0.0.1 The Cantor Set

Construction of the Cantor Set

Let $\mathcal{C}_0 = [0,1]$, $\mathcal{C}_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$, $\mathcal{C}_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{1}{3}] \cup [\frac{2}{3},\frac{7}{9}] \cup [\frac{7}{9},1]$, and in general, let $\mathcal{C}_{n+1}$ be the union of the 2^{n+1} closed intervals, each of length $(\frac{1}{3})^{n+1}$, obtained by removing the open middle thirds of the 2^n closed intervals of $\mathcal{C}_n$. We define the Cantor set $\mathcal{C}$ to be the intersection of all the $\mathcal{C}_n$; $\mathcal{C} = \cap_{n=0}^{\infty} \mathcal{C}_n$. Another way to describe this is to say that $\mathcal{C}$ is the set of points in $[0,1]$ that remain after removing the open middle third interval $(\frac{1}{3},\frac{2}{3})$, and then removing the open middle thirds from the remaining two closed intervals $[0,\frac{1}{3}]$, $[\frac{2}{3},1]$, and then removing the open middle thirds from the remaining four closed intervals, etc ad infinitum. Note that the sets $\mathcal{C}_n$ are approximations of $\mathcal{C}$ in the sense that $\lim_{n\to\infty} \mathcal{C}_n = \mathcal{C}$ so we can get an impression of what $\mathcal{C}$ looks like by looking at $\mathcal{C}_n$ as n gets large (see Figure 2 and Exercise 1). figure exercise

This construction of $\mathcal{C}$ removes infinitely many intervals from $[0,1]$, so we might wonder if there are any points in $\mathcal{C}$. Some obvious points are the end points of the open middle third intervals that were removed; $\{0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \ \dots \}$. So there are at least a countably infinite number of points in $\mathcal{C}$. In fact, we will see that there are *many* more points in $\mathcal{C}$ than these. But let's first describe some of the properties of the Cantor set.

Since $[0,1] = \mathcal{C} \cup \{\text{intervals removed}\}$ (a disjoint union, notice), the length of $\mathcal{C} = 1 -$ (total length of the intervals removed). The length removed in the first stage is $\frac{1}{3}$, the length removed in the second stage is $2 \cdot (\frac{1}{3})^2$, the length removed in the third stage is $2^2 \cdot (\frac{1}{3})^3$, etc, so the total length of the intervals removed is $\sum_{n=1}^{\infty} 2^{n-1}(\frac{1}{3})^n = \frac{1}{2}\sum_{n=1}^{\infty}(\frac{2}{3})^n = \frac{1}{2}\left(\frac{2/3}{1/3}\right) = \frac{1}{2}(2) = 1$. Thus, the length of $\mathcal{C}$ is 0. This implies that $\mathcal{C}$ cannot contain any intervals, i.e., that is is 'dust' (between any two points in $\mathcal{C}$ is a point that is not in $\mathcal{C}$).

We've noted that the end points of the intervals removed during the construction are in $\mathcal{C}$, but what other points are in $\mathcal{C}$, if any? It's very difficult to see what other points are in $\mathcal{C}$ by relying on this geometric construction. For example, can you see why the numbers $\frac{1}{4}$ and $\frac{3}{4}$ are in $\mathcal{C}$? They are not the end points of any interval that was removed, yet they are never removed in the construction of $\mathcal{C}$. To see exactly what numbers are in $\mathcal{C}$, it's much more convenient to represent numbers in a way that reflects the structure of $\mathcal{C}$. Going back to the construction, note that the points in $\mathcal{C}_1$ are precisely the numbers x in $[0,1]$ that have no 1 in the first place of their ternary expansion[1] $[x]_3$. (Here we need to resolve the ambiguity of which ternary expansion to choose for those numbers that have two expansions. First note that in base 3, $0.1 = 0.0\overline{22}$, and more generally, $0.\underbrace{00\cdots0}_{k}\overline{22} = 0.\underbrace{00\cdots0}_{k-1}1\overline{00}$. In these cases we choose the expansion that contains only $2's$. For example we choose $[\frac{1}{3}]_3 = 0.0\overline{22}$ (rather than 0.1) and $[\frac{2}{3}]_3 = 0.2$ (rather than $0.1\overline{22}$). The numbers that have two such ternary expansions are numbers of the form $\frac{1}{3^n}$ and $\frac{2}{3^n}$.) Similarly, $\mathcal{C}_2$ are the numbers in $[0,1]$ that have no 1 in either the first or second places of their ternary expansion. So we see that $\mathcal{C}_n$ is

[1] See Appendix X for a discussion of ternary, and other base, expansions

precisely the numbers in $[0,1]$ that have no 1 in any of the first n places of their ternary expansion. Since $\mathcal{C}$ is the limit of the $\mathcal{C}_n$, the numbers in $\mathcal{C}$ are the numbers in $[0,1]$ that have no 1 in their ternary expansion; see Figure 3. figure

For example, $[\frac{1}{4}]_3 = 0.02\overline{02}$ and $[\frac{3}{4}]_3 = 0.2\overline{02}$ so both $\frac{1}{4}$ and $\frac{3}{4}$ are in $\mathcal{C}$. Moreover, if $\vec{a} = b_1 b_2 b_3 \ldots$ is a sequence of $0'^s$ and $2'^s$, then the number x with $[x]_3 = 0.\vec{a}$ is a number in $\mathcal{C}$. Just how many of these numbers are there? To answer this we observe that we can match elements in the set $\mathcal{B}$ of all sequences (binary expansions) of the form $0.\vec{b}$, where $\vec{b} = b_1 b_2 b_3 \cdots$ is a sequence of $0'^s$ and $1'^s$, with numbers in $[0,1]$ via binary expansions (however, not in a one-to-one manner due to the non-uniqueness of some expansions mentioned in the previous paragraph[2]; see also Exercise 1.2.1). That is, if $\vec{b} = b_1 b_2 b_3 \cdots$ is any sequence of exercise $0'^s$ and $1'^s$, then there is a unique number $x \in [0,1]$ such that $[x]_2 = 0.\vec{b}$, in fact $x = \frac{b_1}{2} + \frac{b_2}{2^2} + \frac{b_3}{2^3} + \cdots$. So the cardinality of $[0,1]$ is at least as large as the cardinality of $\mathcal{B}$. The matching $x \in [0,1] \mapsto [x]_2 \in \mathcal{B}$ (choosing one of the two expansions if x has two binary expansions) shows that the cardinality of $\mathcal{B}$ is at least as large as that of $[0,1]$. Hence the cardinality of $[0,1]$ is the same as that of $\mathcal{B}$.

Now, the set $\mathcal{S}$ of all sequences of the form $0.\vec{a}$ where $\vec{a}$ is a sequence of $0'^s$ or $2'^s$ has the same cardinality as the set $\mathcal{B}$; just match each $0.\vec{b} \in \mathcal{B}$ with an element $0.\vec{a} \in \mathcal{S}$ by changing every 1 in $\vec{b}$ to a 2, and visa versa, match each element $0.\vec{a} \in \mathcal{S}$ with an element $0.\vec{b} \in \mathcal{B}$ by changing every 2 in $\vec{a}$ to a 1. Since $\mathcal{B}$ has the same cardinality as $[0,1]$, and since $\mathcal{S}$ has the same cardinality as $\mathcal{C}$ (same argument as for $\mathcal{B}$ and $[0,1]$), the Cantor set $\mathcal{C}$ has the same cardinality as the interval $[0,1]$! This seems bizarre because in some sense $\mathcal{C}$ is a 'small' subset of $[0,1]$ (it is a proper subset of length 0). This shows you that by 'rearranging' the points in $[0,1]$ we can obtain a set of length zero (there are also 'generalized Cantor sets' which have lengths anywhere between 0 and 1, so more generally we can rearrange the points in $[0,1]$ to obtain a set of any length between 0 and 1, including 0 and 1; cf. Excercise xx)[3]. exercise

How are the numbers in $[0,1]$ rearranged to obtain $\mathcal{C}$? The discussion in the previous paragraph explained how we could determine the cardinality of $\mathcal{C}$ by matching each number x in the interval $[0,1]$ with a number y in $\mathcal{C}$ in a one-to-one manner;

$$[0,1] \ni x \mapsto \overbrace{0.\vec{b} = [x]_2 \mapsto 0.\vec{a}}^{\text{change 1's to 2's}} \in \mathcal{S} \mapsto y = \frac{a_1}{3} + \frac{a_2}{3^2} + \cdots \in \mathcal{C} \tag{1}$$

(we make the convention that if x has two binary expansions we take the one that ends in zeros). If you look more closely at this matching, you'll see that some numbers in $\mathcal{C}$ are actually missed (so this mapping is not one-to-one). For example, $\frac{1}{3} \in \mathcal{C}$ is not matched with any number in $[0,1]$; $[\frac{1}{3}]_3 = 0.0\overline{22}$, but $\frac{1}{2}$ (where we choose $[\frac{1}{2}]_2 = 0.1$) is mapped to $\frac{2}{3} \in \mathcal{C}$; see exercise xxx. So this matching is actually onto a proper exercise subset of $\mathcal{C}$ (which is sufficient to prove that the cardinality of $\mathcal{C}$ is at least as large as the cardinality of

[2]The non-one-to-oneness of the matching $\mathcal{B} \to$ [0,1] occurs 'only' for a countable number of points in $[0,1]$. Since $[0,1]$ is uncountable, this does no affect the conclusion of our argument here. See Appendix X.

[3]A simpler way to rearrange the points in [0,1] to obtain a set of small length is given below.

[0, 1]). However, 'most' of the numbers of $\mathcal{C}$ are matched with a number in [0, 1] (exercise xx) so this way of matching the two sets gives us a good impression of how to rearrange [0, 1] to produce $\mathcal{C}$. exercise

Generally, we can represent any rearrangement of [0, 1] by drawing the graph of the function $\varphi(x)$ that represents the rearrangement (i.e., $\varphi(x) = y$ means the rearrangement moves x to y). Now, it's no mystery how one can rearrange [0, 1] to obtain a set of small length. Let ε be any small positive number. Then the function $\varphi_\varepsilon(x) = \varepsilon x$ rearranges [0, 1] into a set of length ε, namely the set $[0, \varepsilon]$. Notice that the slope of the graph of $\varphi_\varepsilon(x)$ is small; the slope of the graph of any function that rearranges [0, 1] into a set of small length must necessarily be small. Since the length of $\mathcal{C}$ is zero, the graph of the function $\varphi(x)$ that represents the rearrangement of [0, 1] into $\mathcal{C}$ must in some sense have zero slope. But the graph of this function cannot be flat on any interval because we know that if $x_1 \neq x_2$, then $\varphi(x_1) \neq \varphi(x_2)$. So it's not obvious what the graph of this function looks like; it begins at (0, 0), ends at (1, 1), its range is $\mathcal{C}$ (so if you projected the graph onto the y-axis it would be $\mathcal{C}$), and is 'flat'!

To get an idea of what that graph looks like, let's define a function $\varphi_\mathcal{C}(x)$ which matches the numbers in [0, 1] to numbers in $\mathcal{C}$ as described above in equation (2.1). Since 'most' numbers in $\mathcal{C}$ are matched in this way with a number in [0, 1], the graph of $\varphi_\mathcal{C}(x)$ will give an accurate impression of the way [0, 1] is rearranged to make $\mathcal{C}$.

Figure 4 shows the graph of $\varphi_\mathcal{C}(x)$. It was obtained by taking $x \in [0, 1]$, computing $[x]_2$, changing every 1 in $[x]_2$ to a 2, then summing up the resulting ternary expansion to obtain $y = \varphi_\mathcal{C}(x)$. If you look closely you'll see that the graph appears to be flat everywhere, but also has lots of jumps. The jumps are precisely at the points x where $x = (\frac{m}{2^n})$ for some positive integer n, and positive integer $m < 2^n$ (these are the numbers which have a binary expansion that ends in zeros). Note that these points are dense in [0, 1], so the graph of $\varphi_\mathcal{C}(x)$ has a jump almost everywhere, and is 'flat' everywhere else. figure

Let $\mathcal{E} = \{0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \ \ldots\}$ be the set of edges of the intervals removed in the construction of the Cantor set $\mathcal{C}$.

Claim: $\mathcal{E}$ is a 'small' subset of $\mathcal{C}$, i.e., 'most' of the numbers in $\mathcal{C}$ are not edge points.

Proof: If $x \in \mathcal{E}$, then $[x]_3$ ends in $\overline{00}$ because $x = \frac{m}{3^n}$ for some positive integer $m < 3^n$ (exercise xx).If $\mathcal{S}_o$ is the subset of $\mathcal{S}$ of sequences that end in $\overline{00}$, then $\mathcal{S}_0$ is a 'small' subset of $\mathcal{S}$ in the sense that the cardinality of $\mathcal{S} \setminus \mathcal{S}_0$ is the same as the cardinality of $\mathcal{S}$ (exercise xx) $\square$ exercise exercise

So we could have removed the *closed* middle thirds in the construction of $\mathcal{C}$ and still have obtained essentially $\mathcal{C}$. However, the set $\mathcal{E}$ shows us where the points of $\mathcal{C}$ are.

Claim: The set $\mathcal{E}$ is dense in $\mathcal{C}$; $\overline{\mathcal{E}} = \mathcal{C}$.

Proof: Exercise xx. exercise

In other words, the edge points $\mathcal{E}$ accumulate to $\mathcal{C}$; if $x \in \mathcal{C}$ is any point in the Cantor set, then there is an infinite sequence of points from $\mathcal{E}$ that converge to x. So although $\mathcal{E}$ is a negligibly small subset of $\mathcal{C}$, the edge points do show us exactly where the points in $\mathcal{C}$ are, and so sketching the edge points gives us an accurate impression of what $\mathcal{C}$ looks like (however, sketching $\mathcal{E}$ is no easy task!).

Summary of properties of the Cantor set:

- the length of $\mathcal{C}$ is zero
- $\mathcal{C}$ is totally disconnected (is 'dust')
- $\mathcal{C}$ is a closed set (Exercise xx) exercise
- $\mathcal{C}$ has the same cardinality as $[0, 1]$
- every point in $\mathcal{C}$ is a limit of a sequence of end points $\mathcal{E}$
- $\mathcal{C}$ is self-similar

Some Useful Formulae

Geometric series

For any number $r \neq 1$,

$$\begin{aligned} 1 + r + r^2 + r^3 + \cdots + r^n &= \sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1 - r}, \quad \text{and} \\ r + r^2 + r^3 + \cdots + r^n &= \sum_{i=1}^{n} r^i = \frac{r - r^{n+1}}{1 - r} \end{aligned}$$

Thus, if $|r| < 1$ then taking the limit $n \to \infty$,

$$\begin{aligned} 1 + r + r^2 + r^3 + \cdots &= \sum_{i=0}^{\infty} r^i = \frac{1}{1 - r}, \quad \text{and} \\ r + r^2 + r^3 + \cdots &= \sum_{i=1}^{\infty} r^i = \frac{r}{1 - r} \end{aligned}$$

Decimal, binary, and ternary expansions

Let x be a number in $[0,1]$, a a positive integer, and $[x]_a$ denote the expansion of x in base a. The relations between $[x]_a$ and x are as follows.

- **Decimal expansion** ($a = 10$):

$$\begin{aligned} \text{if } [x]_{10} &= 0.d_1 d_2 d_3 \ldots, \quad \text{where } d_i = 0, 1, 2, \ldots, 9, \\ \text{then } x &= \frac{d_1}{10} + \frac{d_2}{10^2} + \frac{d_3}{10^3} + \cdots \end{aligned}$$

- **Binary expansion** ($a = 2$):

$$\begin{aligned} \text{if } [x]_{2} &= 0.b_1 b_2 b_3 \ldots, \quad \text{where } b_i = 0 \text{ or } 1, \\ \text{then } x &= \frac{b_1}{2} + \frac{b_2}{2^2} + \frac{b_3}{2^3} + \cdots \end{aligned}$$

- **Ternary expansion** ($a = 3$):

$$\begin{aligned} \text{if } [x]_{3} &= 0.c_1 c_2 c_3 \ldots, \quad \text{where } c_i = 0, 1 \text{ or } 2, \\ \text{then } x &= \frac{c_1}{3} + \frac{c_2}{3^2} + \frac{c_3}{3^3} + \cdots \end{aligned}$$

For numbers x greater than 1 there would be digits to the left of the point corresponding to positive powers of a.

Note that in any base some numbers will have two expansions. For example $[\frac{1}{10}]_{10} = 0.1$ and $0.0\overline{99}$, $[\frac{3}{4}]_2 = 0.11$ and $0.10\overline{11}$, $[\frac{7}{9}]_3 = 0.021$ and $0.020\overline{22}$. This can be seen by using the formulae above for adding geometric series.

Here's an algorithm for computing the ternary expansion of a number $x \in [0,1]$ (there's a similar algorithm if the number is greater than 1 and for other bases). We want to determine the c_i, $c_i \in \{0,1,2\}$, such that $x = c_1/3 + c_2/3^2 + c_3/3^3 + \cdots$

- Determine c_1;
  - if $0 \le x < 1/3$, then $c_1 = 0$
  - if $1/3 \le x < 2/3$, then $c_1 = 1$
  - if $2/3 \le x < 1$, then $c_1 = 2$
  - let $x_1 = x - c_1/3$ (note that $x_1 < 1/3$)

- Determine c_2;
    - if $0 \le x_1 < 1/3^2$, then $c_2 = 0$
    - if $1/3^2 \le x_1 < 2/3^2$, then $c_2 = 1$
    - if $2/3^2 \le x_1 < 1/3$, then $c_2 = 2$
    - let $x_2 = x_1 - c_2/3^2$ (note that $x_2 < 1/3^2$)
- Determine c_3;
    - if $0 \le x_2 < 1/3^3$, then $c_3 = 0$
    - if $1/3^3 \le x_2 < 2/3^3$, then $c_3 = 1$
    - if $2/3^3 \le x_2 < 1/3^2$, then $c_3 = 2$
    - let $x_3 = x_2 - c_3/3^3$ (note that $x_3 < 1/3^3$)
- etc.

Here's a pseudocode that can be implemented on a computer to calculate the expansion of a number $x \in [0,1]$ in base b, $[x]_b = a_1 a_2 a_3 \ldots$

$$\begin{aligned}
a_1 &= \text{floor}(bx); \quad (\text{since } bx = a_1 + a_2/b + a_3/b^2 + \cdots \text{ and } a_2/b + a_3/b^2 + \cdots < 1) \\
a_2 &= \text{floor}(b^2x - ba_1)); \quad (\text{since } b^2x = ba_1 + a_2 + a_3/b + a_4/b^2 + \cdots \text{ and } a_3/b + a_4/b^2 + \cdots < 1) \\
a_3 &= \text{floor}(b^3x - b^2a_1 - ba_2); \\
a_4 &= \text{floor}(b^4x - b^3a_1 - b^2a_2 - ba_3); \\
\cdots & \quad \cdots \\
\cdots & \quad \cdots
\end{aligned}$$

Here, floor(x) (the floor function, sometimes denoted by $\lfloor x \rfloor$) is the largest integer less than or equal to x (eg., floor(2.412)=2, floor(3)=3, floor(-1.23)=-2).

This gives us a quick way to compute base b expansions of numbers. First multiply the number by b. Then a_1 is the integer part of this number. Now multiply the fractional part by b and take the integer part of this number; this is a_2. Now carry one like this. For example, let's compute the base 3 expansion of 1/7. First, $3(1/7) = 3/7$ so $a_1 = 0$. Then, $3(3/7) = 9/7 = 1 + 2/7$ so $a_2 = 1$. Then, $3(2/7) = 6/7$ so $a_3 = 0$. Then, $3(6/7) = 18/7 = 2 + 2/7$ so $a_4 = 2$. Now the pattern starts repeating (for this particular number), so $a_5 = 0, a_6 = 2, a_7 = 0, \ldots$. Thus, $[1/7]_3 = 0.01\overline{02}$.

<u>Some facts</u>:

Every rational number has an eventually repeating expansion in any base, and every irrational number has a non-repeating expansion in any base.

Suppose x and y are two number between 0 and 1. Let b be a positive integer and suppose the base b expansions of x and y agree to the k^{th} place, i.e.,

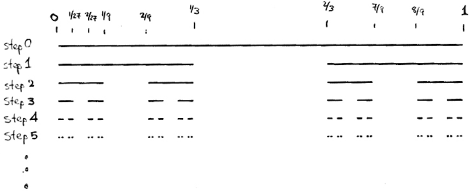
$$\begin{aligned} x &= 0.a_1 a_2 \cdots a_k a_{k+1} a_{k+2} \cdots \\ y &= 0.a_1 a_2 \cdots a_k e_{k+1} e_{k+2} \cdots \end{aligned}$$

where $e_{k+1} \neq a_{k+1}$. Then

$$\begin{aligned} x &= \frac{a_1}{b} + \frac{a_2}{b^2} + \cdots + \frac{a_k}{b^k} + r_x \\ y &= \frac{a_1}{b} + \frac{a_2}{b^2} + \cdots + \frac{a_k}{b^k} + r_y \end{aligned}$$

So $x - y = r_x - r_y$. Now note that both r_x and r_y are in $[0, 1/b^k]$ (since $(b-1)/b^{k+1} + (b-1)/b^{k+2} + (b-1)/b^{k+3} + \cdots = 1/b^k$; use the formula for a geometric sum), so the largest $|r_x - r_y| = |x - y|$ can be is $1/b^k$. That is, if two numbers in $[0, 1]$ have base b expansions that agree to the k^{th} place, then these two numbers are no further apart than $1/b^k$. Conversely, if $0.a_1 a_2 a_3 \cdots$ and $0.e_1 e_2 e_3 \cdots$ are two expansions in base b that agree to the k^{th} place, then the two numbers with these representations are no further apart than $1/b^k$. This is used several times in our discussions.

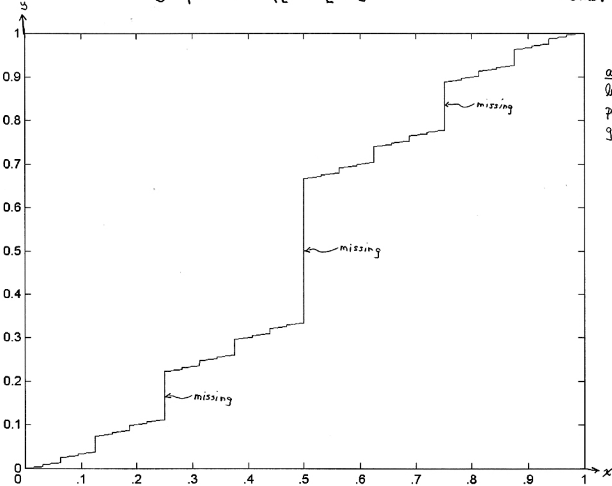
Construction of the Cantor Set



May 19

graph of $\varphi_c : [0,1] \longrightarrow \mathcal{C}$ The "Cantor Staircase"

range $(\varphi_c) = \mathcal{C}$



all vertical lines are **NOT** part of the graph of φ_c

$\varphi_c(x) = y$; $[0,1] \ni x \longmapsto [x]_2 \longrightarrow [y]_3 \longmapsto y \in \mathcal{C}$

↑ change "1's" to "2's"