MATH 201 3.1 - DISCUSSION: THE CANTOR SET

- Today I get to share one of my favorite sets with you!
- This set was discovered by Georg Cantor, the same guy who proved that $\mathbb{R}$ is uncountable.

Definition 1. Let C_0 be the closed interval $[0, 1]$.

(1) Define C_1 to be the set that results when the open middle third is removed; that is,

$$C_1 = C_0 \setminus \left(\frac{1}{3}, \frac{2}{3}\right) = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

[draw]

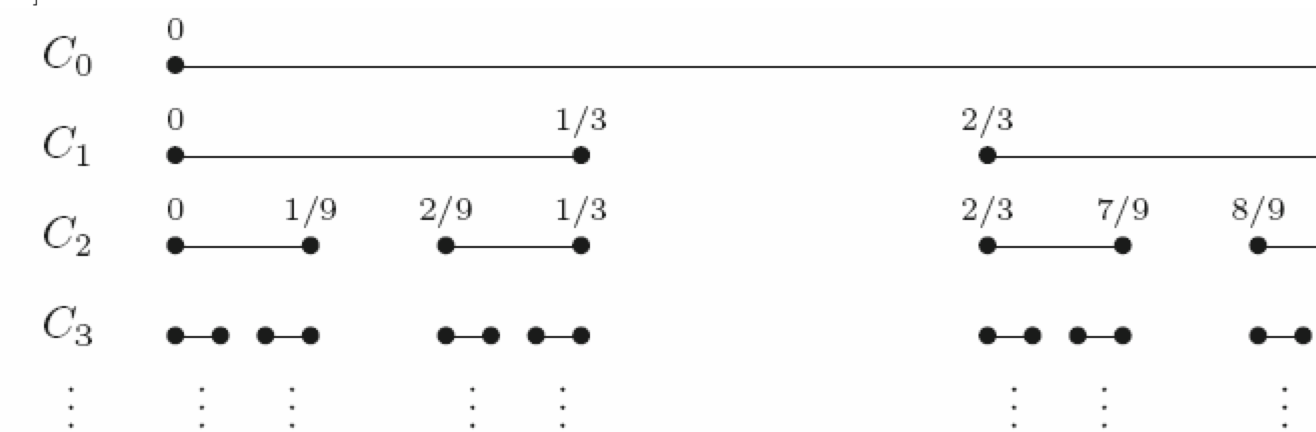

Figure 3.1: Defining the Cantor set; $C = \bigcap_{n=0}^{\infty} C_n$.

(2) Next, construct C_2 in a similar way by removing the open middle third of each of the two components of C_1:

$$C_2 = \left(\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right]\right) \cup \left(\left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]\right).$$

(3) Continue this process inductively, so for $n = 0, 1, 2, \ldots$ we get a set C_n consisting of 2^n closed intervals each having length $1/3^n$.

(4) We define the **Cantor set** C to be the intersection

$$C = \bigcap_{n=0}^{\infty} C_n.$$

Exercise 2. (reading question)

(1) Define the "length" of the Cantor set to be the length of the interval $[0, 1]$ minus the total lengths of all intervals removed to construct C. What is the length of C? Explain your answer.

By definition,

$$C = [0, 1] \setminus \left[\left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \left(\frac{1}{27}, \frac{2}{27}\right) \cup \left(\frac{4}{27}, \frac{5}{27}\right) \cup \left(\frac{19}{27}, \frac{20}{27}\right) \cup \left(\frac{25}{27}, \frac{26}{27}\right) \ldots\right].$$

The total length of all the removed intervals is

$$\begin{aligned}\frac{1}{3}+2\Big(\frac{1}{9}\Big)+4\Big(\frac{1}{27}\Big)+\cdots+2^{n-1}\Big(\frac{1}{3^n}\Big)+\ldots &= \sum_{n=1}^{\infty}\frac{2^{n-1}}{3^n}\\ &= \frac{1}{3}\sum_{n=0}^{\infty}\Big(\frac{2}{3}\Big)^n\\ &= \frac{1}{3}\Big(\frac{1}{1-\frac{2}{3}}\Big)=1.\end{aligned}$$

Therefore, the "length" of the Cantor set is $1-1=0$!

(2) Is the Cantor set empty? If so, prove it. If not, name two elements that are contained in the Cantor set.

By the way the Cantor set is defined, we're only removing **open** intervals. Therefore, $0,1 \in C$.

- In fact, if y is the endpoint of some closed interval of some particular set C_n, then it is also an endpoint of one of the intervals of C_{n+1}. Because, at each stage, endpoints are never removed, it follows that $y \in C_n$ for all n.
- Thus, C contains at least the endpoints of all the intervals that make up each of the sets C_n.

Question 3. *What other questions might we ask about the Cantor set?*

- *Is C countable?*
- *Does C contain any intervals? [can't; it's 0 length]*
- *Are there points of C that aren't endpoints of removed intervals?*
- *Any irrational numbers?*

- Right now, it seems we're only left with "Cantor dust"—a small, thin set with a bunch of points but nothing else in it.

Theorem 4. *C is* uncountable*, with cardinality equal to the cardinality of $\mathbb{R}$.*

- We'll create a 1-to-1 correspondence between C and the set S of sequences of 0s and 1s, which we learned earlier has cardinality equal to that of $\mathbb{R}$.

Exercise 5. How could we mimic the proof of the fact that $S \sim (0,1)$ or the Bolzano-Weierstrass Theorem to show that $C \sim S$?

Lemma 6. *The Cantor set $C \sim S$, where S is the set of all sequences of 0s and 1s.*

Proof. (Lemma) For each $c \in C$, set $a_1 = 0$ if c falls in the left-hand component of C_1 and $a_1 = 1$ if c falls in the right-hand component.

Now, there are two possible components of C_2 that contain C. Set $a_2 = 0$ if c is in the left half of either of these two components of C_2 and 1 if it's in the right half.

In general, every element $c \in C$ yields a sequence $(a_1, a_2, a_3, \ldots)$ of zeroes and ones that serve as "directions" for how to locate c within C.

Exercise 7. Give the sequence of "directions" to locate $\frac{1}{3}$ in C.

- $\frac{1}{3} \mapsto (0,1,1,1,1,1,1,1,\ldots)$ ("binary expansion")

Define a function $f : C \to S$ via setting $f(c)$ equal to the set of "directions" for locating c within C.

- f is 1-1: if $f(c_1) = f(c_2)$, then at each stage of the process of subdividing $[0,1]$, c_1 and c_2 are in the same component. This means that c_1, c_2 are both $\leq \frac{1}{3}$ or both $\geq \frac{2}{3}$. Assume they're both $\leq \frac{1}{3}$, they are both $\leq \frac{1}{9}$ or $\geq \frac{2}{9}$. In general, for all n, c_1 and c_2 are both in the same interval of length $\frac{1}{3^n}$. Since $\frac{1}{3^n} \to 0$, for any $\epsilon > 0$, there exists N such that
$$|c_2 - c_1| < \left|\frac{1}{3^n} - 0\right| < \epsilon.$$
Thus, $c_1 = c_2$.
- f is onto: suppose (a_n) is a sequence of 0s and 1s. Then (a_n) gives us a set of "directions" for finding a $c \in C$ so that $f(c) = (a_n)$.

This shows that f is a 1-1 correspondence and that $C \sim S$, as desired. $\square$

Lemma 8. *The set S of sequences of 0s and 1s satisfies $S \sim (0,1)$ (and hence $S \sim \mathbb{R}$).*

Proof. We define 1-to-1 functions $f : S \to (0,1)$ and $g : (0,1) \to S$. Then, by the Bernstein-Schroeder Theorem, there exists a 1-to-1 correspondence $S \to (0,1)$.

To define f, we map each sequence $a = (a_1, a_2, \dots)$ to the decimal $f(a) = 0.a_1a_2\dots$. Then, since each $a_i \in \{0,1\}$, if $f(a) = f(b)$, it must be that $a = b$.

Exercise 9. Finish the proof of the Lemma by mimicking the proof that $C \sim S$ to define g. You'll have to give a set of "directions" for locating each point of $(0,1)$, then use those "directions" to map each $x \in (0,1)$ to a unique sequence in S.

To define g, given $x \in (0,1)$, we set the first term of $g(x)$ to be 0 if $x \in (0, \frac{1}{2}]$ and 1 if $x \in (\frac{1}{2}, 1)$. Then, if x is in the left half of the previous interval, we set the second term of $g(x)$ to be 0, and if x is in the right half of the previous interval, we set the second term of $g(x)$ to be 1. Since this set of "directions" $\square$

- This theorem implies the following:

Corollary 10. *There are uncountably many points of C that aren't endpoints of removed intervals. Moreover, C must contain uncountably many irrational numbers.*

Exercise 11. Prove it.

- Assume for contradiction that C contains finitely or countably many irrational numbers. Define $Q = \{r \in [0,1] : r \in \mathbb{Q}\}$ and $I = \{c \in C : c \in \mathbb{I}\}$. Since $\mathbb{Q}$ is countable and Q is an infinite subset of $\mathbb{Q}$, Q is countable as well. Moreover, $C = Q \cup I$ is the union of a countable set and a set that is either finite or countable. In either case, C must then be countable, a contradiction.
- Since all endpoints of removed intervals are rational numbers, C must contain uncountably many points that aren't endpoints of removed intervals.

Question 12. *Assuming a point has dimension 0, a line segment has dimension 1, a square has dimension 2, and a cube has dimension 3, what is the "dimension" of C?*

- Let's consider what happens when we magnify various sets by a factor of 3.

Exercise 13. How does the "size" (length, area, volume, etc.) of each of the point, line segment, square, and cube change when we multiply each side by a factor of 3? What pattern(s) do you see between these changes and the dimension of each of these sets?

- A point undergoes no change at all.
- A line segment triples in length.
- Magnifying each side of a square by 3 results in multiplying the area by a factor of 9.
- The "tripled cube" contains 27 copies of the original cube within its volume.
- In each case, to compute the number of copies of the old set contained in the new set, we take 3 to the power of the dimension:
  - point: $3^0 = 1$ copy
  - segment: $3^1 = 3$ copies
  - etc.

- What happens when we apply the "times 3" transformation to the Cantor set?
  - C_0 becomes $[0,3]$.
  - [draw] Deleting the middle third leaves $[0,1] \cup [2,3]$, which is where we started in the original construction except that we now stand to produce an additional copy of C in the interval $[2,3]$.
  - Magnifying the Cantor set by a factor of 3 yields 2 copies of the original set.
- Thus, if x is the dimension of C, x should satisfy $2 = 3^x$, or $x = \log 2/\log 3 \approx .631$.
- [slides] The notion of a fractional dimension is the impetus behind the term "fractal", coined in 1975 by Benoit Mandlebrot to describe a class of sets whose intricate structures have much in common with the Cantor set.
- In the coming class sessions, we'll use the Cantor set as a testing ground for the upcoming theorems and conjectures about the nature of subsets of the real line.

Example 14. How could we tweak the Cantor set to end up with some length?

Exercise 15. Next time, I'll ask you to think about a tweak that gives us a set with length.