The Cantor Set and the Cantor Function

TMA4225 - Foundations of Analysis

"Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line."
Benoit Mandelbrot, The Fractal Geometry of Nature

Definition of Cantor's Set

- Step 0: we begin with the interval $[0, 1]$.

- Step 1: we divide $[0, 1]$ into 3 subintervals and <u>delete</u> the *open* middle subinterval $(\frac{1}{3}, \frac{2}{3})$.

- Step 2: we divide each of the 2 resulting intervals above into 3 subintervals and <u>delete</u> the *open* middle subintervals $(\frac{1}{9}, \frac{2}{9})$ and $(\frac{7}{9}, \frac{8}{9})$.

We continue this procedure indefinitely. At each step, we delete the *open* middle third subinterval of each interval obtained in the previous step.

Definition of Cantor's Set

Cantor's set is the set C left after this procedure of deleting the open middle third subinterval is performed infinitely many times.

- Is there anything left?

Yes, at least the endpoints of the deleted middle third subintervals. There are countably many such points.

- Are there any *other* points left?

Yes, in some sense, a whole lot more. But in some other sense, just some dust - which in some ways is *scattered*, in some other ways it is *bound together*.

We will describe different ways to "measure" the dust left. This will take us through several mathematical disciplines: set theory, measure theory, topology, geometric measure theory, real analysis.

Ternary Representation of Cantor's Set

- Every real number can be represented by an infinite sequence of digits:

$$\begin{aligned} \frac{1}{3} &= 0.33333\ldots \\ \text{golden ratio} &= 1.6180339887498948482045\ldots \\ \frac{1}{10} &= 0.10000\ldots = 0.09999\ldots \end{aligned}$$

  This is the decimal (base **10**) representation:
  - ◇ Numbers described using powers of **10** and
  - ◇ Digits used: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
  - ◇ *Some* numbers can be represented in <u>two</u> ways, with one representation having only 9's from some point on.
- Computers use the binary (base **2**) representation: every number is described using powers of **2** and digits 0, 1.

$$\begin{aligned} \text{golden ratio} &= 1.10011110001101110111\ldots_{(2)} \\ \frac{1}{2} &= 0.1000\ldots = 0.011111\ldots_{(2)} \end{aligned}$$

Ternary Representation of Cantor's Set

We can represent real numbers in any base. We will use the *ternary* (base 3) representation, because Cantor's set has a special representation in base 3.

$$
\begin{aligned}
\frac{1}{3} &= 1 \cdot 3^{-1} = 0.10000\ldots_{(3)} = 0.022222\ldots_{(3)} \\
\frac{2}{3} &= 2 \cdot 3^{-1} = 0.20000\ldots_{(3)} \\
\frac{7}{9} &= 2 \cdot 3^{-1} + 1 \cdot 3^{-2} = 0.210000\ldots_{(3)} = 0.20222\ldots_{(3)} \\
\frac{8}{9} &= 2 \cdot 3^{-1} + 2 \cdot 3^{-2} = 0.220000\ldots_{(3)}
\end{aligned}
$$

A number is in Cantor's set if and only if its ternary representation contains only the digits 0 and 2 (in other words, it has no 1's).

$$\boxed{C = \{x \in [0,1] : x = 0.c_1 c_2 c_3 \ldots c_n \ldots_{(3)} \text{ where } c_n = 0 \text{ or } 2\}}$$

We already know that Cantor's set is <u>infinite</u>: it contains all endpoints of deleted intervals. There are only *countably* many such endpoints.

We will show that in fact Cantor's set has a much larger cardinality (i.e. "number" of elements).

Theorem: The cardinality of Cantor's set is the *continuum*. That is, Cantor's set has the same cardinality as the interval $[0, 1]$.

Proof.

The function $f : C \to [0, 1]$ defined by:

$$f(0.c_1 c_2 c_3 \dots c_n \dots_{(3)}) := 0.\frac{c_1}{2}\frac{c_2}{2}\frac{c_3}{2} \dots \frac{c_n}{2} \dots_{(2)}$$

is onto, so card $C \geq$ card $[0, 1]$.
But clearly card $C \leq$ card $[0, 1]$.
Then card $C =$ card $[0, 1]$ by Cantor-Bernstein-Schroeder theorem. $\square$

Measure Theory ⇝ Cantor's set is negligible

- One way to measure Cantor's set (by "counting" its elements) shows that it is a very large set - as large as the whole interval it is part of.
- Another way to measure it is by looking at the amount of space it occupies on the line.

Theorem: Cantor's set is *negligible*.
In other words, its "length" / Lebesgue measure is 0.

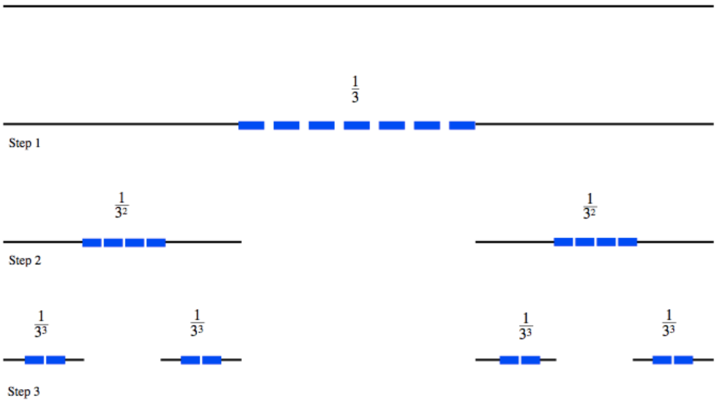
Proof.

Cantor's set is obtained by successively removing intervals.

We will measure the intervals removed.

At each step the number of intervals doubles and their length decreases by 3.

Measure Theory $\rightsquigarrow$ Cantor's set is negligible



Total length / measure of intervals removed:

$$\frac{1}{3} + 2 \cdot \frac{1}{3^2} + 2^2 \cdot \frac{1}{3^3} + \ldots = \sum_{n=0}^{\infty} 2^n \cdot \frac{1}{3^{n+1}} = \ldots = 1$$

"Length" / measure of Cantor's set $= 1 - 1 = 0$. $\square$

Topology $\rightsquigarrow$ Structure of Cantor's set

- **Theorem:** Cantor's set has no interior points / it is nowhere dense.

In other words, it is just "dust".

That's because its length is 0, so it contains no continuous parts (no intervals).

- **Theorem:** Cantor's set is bounded.

That's because it lives inside the interval $[0, 1]$.

- **Theorem:** Cantor's set is closed.

That's because it is the complement relative to $[0, 1]$ of open intervals, the ones removed in its construction.

$\rightsquigarrow$ Bounded + Closed on the real line $\Longrightarrow$ compact set
(Heine-Borel theorem)
Morally compact means that every task that may theoretically require an *infinite* number of steps/*infinite* amount of data, can be accomplished in a finite number of steps /with finite resources.

Topology $\rightsquigarrow$ Structure of Cantor's set

- **Theorem:** Cantor's set has no isolated points.

That is, in any neighborhood of a point in Cantors's set, there is another point from Cantor's set.

"Proof." Given say $a = 0.0220020202\ldots_{(3)} \in C$ one could find another element $b = 0.0220022202\ldots_{(3)} \in C$ which is near a. $\square$

$\rightsquigarrow$ In topology, a set which is compact and has *no isolated points* is called a perfect set

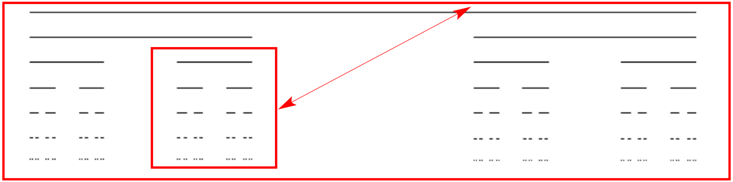
- **Theorem:** Cantor's set is totally disconnected.

In other words, given any two elements $a, b \in C$, Cantor's set can be divided into two disjoint and closed neighborhoods A and B, one containing a and the other containing b.

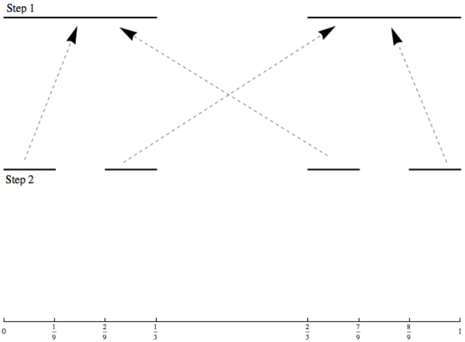
"Proof." Given say the numbers a and b from above:
Neighborhood $A =$ all elements of C whose 7th digit is 0.
Neighborhood $B =$ all elements of C whose 7th digit is 2.

Geometric Measure Theory ⇝ Cantor's Set is a Fractal

- **Theorem:** Cantor's set is self-similar.



More accurately: magnify Cantor's set by **3**, get **2** copies of itself.



Back to Real Analysis $\rightsquigarrow$ Cantor's Function

The function defined earlier, $f: C \to [0,1]$

$$f(0.c_1c_2c_3\ldots c_n\ldots_{(3)}) := 0.\frac{c_1}{2}\frac{c_2}{2}\frac{c_3}{2}\ldots\frac{c_n}{2}\ldots_{(2)}$$

has the following properties:

- It is onto.
- It is increasing
- It is not one to one. For instance:

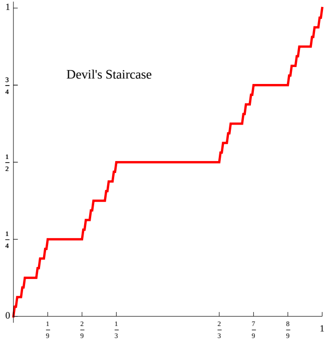
$$\begin{aligned} f(\frac{1}{3}) &= f(0.0222\ldots_{(3)}) = 0.01111\ldots_{(2)} = 0.1_{(2)} = \frac{1}{2} \\ f(\frac{2}{3}) &= f(0.2000\ldots_{(3)}) = 0.10000\ldots_{(2)} = 0.1_{(2)} = \frac{1}{2} \end{aligned}$$

Two inputs of f have the same outputs if and only if they are the endpoints of an interval removed - like $(\frac{1}{3}, \frac{2}{3})$ or $(\frac{1}{9}, \frac{2}{9})$ etc.

Back to Real Analysis $\rightsquigarrow$ Cantor's Function

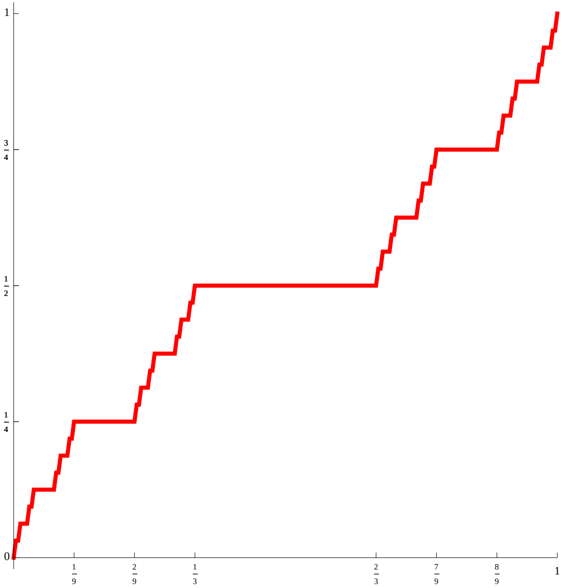
Extend f to the whole interval $[0, 1]$ by making it constant on these removed intervals.

The function obtained by this extension is called Cantor's function.



- Cantor's function is onto.
- Cantor's function is increasing, but constant almost everywhere (except on the "dust").
- Cantor's function is continuous.
- The derivative of Cantor's function is 0 almost everywhere.

The Devil's Staircase ...



... and Its Musical Counterpart

- Cantor's function, also called the Devil's Staircase, makes a continuous finite ascent (from 0 to 1) in an infinite number of steps (there are infinitely many intervals removed) while staying constant most of the time.
- Playing the following YouTube video (click the link):
  `http://youtu.be/1ZTaiDHqs5s?t=8s`
  you will hear the musical illustration of this ascent (followed by a descent and by more playing along the staircase). It is called "L'Escalier du Diable" (i.e. "The Devil's Staircase"). It was composed in the 90's by the composer György Ligeti.
- This composition is harmonically self-similar, like Cantor's set.
- Its structure has the musical equivalent of dividing an interval into three parts, changing the middle, repeating the process and matching the pitch arrangements to this structure, just like with Cantor's set and function.
- But unlike its mathematical counterpart, the process is finite, its infinity being only suggested.