# Analysis 1
# Colloquium of Week 4
# Cantor's Set–The Set without the Middle Thirds

Jacob Shapiro

October 8, 2014

**Abstract**

We present the Cantor set, its ternary representation, and show how this proves that it has the same Cardinality as $\mathbb{R}$. We follow Rudin and Koenigsberger.

## 1 Preliminaries

### 1.1 The Cantor Set

We follow the definition found in Walter Rudin's *Principles of Mathematical Analysis.*

#### 1.1.1 Definition

We define an infinite sequence of sets of real numbers. The intersection of all sets in the sequence *is* the Cantor set.

1. Define the first element in the sequence, the set $E_0 := [0, 1]$.
2. Define recursively, for all $n \in \mathbb{N}$, $E_n = \frac{1}{3}E_{n-1} \cup \left(\frac{1}{3}E_{n-1} + \frac{2}{3}\right)$.
   Observe that what this is doing is taking just two miniature (scaled by $\frac{1}{3}$) copies of the previous step, placing the first from 0 and the second from $\frac{2}{3}$.
3. The Cantor set is then defined as $P := \bigcap_{n \in \mathbb{N}} E_n$.

#### 1.1.2 A Few Examples

1. For $n = 1$ we have

$$\begin{aligned} E_1 &= \frac{1}{3}E_0 \cup \left(\frac{1}{3}E_0 + \frac{2}{3}\right) \\ &= \frac{1}{3}[0, 1] \cup \left(\frac{1}{3}[0, 1] + \frac{2}{3}\right) \\ &= \left[0, \frac{1}{3}\right] \cup \left(\left[0, \frac{1}{3}\right] + \frac{2}{3}\right) \\ &= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \end{aligned}$$

Note that $E_1 = E_0 \backslash \left(\frac{1}{3}, \frac{2}{3}\right)$.

2. For $n = 2$ we have

$$\begin{aligned}
E_2 &= \frac{1}{3}E_1 \cup \left(\frac{1}{3}E_1 + \frac{2}{3}\right) \\
&= \frac{1}{3}\left(\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]\right) \cup \left(\frac{1}{3}\left(\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]\right) + \frac{2}{3}\right) \\
&= \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right]\right) \cup \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] + \frac{2}{3}\right) \\
&= \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right]\right) \cup \left(\left[\frac{2}{3}, \frac{2}{3} + \frac{1}{9}\right] \cup \left[\frac{2}{3} + \frac{2}{9}, \frac{2}{3} + \frac{1}{3}\right]\right) \\
&= \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right]\right) \cup \left(\left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]\right) \\
&= \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]
\end{aligned}$$

Note that $E_2 = E_1 \backslash \left(\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)\right)$.

### 1.1.3 A Few Properties

1. *Claim:* In general, $E_n$ will a union of $2^n$ intervals, each of which of length $3^{-n}$.
*Proof:*

   (a) We proceed by induction. $E_0$ is indeed a union of $2^0 = 1$ intervals (just the interval $[0, 1]$) and this interval has length $3^{-0} = 1$ indeed.

   (b) Let $n \in \mathbb{N}$ be given. Assume the statement holds for $E_{n-1}$. We want to show that it holds $E_n$.

   (c) $E_n \equiv \frac{1}{3}E_{n-1} \cup \left(\frac{1}{3}E_{n-1} + \frac{2}{3}\right)$. From this formula we see that because in $E_{n-1}$ we have the nion of $2^{n-1}$ intervals, and $E_n$ is a union of two disjoint copies of those intervals, in $E_n$ we have then exactly the union of $2^n$ intervals. In addition, because each interval in $E_{n-1}$ is of length $3^{-(n-1)}$, and in $E_n$ we are scaling all intervals by $\frac{1}{3}$, the intervals in $E_n$ will each be of length $\frac{1}{3} \times 3^{-(n-1)} = 3^{-n}$.

■

2. Thus, the length $E_n$ in total is $2^n \times 3^{-n} = \left(\frac{2}{3}\right)^n$, which approaches zero as $n \to \infty$. Thus, as we take intersections of larger and larger $n$, the total length of the intervals in $E_n$ is smaller and smaller, and indeed, the "length" of $P$ is zero, at least from one mathematical perspective called "measure theory". In light of this, we may ask, is something left at all after we remove some many middle thirds? This is the subject of this exercise.

3. *Claim:* $E_n \supset E_{n+1}\ \forall n \in \mathbb{N}$.
*Proof:*

   - Observe that

$$\begin{aligned}
E_n &\equiv \frac{1}{3}E_{n-1} \cup \left(\frac{1}{3}E_{n-1} + \frac{2}{3}\right) \\
&= E_{n-1} \backslash \left(\bigcup_{k=0}^{3^{n-1}-1} \left(\frac{3k+1}{3^n}, \frac{3k+2}{3^n}\right)\right)
\end{aligned}$$

as we've seen in the two examples above.

- Because in every step we are removing the open middle third of each sub-interval, we never add any points on top of the previous step and so every step is a subset of its predecessor.

## 1.2 Ternary Representation of the Cantor Set Numbers

- As you may well know, we can represent any real number in $[0, 1]$ in any base system we want–it doesn't have to be base 10. Thus, instead of writing $(0.5)_{10}$, we may write $(0.1)_2$ because $\frac{1}{2} = 1 \times 2^{-1}$, and $(0.75)_{10} = (0.11)_2$ because $0.75 = \frac{1}{2} + \frac{1}{4} = 1 \times 2^{-1} + 1 \times 2^{-2} = (0.11)_2$ and $(0.011)_2 = 0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3} = \frac{1}{4} + \frac{1}{8} = \frac{3}{8} = (0.375)_{10}$.
- As you may *also* know, decimal representation of fractions is not unique: $0.9999\ldots \equiv 1$.
- This problem also exists in ternary representation: $(0.111111\ldots)_3 = (0.2)_3$, so there might be two possible representations for each ternary represented number.
- *Claim:* $P = \left\{ x \in [0, 1] \,\middle|\, \exists a : \mathbb{N} \to \{0, 2\} \text{ such that } x = \sum_{n\in\mathbb{N}} \frac{a(n)}{3^n} \right\}$. (That is, the numbers in $P$ are exactly those numbers between 0 and 1 which can be written without the digit 1 in their ternary representation.
  *Proof:*
  - At each step of the Cantor construction, we are removing the open middle third interval.
  - Proof by induction on $n \in \mathbb{N}$ in $E_n$ (because $P = \bigcap_{n\in\mathbb{N}} E_n$, any statement that holds for all $E_n$ holds for $P$). In particular, the induction statement is that in $E_n$ we have removed exactly all numbers where the $n$-th digit necessarily starts with 1 (there is no other representation for that number).
  - For the $n = 1$ case:
    * In the very first step, when going from $[0, 1]$ to $\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$, we are removing $\left(\frac{1}{3}, \frac{2}{3}\right)$.
    * But those numbers in this interval we are removing are exactly those which *must* have the first digit as 1: $\left(\frac{1}{3}\right)_{10} = (0.1)_3$ and $\left(\frac{2}{3}\right)_3 = (0.2)_3$. So we are removing all the numbers between $(0.1)_3$ and $(0.2)_3$, which are exactly all the numbers that *start* with $(0.1)_3$:
      $(0.11)_3$, $(0.12)_3$, $(0.10000001)_3$ and so on.
    * Note that we are not removing the actual end points and this is fine because they either don't contain the digit 1 at all, or they can take on another representation: $(0.1)_3 = (0.02222\ldots)_3$ and $(0.2)_3$ is not removed, even though it can also be represented by $(0.1111\ldots)_3$: recall, we are trying to prove that numbers that *must* be represented by the first digit being 1 are being removed, not numbers that may have (in addition to others) such a representation.
  - Let $n \in \mathbb{N}$ be given. Assume the statement is true for $E_{n-1}$, as we have seen above,

$$
\begin{aligned}
E_n &= E_{n-1} \setminus \left( \bigcup_{k=0}^{3^{n-1}-1} \left( \frac{3k+1}{3^n}, \frac{3k+2}{3^n} \right)_{10} \right) \\
&= E_{n-1} \setminus \left( \bigcup_{k=0}^{3^{n-1}-1} \left( \underbrace{k \times 3^{-n+1}}_{\text{in the first } n-1 \text{ digits}} + \underbrace{1 \times 3^{-n}}_{\text{in the last digit}}, k \times 3^{-n+1} + \underbrace{2 \times 3^{-n}}_{\text{in the last digit}} \right)_{10} \right)
\end{aligned}
$$

  so we see that the intervals we are removing, $\left(\frac{3k+1}{3^n}, \frac{3k+2}{3^n}\right)_{10}$, are of the form $(0.a_1a_2\ldots a_{n-1}1, 0.a_1a_2\ldots a_{n-1}2)_3$. This is again the middle third of something portion of $[0, 1]$ and so we see that we are removing all numbers that start with $(0.a_1a_2\ldots a_{n-1}1)_3$:

$(0.a_1a_2\ldots a_{n-1}101)_3$, $(0.a_1a_2\ldots a_{n-1}10002)_3$, $(0.a_1a_2\ldots a_{n-1}111111)_3$ and so on. Again, we are *not* removing the end points. In particular, $(0.a_1a_2\ldots a_{n-1}1)_3$ is *not* removed because it can also be represented *without* 1: $(0.a_1a_2\ldots a_{n-1}1)_3 = (0.a_1a_2\ldots a_{n-1}2222222\ldots)_3$.

  - As a result the statement is true for all $n \in \mathbb{N}$: $P$ contains no numbers whose representations *must* start with 1 at some digit. This is exactly what we wanted to prove.

$\blacksquare$

- This ternary representation is quite cool because it immediately reveals that not only is the Cantor set not empty: it contains much more than the end points which we never touch: it contains all numbers whose ternary representation can avoid the digit 1, and as we shall see, this includes "just as many" (cardinality wise) numbers as are initially in $[0,\,1]$.

# 2 Map Onto $[0,\,1]$

- Define a map $\varphi : P \to [0,\,1]$ by the following rule: $a \mapsto \sum_{n=1}^{\infty} \frac{a(n)}{2^{n+1}}$ where $a$ is the representation of $x$ as a function $a : \mathbb{N} \to \{0,\,2\}$ which was exactly defined above. Note that this is a cool map: it maps maps into numbers. (So far we have seen maps that map numbers to numbers. $\varphi$ maps functions ($a : \mathbb{N} \to \{0,\,2\}$) to numbers.)

- *Claim:* $\varphi$ is well-defined.
  *Proof:*

  - In order to show that $\varphi$ is well-defined, we must take two representations of the same element of $P$, let's call it $x \in P$, and the two representaions $a : \mathbb{N} \to \{0,\,2\}$ and $b : \mathbb{N} \to \{0,\,2\}$, and show that necessarily $\varphi(a) = \varphi(b)$. However, because we have restricted ourselves to representations which don't include the digit 1, our representations of $x$ are in fact all unique: the only ambiguities we could have are a sequence of 2's which becomes a 1, but that 1 representation we don't include! Thus if $a$ and $b$ correspond to the same $x$, then $a(n) = b(n)$ for all $n \in \mathbb{N}$ and thus $\varphi(a) = \varphi(b)$.

$\blacksquare$

- *Claim:* $\varphi$ is surjective.
  *Remark:* As a result, we must have that $|P| \geq [0,\,1]$! But $P \subset [0,\,1]$ and so $|P| \leq [0,\,1]$, that is, we will have shown that $|P| = [0,\,1]$.
  *Proof:*

  - Let $y \in [0,\,1]$ be given.
  - Observe that our map is actually $a \mapsto \sum_{n=1}^{\infty} \frac{\left(\frac{a(n)}{2}\right)}{2^n}$ and that $a(n) \in \{0,\,2\}$ so that the numerator is always either 0 or 1, that is, we arrive at fractions in $[0,\,1]$ with binary representation.
  - So given $y \in [0,\,1]$, let $b : \mathbb{N} \to \{0,\,1\}$ be the binary representation of $y$. (That one always exists should be clear). Then $a(n) = 2 \times b(n)$ should do the job, and of course, $a$ defined in this way is indeed a member of the Cantor set according to our rules.

$\blacksquare$

- *Claim:* $\varphi$ is monotonically increasing.
  *Proof:*

  - Let $(x,\,y) \in P^2$ be given, such that $x \leq y$.

- Assume that the ternary representation of $x$ is given by $a$ and the ternary representaion of $y$ is given by $b$.
- Because $x \leq y$, is equivalent to saying that $\exists n \in \mathbb{N}$ such that $a(j) = b(j)$ for all $j \in \{1, \ldots, n-1\}$ and that $a(n) \leq b(n)$.
- But this, in turn, implies that $\frac{1}{2}a(n) \leq \frac{1}{2}b(n)$, because $x \mapsto \frac{1}{2}x$ is monotonically increasing.
- This implies that $\sum_{n=1}^{\infty} \frac{\left(\frac{a(n)}{2}\right)}{2^n} \leq \sum_{n=1}^{\infty} \frac{\left(\frac{b(n)}{2}\right)}{2^n}$ which is $\varphi(a) \leq \varphi(b)$.
  $\blacksquare$

## 2.1 Homework

- *Claim:* $\varphi$ is continuous.
  *Proof:* (can be found in Koenigsberger on page 377)