.$$

Thus the measure of C_0 is equal to $\mu([0, 1]) - \mu(G) = 1-1 = 0$.

In order to define the Cantor function, we shall consider another representation of the Cantor set C_0 . This will also show that C_0 has the same cardinality as [0, 1] and so it is uncountable. We shall make use of the ternary representation of numbers in [0, 1]. There is a unique way of representing real numbers in (0, 1] by non-terminating ternary expansion.

We shall redefine these real numbers as the supremum of a sequence of rational numbers. Take any $1 \geq x > 0$ in [0,1]. Choose the greatest integer a_1 such that $a_1/3 < x \leq a_1/3 + 1/3$. Consider the rational numbers 0/3, 1/3, 2/3. Then take the largest of these which is $< x$. That is we choose a_1 in {0,1,2} such that $\frac{a_1}{3} < x \leq \frac{a_1}{3} + \frac{1}{3}$. Thus $a_1 = 0 \Leftrightarrow 0/3 < x \leq 1/3$, $a_1 = 1 \Leftrightarrow 1/3 < x \leq 2/3$ and $a_1 = 2 \Leftrightarrow 2/3 < x \leq 3/3=1$. Let $d_1 = 0 + \frac{a_1}{3}$. Then $d_1 < x \leq d_1 + \frac{1}{3}$. Now choose integer a_2 in {0,1,2} such that $d_1 + \frac{a_2}{3^2} < x \leq d_1 + \frac{a_2}{3^2} + \frac{1}{3^2}$. Let now $d_2 = d_1 + \frac{a_2}{3^2} = 0 + \frac{a_1}{3} + \frac{a_2}{3^2}$. We have $d_2 < x \leq d_2 + \frac{1}{3^2}$. Continuing like this we shall obtain a sequence $a_0, a_1, a_2, \ldots$, of integers, $0 \leq a_i \leq 2$, such that if $d_n = d_{n-1} + \frac{a_n}{3^n} = 0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_n}{3^n}$ then $d_n < x \leq d_n + \frac{1}{3^n}$. The symbol $0 \cdot a_1\ a_2\ a_3 \ldots$ is called the *non-terminating ternary expansion* of x. So the set $\{d_1, d_2, d_3, d_4 \ldots\}$ is bounded above by x. Then the least upper bound or supremum of this set is x. We deduce this as follows. If $x \neq \sup\{d_1, d_2, d_3, d_4 \ldots\} = M$, then $x > M$. Then by the Archimedean property of the set of real numbers, there exists a counting number m such that $\frac{1}{m} < x - M$ and so $M + \frac{1}{m} < x$. Now take a non-terminating ternary expansion $0 \cdot c_1\ c_2\ c_3 \ldots$ for $\frac{1}{m}$. Let L be the first integer such that $c_L \neq 0$. Then $\frac{1}{3^L} < \frac{1}{m}$. Therefore, $M + \frac{1}{3^L} < M + \frac{1}{m} < x$. Hence since $M = \sup\{d_1, d_2, d_3, d_4 \ldots\}$, $d_L + \frac{1}{3^L} \leq M + \frac{1}{3^L} < x$ and this contradicts $x \leq d_L + \frac{1}{3^L}$. Hence $x = \sup\{d_1, d_2, d_3, d_4 \ldots\}$.

This representation of the real number x in (0, 1] is unique. Suppose two non-terminating ternary expressions $0\cdot a_1\ a_2\ a_3 \ldots$ and $0\cdot b_1\ b_2\ b_3 \ldots$ are such that for some integer $j \geq 0$, $a_j \neq b_j$ and $a_i = b_i$ for $i \leq j - 1$. If $0\cdot a_1\ a_2\ a_3 \ldots$ represents x and $0\cdot b_1\ b_2\ b_3 \ldots$ represents y we shall show that $x \neq y$. Suppose that $a_j < b_j$. By the hypothesis we have

$$0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_{j-1}}{3^{j-1}} + \frac{a_j}{3^j} < x \leq 0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_{j-1}}{3^{j-1}} + \frac{a_j + 1}{3^j}.$$

But $0 + \frac{a_1}{3} + \frac{a_2}{3^2} + \ldots + \frac{a_{j-1}}{3^{j-1}} + \frac{a_j+1}{3^j} = 0 + \frac{b_1}{3} + \frac{b_2}{3^2} + \ldots + \frac{b_{j-1}}{3^{j-1}} + \frac{a_j+1}{3^j}$

$$\leq 0 + \frac{b_1}{3} + \frac{b_2}{3^2} + \ldots + \frac{b_{j-1}}{3^{j-1}} + \frac{b_j}{3^j} < y.$$

Therefore, $x < y$ and so $x \neq y$. Similarly if $a_j > b_j$, we can show that $x \neq y$. Thus, this way of representing any real number in (0, 1] is unique. Let 0 be represented by the terminating ternary expansion 0.00. Notice that for $x = \frac{1}{3}$, the corresponding non-terminating expansion is the expression 0.022222…… which is a non-terminating expression with recurring '2'. $\frac{2}{3}$ is represented by 0.122222....... a non-terminating expression with recurring '2'. 1 is represented by 0.22222........ a non-terminating expression with recurring '2'. We are going to change our representation a little. We are going to use some terminating expansion in our representation as follows. If the non-terminating expansion of x consists of exactly one '1' in its expansion, say $0\cdot a_1\ a_2\ a_3$ .... has precisely $a_j = 1$ and that $a_k = 2$ for $k > j$, then we replace it by the terminating expansion $0\cdot b_1\ b_2\ b_3 \ldots b_{j-1}\, 2$ with $b_k = a_k$ *for* $k \leq j-1$, $b_j = 2$

and $b_k = 0$ for $k > j$, hence this representation has no '1' in it. Thus we are going to use the following convention: if a number x in [0, 1] has two ternary expansion, one with no 1's and one with at least one '1', then it is the one that has no 1's that is to be used. We refer to this as our system of expansion. We are going to use this representation to describe the set G in [0, 1].

$x \in I(1,1) = (1/3, 2/3) \Leftrightarrow$ the non terminating expansion of x has $a_1 = 1$ and a_2 and subsequent a_j are not all equal to 2 $\Leftrightarrow$ the system of expansion used for x has $a_1 = 1$. (Thi is because the terminating expansion excluded here has no '1', namely 2/3 which has terminating expansion 0.2 and recurring expansion $0.1\underline{2}$, where the underscore denotes repeating infinitely many times the number underscored.)

Now dividing by 3 has the effect of shifting the expansion by one place to the right and introducing a zero, i.e., it has the effect of moving the ternary point to the left. Thus,

$x \in I(2,1) = 1/3\ I(1,1) = (1/9, 2/9)$

$\Leftrightarrow$ the system of expansion used for x has $a_1 = 0$ and $a_2 = 1$.

Therefore, $x \in I(2,2) = I(2,1) + 2/3$

$\Leftrightarrow$ the system of expansion used for x has $a_1 = 2$ and $a_2 = 1$.

Hence $x \in I(2,1) \cup I(2,2) \Leftrightarrow$ the system of expansion used for x has *the first* '1' occurring in the second ternary place.

Then since $I(3,1) = 1/3\ I(2,1)$ and $I(3,2) = 1/3\ I(2,2)$, $x \in I(3,1) \cup I(3,2) \Leftrightarrow$ the system of expansion used for x has $a_1 = 0$ and has *the first* '1' occurring in the third ternary place. Since $I(3, 3) = I(3,1) + 2/3$ and $I(3, 4) = I(3,2) + 2/3$, $x \in I(3, 3) \cup I(3, 4) \Leftrightarrow$ the system of expansion used for x has $a_1 = 2$ and has *the first* '1' occurring in the third ternary place. Therefore, $x \in I(3,1) \cup I(3,2) \cup I(3, 3) \cup I(3, 4) \Leftrightarrow$ the system of expansion used for x has the *first* '1' occurring in the third ternary place. Thus by induction we see that $x \in \bigcup_{j=1}^{2^{k-1}} I(k,j) \Leftrightarrow$ the system of expansion used for x has the *first* '1' occurring in the k-th ternary place. Therefore, $x \in G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j) \Leftrightarrow$ the system of expansion used for x has at least one '1'. Hence $x \in C_0 = [0, 1] - G \Leftrightarrow$ the system of expansion used for x has no 1's. We have thus proved the following:

Lemma 1. C_0 consists of those numbers in [0, 1] whose representation in the above system of ternary expansion has no 1's. Therefore, we can write for any x in C_0 $x = \sum_{k=1}^{\infty} \frac{2b_k}{3^k}$, where $b_k = 0$ or 1.

Lemma 2. C_0 is uncountable.

Proof. Define g : $C_0 \to [0, 1]$ by $g(x) = \sum_{k=1}^{\infty} \frac{b_k}{2^k}$, where $x = \sum_{k=1}^{\infty} \frac{2b_k}{3^k}$. Then since every real number y in [0, 1] has a non-terminating binary representation of the form

$\sum_{k=1}^{\infty}\frac{b_k}{2^k}$ or $y = 0$, we can take x in C_0 to be $\sum_{k=1}^{\infty}\frac{2b_k}{3^k}$ or 0 if $y = 0$ and so $g(x) = y$. Hence, g is surjective. Since [0, 1] is uncountable, C_0 is uncountable.

Next we would like to extend the function to all of [0, 1]. Let $I(i, j)$ be denoted by the open interval $(a(i, j), b(i, j))$. Then $\{a(i, j), b(i, j)\} \subseteq C_0$. We shall show that $g(a(i, j)) = g(b(i, j))$. It is easily seen that in our system of ternary representation

$$a(i,j) = \sum_{k=1}^{i-1}\frac{2b_k}{3^k} + \frac{0}{3^i} + \sum_{k=i+1}^{\infty}\frac{2}{3^k} \text{ and}$$

$$b(i,j) = \sum_{k=1}^{i-1}\frac{2b_k}{3^k} + \frac{2}{3^i} \text{ , for some } b_k = 0 \text{ or } 1,\ k = 1,2,\ldots i-1.$$

Hence $g(a(i,j)) = \sum_{k=1}^{i-1}\frac{b_k}{2} + \frac{0}{2^i} + \sum_{k=i+1}^{\infty}\frac{1}{2^k} = \sum_{k=1}^{i-1}\frac{b_k}{2} + \frac{1}{2^i} = g(b(i,j))$. We define for x in $I(i, j)$, $g(x) = g(b(i, j)) = g(a(i, j))$. We shall next show that g is a non decreasing function, that is, g is a monotonic increasing function (not necessarily strictly increasing).

Lemma 3. Represent the real numbers in [0, 1] by non-terminating ternary expansions except for 0. In this representation suppose $x = \sum_{k=1}^{\infty}\frac{a_k}{3^k}$ and $y = \sum_{k=1}^{\infty}\frac{b_k}{3^k}$. Then

1. $x = y \Leftrightarrow a_i = b_i$ for all $i \geq 1$.
2. $x < y \Leftrightarrow a_1 < b_1$ or there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k < b_k$.

Proof. Part 1 follows from the uniqueness of the representation of the numbers either by the non terminating ternary expansion or by the system of representation used above.

Suppose $a_1 < b_1$. Then $x \leq \frac{a_1}{3} + \frac{1}{3} = \frac{a_1+1}{3} \leq \frac{b_1}{3} < y$. Similarly, if there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k < b_k$, then

$$x \leq \sum_{i=1}^{k-1}\frac{a_i}{3^i} + \frac{a_k}{3^k} + \frac{1}{3^k} = \sum_{i=1}^{k-1}\frac{b_i}{3^i} + \frac{a_k+1}{3^k} \leq \sum_{i=1}^{k-1}\frac{b_i}{3^i} + \frac{b_k}{3^k} < y.$$

Now if $x < y$, then $y - x > 0$. Then by the Archimedean property of the real number system, there exists an integer $k \geq 1$ such that $\frac{1}{3^k} < y - x$. Then

$$x \leq \sum_{i=1}^{k}\frac{a_i}{3^i} + \frac{1}{3^k} < \sum_{i=1}^{k}\frac{a_i}{3^i} + y - x < y \leq \sum_{i=1}^{k}\frac{b_i}{3^i} + \frac{1}{3^k}.$$

Therefore, $\sum_{i=1}^{k}\frac{a_i}{3^i} < \sum_{i=1}^{k}\frac{b_i}{3^i}$. Then $a_1 \leq b_1$. For if $a_1 > b_1$, then $\frac{a_1}{3} > \frac{b_1+1}{3} > \sum_{i=1}^{k}\frac{b_i}{3^i}$ and so $\sum_{i=1}^{k}\frac{a_i}{3^i} > \sum_{i=1}^{k}\frac{b_i}{3^i}$ contradicting $\sum_{i=1}^{k}\frac{a_i}{3^i} < \sum_{i=1}^{k}\frac{b_i}{3^i}$. Hence either $a_1 < b_1$ or $a_1 = b_1$. If

$a_1 = b_1$, then $\sum_{i=2}^{k} \frac{a_i}{3^i} < \sum_{i=2}^{k} \frac{b_i}{3^i}$. Therefore, we conclude again that $a_2 \le b_2$. If $a_2 < b_2$, then we get the conclusion of the Lemma. Because $\sum_{i=1}^{k} \frac{a_i}{3^i} < \sum_{i=1}^{k} \frac{b_i}{3^i}$, $a_i = b_i$ cannot hold for all $1 \le i \le k$. Thus for some j with $1 \le j \le k$ $a_i = b_i$ *for* $1 \le i < j$ and $a_j < b_j$. This completes the proof.

Next we shall see how our system of representation also gives the same conclusion. Observe that the convention used in our system of representation deviates from a non-terminating expansion only if the number is zero or has two possible ternary expansions with one of them involving no 1's. Thus if y has a terminating expansion either $y = 0$ or the terminating expansions has no 1's and ends in the number '2'. Thus if $x < y$, and y has a terminating ternary expansion,

$$y = \sum_{i=1}^{l-1} \frac{b_i}{3^i} + \frac{2}{3^l} = \sum_{i=1}^{l} \frac{b_i'}{3^i}, \text{ where } b_i' \text{ is equal to 0 or 2 and } b_l' = 2.$$

Then we can write $y = \sum_{i=1}^{l-1} \frac{b_i}{3^i} + \frac{1}{3^l} + \sum_{i=l+1}^{\infty} \frac{2}{3^i} = \sum_{i=1}^{\infty} \frac{c_i}{3^i}$. Thus if $x = \sum_{i=1}^{\infty} \frac{a_i}{3^i}$, $x < y$ implies that there exists integer $k \ge 1$ such that $a_i = c_i$ for all $1 \le i \le k-1$ and $a_k < c_k$ or $a_1 < c_1$, if $k = 1$. If $k \le l-1$, then we have $a_i = c_i = b_i'$ for all $1 \le i \le k-1$ and $a_k < b_k'$. If $k = l$, then we have $a_i = b_i'$ for all $1 \le i \le l-1$ and $a_l < c_l = 1 < 2 = b_l'$. If $k > l$, then we have $a_i = b_i'$ for all $1 \le i \le l-1$ and $a_l = c_l = 1 < 2 = b_l'$. Hence, in all cases, we obtain that $x < y$ implies that there exists integer $l \ge k > 1$ such that $a_i = b_i'$ for all $1 \le i \le k-1$ and $a_k < b_k'$ or $a_1 < b_1'$.

Now if x has the terminating ternary expansion $x = \sum_{i=1}^{p-1} \frac{a_i}{3^i} + \frac{2}{3^p} = \sum_{i=1}^{p} \frac{a_i'}{3^i}$, we can write it as $x = \sum_{i=1}^{p-1} \frac{a_i}{3^i} + \frac{1}{3^p} + \sum_{i=p+1}^{\infty} \frac{2}{3^i} = \sum_{i=1}^{\infty} \frac{d_i}{3^i}$. Hence by what we have just proved, there exists integer $l \ge k > 1$ such that $d_i = b_i'$ for all $1 \le i \le k-1$ and $d_k < b_k'$ or $d_1 < b_1'$. If $p = 1$, then $d_1 = 1$ and if $d_1 < b_1'$, $b_1' = 2$ and l must be greater than 1, otherwise $x = y$. Thus if $p = 1$, then $a_1' = 2 = b_1'$ and there exists an integer $l \ge k > 1$ such that $a_i = b_i'$ for all $1 \le i \le k-1$ and $a_k' < b_k'$. We now assume that $p > 1$. If $k < p$, then $a_i' = d_i = b_i'$ for all $1 \le i \le k-1$ and $a_k' = d_k < b_k'$. If If $k = p$, then $a_i' = d_i = b_i'$ for all $1 \le i \le k-1$ and $d_k = d_p = 1 < b_p'$. Hence $b_p' = 2$ and so $a_p' = 2 = b_p'$. Then it follows that $l > p$ and that there exists q such that $p < q \le l$ and $a_i' = b_i'$ for all $1 \le i \le q-1$ and $a_q < b_q'$. Now we shall see that $k \le p$. if $k > p$, then we have $a_i' = d_i = b_i'$ for all $1 \le i \le k-1$ and $d_k < b_k'$. Therefore, $a_i' = d_i = b_i'$ for $1 \le i \le p-1$, $1 = d_p = b_p'$, hence contradicting that b_p' is even. Hence we have shown that the same conclusion for Lemma 3 is true also for our system of representation of real numbers in [0, 1] by ternary expansion. The converse of the statement is obvious.

We summarise what we have proved in the following lemma.

Lemma 4. Represent the real numbers in [0, 1] by the system of ternary expansion as described above. In this representation suppose $x = \sum_{k=1}^{\infty} \frac{a_k'}{3^k}$ and $y = \sum_{k=1}^{\infty} \frac{b_k'}{3^k}$. Then

1. $x = y \Leftrightarrow a_i' = b_i'$ for all $i \geq 1$.

2. $x < y \Leftrightarrow a_1' < b_1'$ or there exists integer $k \geq 2$ such that $a_i' = b_i'$ for all $1 \leq i \leq k-1$ and $a_k' < b_k'$.

Now we are ready to investigate the monotonicity of the function g:[0, 1] →[0, 1]. We have shown that g is a surjective map.

Proposition 5. The map, g:[0, 1] →[0, 1], defined previously, is a bounded surjective monotonic increasing map. That is to say $x < y \Rightarrow g(x) \leq g(y)$.

Proof. We have already seen that g is surjective. It is obviously bounded. Now suppose x and y are in C_0 and that $x < y$. Then x and y have the representation, $x = \sum_{i=1}^{\infty} \frac{2a_i}{3^i}$ and $y = \sum_{i=1}^{\infty} \frac{2b_i}{3^i}$, where the a_i's and b_i's are either 0 or 1. By Lemma 4 either $a_1 = 0$ and $b_1 = 1$ or there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k = 0 < b_k = 1$. Note that $g(x) = \sum_{i=1}^{\infty} \frac{a_i}{2^i}$ and $g(y) = \sum_{i=1}^{\infty} \frac{b_i}{2^i}$. Hence if $a_1 = 0$ and $b_1 = 1$, then $g(x) = \frac{0}{2} + \sum_{i=2}^{\infty} \frac{a_i}{2^i} \leq \frac{1}{2} + \sum_{i=2}^{\infty} \frac{b_i}{2^i} = g(y)$. If there exists integer $k \geq 2$ such that $a_i = b_i$ for all $1 \leq i \leq k-1$ and $a_k = 0 < b_k = 1$, then

$$g(x) = \sum_{i=1}^{k-1} \frac{a_i}{2^i} + \frac{0}{2^k} + \sum_{i=k+1}^{\infty} \frac{a_i}{2^i} \leq \sum_{i=1}^{k-1} \frac{a_i}{2^i} + \frac{1}{2^k} + \sum_{i=k+1}^{\infty} \frac{b_i}{2^i} = g(y).$$

Hence, if $x < y$ and x and y are in C_0, then $g(x) \leq g(y)$. Suppose now that $x < y$, $x \notin C_0$ and y is in C_0. Then $x \in I(i, j) = (a(i, j), b(i, j))$ for some (i, j). Obviously $x < y$ implies that $a(i, j) < y$ and so since both $a(i, j)$ and y are in C_0 by what we have just proved, $g(a(i, j)) \leq g(y)$. Therefore, by definition of g, $g(x) = g(a(i, j)) \leq g(y)$.

Similarly, if $x < y$, $y \notin C_0$ and x is in C_0, then $y \in I(i, j) = (a(i, j), b(i, j))$ for some (i, j) and $x < b(i, j)$. Again since both x and $b(i, j)$ are in C_0, we have $g(x) \leq g(b(i, j)) = g(y)$.

Suppose now $x < y$, $x, y \notin C_0$. Then for some (i, j) and (i', j'), $x \in I(i, j)$, $y \in I(i', j')$. Then $a(i, j) < x < y < b(i', j')$. Therefore, since both $a(i, j)$ and $b(i', j')$ are in C_0, $g(a(i, j)) \leq g(b(i', j'))$ and so $g(x) = g(a(i, j)) \leq g(b(i', j')) = g(y)$. Hence g is monotonically increasing. This completes the proof.

Next we shall state a general result concerning bounded monotone function whose range is an interval.

Theorem 6. Suppose f: $[a, b] \to \mathbf{R}$ is an increasing function. Then f is continuous, if and only if, the range of f, $J = f([a, b])$ is an interval.

Proof. If f is continuous, then by the Intermediate Value Theorem, the range $f([a, b])$ is an interval. Now if f is increasing, then by Theorem 2 of *Monotone Function, Function of Bounded Variation, Fundamental Theorem of Calculus,* the discontinuity of f can only be jump discontinuity. Suppose the range J is an interval. Suppose f is discontinuous at $x = k$ in (a, b). Then we have

$$\lim_{x\to k^-} f(x) = f(k^-) \leq f(k) \leq f(k^+) = \lim_{x\to k^+} f(x) \text{ and } f(k^-) < f(k^+).$$

Note that for any $y < k$, because f is increasing $f(y) \leq f(k^-)$. Also for any $z > k$, $f(k^+) < f(z)$. Therefore, $(f([a, k)\cup(k, b]))\cap(f(k^-), f(k^+)) = \varnothing$. But by assumption the range J is an interval and so $(f(k^-), f(k^+)) \subseteq J$ and $(f([a, k)\cup(k, b]))\cap(f(k^-), f(k^+)) = (J - \{f(k)\})\cap(f(k^-), f(k^+)) \neq \varnothing$. This contradicts $(f([a, k)\cup(k, b]))\cap(f(k^-), f(k^+)) = \varnothing$ and so we have $f(k^-) = f(k^+) = f(k)$ and so $\lim_{x\to k} f(x) = f(k)$ and that means f is continuous at $x = k$. We can similarly derive a contradiction when $k = a$ *or* b. If f is not continuous at $k = a$, then $f(a) < f(a^+)$ and so $(f(a), f(a^+))\cap J = \varnothing$, contradicting $(f(a), f(a^+)) \subseteq J$. Thus f must be continuous at $x = a$. If f is not continuous at $k = b$, then $f(b^-) < f(b)$ and so $(f(b^-), f(b))\cap J = \varnothing$, contradicting $(f(b^-), f(b)) \subseteq J$. Thus, f must be continuous at $x = b$. Therefore, f cannot have any discontinuity and so it is continuous. This completes the proof.

Proposition 7. The function $g:[0, 1] \to [0, 1]$ is continuous.

Proof. By Proposition 5, the map g is onto and increasing. Since the range [0, 1] is an interval, by Theorem 6, g is continuous.

Next we shall reveal some interesting facts concerning the Cantor set C_0 .

Theorem 8. The Cantor set C_0 is

(1) compact,

(2) no where dense, i.e., it contains no open intervals,

(3) its own boundary points,

(4) perfect, i.e.,it is its own set of accumulation point,

(5) totally disconnected and

(6) between any two points in C_0 , there is an open interval not contained in C_0.

Proof. (1) Since C_0 is closed and bounded, it is compact by the Heine-Borel Theorem.

(2) Let $G_p = \bigcup_{k=1}^{p} \bigcup_{j=1}^{2^{k-1}} I(k,j)$ denote the union of the collection of disjoint open intervals deleted after the p-th stage in the construction. Suppose C_0 contains an open interval, then there is a point x in C_0 and a $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq C_0$. Hence $(x - \delta, x + \delta)\cap G = (x - \delta, x + \delta)\cap([0, 1] - C_0) = \varnothing$. Since $(x - \delta, x + \delta)$ and G are disjoint and are contained in [0,1], $\mu(G) + \mu((x - \delta, x + \delta)) \leq 1$. Hence $1 + 2\delta \leq 1$

implies $\delta \le 0$, contradicting $\delta > 0$. Therefore, C_0 cannot contain any open interval and so is nowhere dense. (We can also use $\mu((x-\delta, x+\delta)) \le \mu(C_0) = 0$ and so $2\delta \le 0$ contradicting $\delta > 0$.)

(3). C_0 is closed, by a characterization of closed set, because $C_0 = [0, 1] - G = \bigcap_{p=1}^{\infty}([0,1]-G_p)$ and each $[0, 1] - G_p$ is closed. Hence the closure of C_0 is C_0 and so its set of cluster points C_0' is contained in C_0. Let x be in $[0, 1]$. Then for any open set J containing x, say $J = U \cap [0, 1]$, where U is an open interval, J is non empty and contains more than one point and $J \cap G \neq \varnothing$ because $J \not\subset [0, 1] - G = C_0$ by part (2). This is because if $J \subseteq C_0$, then the interior of J, which is a non empty open interval is contained in C_0 contradicting part (2). Thus the closure of G, $\bar{G} = [0, 1]$. Therefore, the boundary of C_0, $\partial C_0 = \bar{C}_0 \cap \bar{G} = C_0 \cap [0,1] = C_0$.

(4) Since C_0 is closed the set of accumulation points of C_0, C_0' is a subset of C_0. We claim that $C_0' = C_0$. Let $x \in C_0$. Suppose $x \notin C_0'$. Then there exists $\delta > 0$, such that $(x-\delta, x+\delta) \cap C_0 = \{x\}$. Thus $(x-\delta, x+\delta) \cap ([0, 1] - C_0) = (x-\delta, x) \cup (x, x+\delta)$. That means $(x-\delta, x) \subseteq ([0, 1] - C_0) = G = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{2^{k-1}} I(k,j)$ and so

$$(x-\delta, x) = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{2^{k-1}} I(k,j) \cap (x-\delta, x).$$

Since the collection $\{I(k,j) : k = 1, \cdots, \infty; j = 1, 2, \cdots, 2^{k-1}\}$ is a collection of disjoint open intervals, and $(x-\delta, x)$ is connected, $(x-\delta, x) \subseteq I(i,j)$ for some i and j. This is because otherwise $(x-\delta, x)$ would be a disjoint union of open intervals contradicting that it is connected. Since $x \notin I(i,j)$, $x = \sup I(i,j) = b(i,j)$. Similarly,

$$(x, x+\delta) = \bigcup_{k=1}^{\infty}\bigcup_{j=1}^{2^{k-1}} I(k,j) \cap (x, x+\delta).$$

We deduce as before then that for some p and q, $(x, x+\delta) \subseteq I(p, q)$. Thus because $x \notin I(p, q)$, $x = \inf I(p, q) = a(p, q)$. Therefore, $a(p, q) = x = b(i,j)$, contradicting that $a(p, q) \neq b(i,j)$ by virtue of the definition of $I(i,j)$'s . Therefore $x \in C_0'$. Thus $C_0' = C_0$.

(5) Since by (2) C_0 does not contain any open interval and since the connected subsets of **R** are either the singleton sets or the intervals, the only components of C_0 are the singletons $\{x\}$, $x \in C_0$. Therefore, C_0 is totally disconnected.

(6) Suppose $x < y$ and x , y are in C_0. Then $(x, y) \cap ([0, 1] - C_0) \neq \varnothing$ and is a disjoint union of open intervals and so (x, y) contains at least one open interval.

The Cantor Set C_k

Now we turn our attention towards constructing a Cantor set of positive measure in [0, 1]. Indeed the construction also applies to give a different Cantor set of measure zero. The procedure of deleting the middle portion of each of the remaining closed sets is followed here but using a different specified length. (This idea can be generalised by not requiring the intervals to be deleted to be the middle portion, to give a *generalised Cantor set*.) Hence the construction is canonical. Let k be any real number with $0 \le k < 1$. Let $\delta = (1 - k) > 0$. Start with the closed unit interval [0, 1]. Let $I(1,1)$ be the middle open interval of length $\delta/2$. Then if we write $I(1,1) = (a(1,1), b(1,1))$, then $a(1,1) = 1/2 - \delta/4$. Then let $G_1 = I(1,1)$ and $F_1 = [0,1] - G_1$. The first stage is to form F_1 . Then F_1 is the disjoint union of 2 closed intervals each of length equal to $a(1,1) = 1/2 - \delta/4$. Let $I(2,1)$ and $I(2,2)$ be respectively the middle open intervals, each of length $\delta/2^3$, of the 2 closed intervals of F_1 . The open intervals are ordered from the left to the right by the second indices. Let $G_2 = I(1,1) \cup I(2,1) \cup I(2,2)$. Suppose $I(2,1) = (a(2,1), b(2,1))$. Then $a(2,1) = 1/2\, a(1,1) - \delta/2^4 = 1/2^2 - \delta(1/2^2 - 1/2^4)$. The second stage is to form $F_2 = [0,1] - G_2 = [0,1] - G_1 - I(2,1) \cup I(2,2)$. Then F_2 is the disjoint union of 2^2 closed intervals each of length equaling $a(2,1) = 1/2^2 - \delta(1/2^2 - 1/2^4)$. Then let $I(3, j), j = 1,..,2^2$ be respectively the middle portion of length $\delta/2^5$ of the 2^2 closed intervals again ordered by the second indices, from left to right in the sense that $I(3, j) < I(3, k)$, if and only if, $j < k,$ if and only if,,there exist x in $I(3, j)$ and y in $I(3, k)$ such that $x < y$. Stage three is to remove from F_2 these 2^2 open intervals, that is to form $F_3 = F_2 - \cup\{ I(3, j), j = 1,..,2^2 \} = [0, 1] - G_3$, where $G_3 = \bigcup_{k=1}^{3} \bigcup_{j=1}^{2^{k-1}} I(k,j)$. Thus F_3 consists of 2^3 disjoint closed intervals, each of length equaling $a(3,1)$, where $I(3,1) = (a(3,1), b(3,1))$. It is easily seen that $a(3,1) = 1/2^3 - \delta(1/2^3 - 1/2^6)$. Let $H_4 = \bigcup_{j=1}^{2^{4-1}} I(4,j)$, where $I(4, j), j = 1,..,2^3$ are the middle open intervals each of length $\delta/2^7$, of each of the disjoint closed intervals. Continuing in this way, at the n-th stage we have $H_n = \bigcup_{j=1}^{2^{n-1}} I(n,j)$, where $I(n, j) = (a(n, j), b(n, j))$ is an open interval of length $\delta/2^{2n-1}$ and the $I(n, j)$'s are ordered from left to right by the second indices according to the natural ordering of elements in the intervals. Note that $a(n,1) = 1/2\, a(n-1,1) - \delta(1/2^{2n}) = 1/2^n - \delta(1/2^n - 1/2^{2n})$. We obtain F_n by deleting from F_{n-1} the uion of open intervals H_n. Hence $F_n = F_{n-1} - H_n = [0,1] - G_n$, where $G_n = \bigcup_{k=1}^{n} \bigcup_{j=1}^{2^{k-1}} I(k,j)$ and F_n is a disjoint union of 2^n closed intervals each of length equaling $a(n,1) = 1/2^n - \delta(1/2^n - 1/2^{2n})$. To obtained F_{n+1} , we shall delete from F_n open intervals each of length equaling $\delta/2^{2n+1}$ from the 2^n closed intervals. That is, if we let $H_{n+1} = \bigcup_{j=1}^{2^n} I(n+1,j)$, where $I(n+1, j),\ j = 1,..,2^n$ are these open intervals to be deleted, then $F_{n+1} = F_n - H_{n+1} = [0, 1] - G_{n+1}$, where $G_{n+1} = \bigcup_{k=1}^{n+1} \bigcup_{j=1}^{2^{k-1}} I(k,j) = \bigcup_{k=1}^{n+1} H_k$. Obviously F_{n+1} is a collection of disjoint closed intervals, each of length equal to $a(n+1,1)$. Note also that for each $n \ge 1$, $F_{n+1} \subseteq F_n$.

Let $G = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j) = \bigcup_{k=1}^{\infty} H_k$. Then G is a disjoint union of open intervals. The length or measure of H_k is given by $2^{k-1}\dfrac{\delta}{2^{2k-1}} = \dfrac{\delta}{2^k}$. Thus the measure or length of G,

$$\mu(G) = \sum_{k=1}^{\infty} \mu(H_k) = \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \frac{\delta}{2} \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = \frac{\delta}{2} \lim_{k\to\infty} \left(\frac{1-\left(\frac{1}{2}\right)^k}{1-\left(\frac{1}{2}\right)} \right) = \delta.$$

The Cantor set C_k is defined to be the complement of G in [0, 1]. That is, $C_k = [0,1] - G = [0,1] - \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{2^{k-1}} I(k,j) = [0,1] - \bigcup_{k=1}^{\infty} G_k = \bigcap_{k=1}^{\infty}([0,1] - G_k) = \bigcap_{k=1}^{\infty} F_k.$ Thus the measure of C_k is equal to $\mu([0,1]) - \mu(G) = 1 - \delta = k$.

Then C_k satisfies the properties stated in Theorem 8.

Theorem 9. The Cantor set C_k defined above for $0 \le k < 1$ is uncountable and

(1) compact,

(2) no where dense, i.e., it contains no open intervals,

(3) its own boundary points,

(4) perfect, i.e., is its own set of accumulation point,

(5) totally disconneceted and

(6) between any two points in C_k , there is an open interval not contained in C_k.

Proof. Part (1) and part (3) to part(6) are proved in exactly the same way as in Theorem 8. Part (2) needs a different approach to prove. Suppose C_k contains an open interval say (c, d). Then $(c,d) \subseteq C_k = \bigcap_{k=1}^{\infty} F_k$ implies that $(c, d) \subseteq F_k$ for each $k \ge 1$. Now since $\lim_{k\to\infty} (\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}})) = 0$, there exists an integer N such that $k > N$ implies that $\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}}) < d - c$. Now we fix a $k > N$. Note first that $(c, d) \cap F_k = (c, d)$. Since the non-trivial interval (c, d) is connected and F_k is a disjoint union of closed intervals, (c, d) must be contained in one of these closed intervals, each of length $\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}})$. Hence, $d - c \le \frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}})$ contradicting $\frac{1}{2^k} - \delta(\frac{1}{2^k} - \frac{1}{2^{2k}}) < d - c$. Therefore, C_k does not contain an open interval. (The main thrust of the argument is that the F_k consists of disjoint closed intervals, whose maximum length tends to zero as k tends to infinity.) This proves part (2). That C_k is uncountable is a consequence of the following proposition.

Proposition 10. There is a continuous strictly increasing bijective map $f : [0, 1] \to [0,1]$ mapping the Cantor set C_k onto the the Cantor set C_0 defined earlier using the "deleted middle third intervals" construction.
Proof. This is given by Lemma 1 of *Composition and Riemann Integrability.* Note that the proof given there applies also to the case $k = 0$.

Since C_0 constructed using the "deleted middle third intervals" construction is uncountable and the function f given by Proposition 10 maps C_k bijectively onto C_0, C_k is also uncountable.