Math 416
1/18/19

The Cantor Set Contains 1/4? Really?
Author(s): Sarah-Marie Belcastro and Michael Green
Source: *The College Mathematics Journal*, Vol. 32, No. 1 (Jan., 2001), pp. 55-56
Published by: Taylor & Francis, Ltd. on behalf of the Mathematical Association of America
Stable URL: https://www.jstor.org/stable/2687224
Accessed: 16-01-2019 18:50 UTC

The Cantor Set Contains $\frac{1}{4}$? *Really?*

sarah-marie belcastro (smbelcas@math.uni.edu), University of Northern Iowa, Cedar Falls, IA, 50614-0506, and Michael Green (Michael.Green@metrostate.edu) Metropolitan State University, St. Paul, MN 55106

The authors have been teaching undergraduate real analysis this year from Marsden's wonderful book [1]. Marsden introduces the Cantor set and many of its properties in an exercise. We find it beneficial for undergraduates to study the Cantor set because it is a good source of counterexamples for what might otherwise seem reasonable conjectures. Also, while the Cantor set is complicated, its definition is fairly simple and easy to understand. As do most undergraduate texts, Marsden uses the middle-thirds definition of the Cantor set.

Definition. Let $F_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, $F_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ and construct F_{k+1} from F_k by deleting the middle third of each interval in F_k. Then the *Cantor set* $C = \bigcap_{n=1}^{\infty} F_n$.

After our students had been thinking about the Cantor set for a couple of weeks, we heard them make the following sort of comment: "In the limit, isn't it true that all points in the Cantor set are endpoints of those intervals?" ... and *then* ... "Our intuition is probably wrong, so there have to be other points in the Cantor set. But they must be some kind of really weird irrational points."

It's easy to see why beginning analysis students would draw such conclusions. After all, if you get rid of the middle third of *every* interval, you should not have *any* intervals in the Cantor set, and the only points left are those of the form $\frac{m}{3^n}$, right?

Wrong.

Analyzing the situation. We immediately knew that this intuition was incorrect—after all, the Cantor set is uncountable, while the set of "the endpoints" is countable. When informed of this, students might naturally ask "So, can you name some point in C other than the endpoints?" As do most mathematicians, when thinking analytically about the Cantor set we use the fact that C is equivalent to the set of points in the interval $[0, 1]$ whose ternary expansion can be written using only the digits 0 and 2. (This is given in most graduate texts such as [2] and [3].) Using this, it is immediate that C is uncountable, and relatively easy to see that there are many rationals other than "the endpoints" in the Cantor set. In particular, the ternary expansion of $\frac{1}{4}$ is $.0202020202\ldots$. We believe that this fact is important! First of all, it is surprising to beginning analysis students (at least it was to our students); secondly, it enhances their intuition and understanding of the Cantor set. Finally, it's just plain cool!

Who knows that $\frac{1}{4}$ is in the Cantor Set? We began to wonder:

- Do most mathematicians know this?
- If so, when do they learn it?
- Do many undergraduates learn this fact?
- How can we convey this to our students without resorting to the not-so-intuitive ternary expansion?

An informal survey (sample size 12) caused us to conclude that most mathematicians either did not know or did not remember this fact (9/12). (Be honest. Did *you* know it?) However, most mathematicians can prove that $\frac{1}{4}$ is in the Cantor set within seconds using the ternary expansion. Generally (12/12) mathematicians discover this fact in graduate school or later.

We surveyed approximately twenty undergraduate analysis texts and found that only one (very recent) text [4] mentioned that $\frac{1}{4}$ is in the Cantor set, and only after having first introduced the ternary expansion. But we can avoid wrangling with the complexities of the ternary expansion!

How we like to present this to our students. We can convince our students that $\frac{1}{4}$ could be in C, by illustrating successive iterations of removing middle-thirds, as in Figure 1:

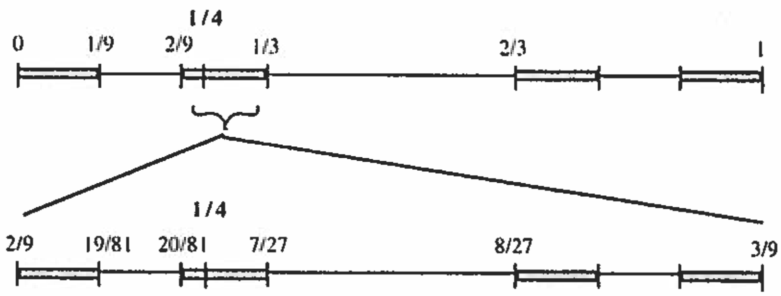

Figure 1.

In the next three iterations, $\frac{1}{4}$ appears in $[\frac{182}{729}, \frac{61}{243}]$, $[\frac{1640}{6561}, \frac{547}{2187}]$, and $[\frac{14762}{59049}, \frac{4921}{19683}]$—exactly one-fourth of the way from the left endpoint to the right endpoint.

What remains is an analytic proof. We use the elegant expansion $\frac{1}{4} = \frac{1}{3} - \frac{1}{3^2} + \frac{1}{3^3} - \frac{1}{3^4} + \ldots$. First, it is clear that the right-hand side of the equality represents an element in the Cantor set, for each partial sum corresponds to travelling to a successive interval endpoint (all of which are in each set in the infinite intersection that is the Cantor set). Figure 2 shows a graphical interpretation.

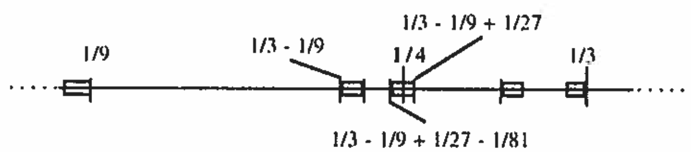

Figure 2.

Then, we can see that the series converges to $\frac{1}{4}$ by examining the geometric series formula $\frac{1}{4} = \frac{A}{(1-r)}$, where in this case $A = \frac{1}{3}$ and $r = \frac{-1}{3}$.

This idea may be introduced by asking students whether there are rationals in the Cantor set other than those of the form $\frac{m}{3^n}$, and later suggesting that perhaps there *are* other rationals, even obvious ones. That leads to showing that $\frac{1}{4}$ is in C. We believe that this beautiful, simple, and somewhat suprising fact should be a standard example in any undergraduate analysis course.

References

1. J. E. Marsden and M. J. Hoffmann, *Elementary Classical Analysis*, W. H. Freeman, 1993.
2. H. L. Royden, *Real Analysis*, Macmillan, 1988.
3. W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, 1976.
4. F. Dangello and M. Seyfried, *Introductory Real Analysis*, Houghton Mifflin, 2000.

Cantor Set

The **Cantor set** is set of points lying on a line segment. It is created by taking some interval, for instance $[0,1]$, and removing the middle third $\left(\frac{1}{3}, \frac{2}{3}\right)$, then removing the middle third of each of the two remaining sections $\left(\frac{1}{9}, \frac{2}{9}\right)$ and $\left(\frac{7}{9}, \frac{8}{9}\right)$, then removing the middle third of the remaining four sections, and so on ad infinitum.

It is a closed set consisting entirely of boundary points, and is an important counterexample in set theory and general topology. When learning about cardinality, one is first shown subintervals of the real numbers, $\mathbb{R}$, as examples of uncountably infinite sets. This might suggest that any uncountably infinite subset of $\mathbb{R}$ must contain an interval; such an assertion is *false*. A counterexample to this claim is the Cantor set $C \subset [0,1]$, which is uncountable despite not containing any intervals. In addition, Cantor sets are uncountable, may have 0 or positive Lebesgue measures, and are nowhere dense. Cantor sets are the only disconnected, perfect, compact metric space up to a homeomorphism.

Contents

Construction

The Cantor set is constructed by removing increasingly small subintervals from $[0,1]$. In the first step, remove $\left(\frac{1}{3}, \frac{2}{3}\right)$ from $[0,1]$. In the second step, remove $\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)$ from what remains after the first step. In general, on the n^{th} step, remove

$$\left(\frac{1}{3^n}, \frac{2}{3^n}\right) \cup \left(\frac{4}{3^n}, \frac{5}{3^n}\right) \cup \cdots \cup \left(\frac{3^n-2}{3^n}, \frac{3^n-1}{3^n}\right)$$

from what remains after the $(n-1)^{\text{th}}$ step. After all $\mathbb{N}$ steps have been taken, what remains is the Cantor set C.

This construction can be formalized as follows. Let $C_0 = [0,1]$. Then, after the first step, what remains is $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. After the second step, what remains is

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

Continuing in this manner, one obtains an infinite collection of sets $\{C_i\}_{i=0}^{\infty}$ such that $C_i \subset C_{i-1}$ for all $i \geq 1$. Then, one defines the Cantor set to be

$$C = \bigcap_{i=0}^{\infty} C_i.$$

Intuitively, this makes sense; since $\bigcap_{i=0}^{n} C_i = C_n$ for all $n \geq 1$, taking an infinite intersection would provide the "limiting set" in this situation.

After seven iterations of the Cantor set's construction. Going from top to bottom, the image depicts the sets C_0 through C_6. [1]

Properties

There is an alternate characterization that will be useful to prove some properties of the Cantor set:

THEOREM

C consists precisely of the real numbers in $[0, 1]$ whose base-3 expansions only contain the digits 0 and 2.

Base-3 expansions, also called ternary expansions, represent decimal numbers on using the digits $0, 1, 2$. For instance, the number 3 in decimal is 10 in base-3. Fractions in base-3 are represented in terms of negative powers of 3, where a fraction N in decimal can be represented as

$$N = c_1 b^{-1} + c_2 b^{-2} + c_3 b^{-3} + \cdots + c_n b^{-n} + \cdots,$$

where $0 \leq c_i < 3$, and the fraction is written in base-3 as $N = 0.c_1 c_2 c_3 \ldots$.
For instance, $\frac{1}{4} = 0.25 = 0 \times 3^{-1} + 2 \times 3^{-2} + 0 \times 3^{-3} + 2 \times 3^{-4} + \cdots = 0.\overline{02}_3$.

Similarly, decimal fractions can be converted by multiplying by 3, taking the result mod 3, then multiplying the remainder by 3, and then taking mod 3 of that... etc. until the remainder is 0:

$$\begin{aligned} 0.25 \times 3 &= \mathbf{0} + 0.75, & c_1 &= 0 \\ 0.75 \times 3 &= \mathbf{2} + 0.25, & c_2 &= 2 \\ 0.25 \times 3 &= \mathbf{0} + 0.75, & c_3 &= 0 \\ 0.75 \times 3 &= \mathbf{2} + 0.25, & c_4 &= 2 \\ & & &\vdots \\ & & &\Rightarrow 0.0202\ldots_3. \end{aligned}$$

In base-3, some points have more than one notation:
$\frac{1}{3} \times 3 = \mathbf{1} + 0,\ c_1 = 1$. So $\frac{1}{3} = 0.1_3$. But we also have the following:

$$\begin{aligned} \frac{1}{3} \times 3 = 0.3333\ldots \times 3 &= \mathbf{0} + 0.9999\ldots, & c_1 &= 0 \\ 0.9999\ldots \times 3 &= \mathbf{2} + 0.9999\ldots, & c_2 &= 2 \\ 0.9999\ldots \times 3 &= \mathbf{2} + 0.9999\ldots, & c_3 &= 2 \\ & & &\vdots \end{aligned}$$

So, $\frac{1}{3} = 0.\overline{3} = 0.1_3 = 0.0\overline{2}_3$. This sometimes means there is ambiguity in C, for if C expansions only contain the digits 0 and 2, then C contains $0.0\overline{2}_3$ but not 0.1_3 even though they are the same. In this case, it's said to contain $0.0\overline{2}_3 = \frac{1}{3}$, which implies it contains 0.1_3 because of the multiple notations.

PROOF

Suppose $x \in [0, 1]$ contains only the digits 0 and 2 in its base-3 expansion. Let x_n be the truncation of x at n places after the decimal point. For example, if $x = 0.020202\cdots_3$, then $x_1 = 0$, $x_2 = 0.02_3$, $x_4 = 0.0202_3$, etc. Certainly, the sequence x_n converges to x as $n \to \infty$. In particular, for every $n \geq 1$, we have

$$x_n \leq x \leq x_n + \frac{1}{3^n}.$$

Note that the numbers in $[0, 1]$

- whose base-3 expansions go on for exactly n digits after the decimal point and
- which use only the digits 0 and 2

are precisely the *left* endpoints of the intervals whose union is C_n. Thus, the interval $\left[x_n, x_n + \frac{1}{3^n}\right]$ is contained in C_n. It follows that $x \in C_n$ for all $n \geq 0$, and hence $x \in C$. □

Conversely, suppose $x \in C$. Then $x \in C_n$ for all $n \geq 0$. Note that the numbers in C_n are precisely those whose n^{th} truncation (i.e., the number obtained by taking only the first n digits after the decimal point) uses only 0 and 2 as digits in base 3. It follows that every truncation of x uses only 0 and 2 as digits. This implies x uses only 0 and 2 in its base-3 expansion. □

TRY IT YOURSELF

Of the given answer choices, which number is in the Cantor set?

○ $\frac{107}{364}$

○ $\frac{111}{364}$

From this theorem, the proofs of the following two properties follow:

- ◯ $\frac{113}{364}$
- ◯ $\frac{115}{364}$

THEOREM

C does not contain any subintervals of $[0,1]$.

PROOF

Let $[a,b] \subset [0,1]$ be an arbitrary interval. Write $a = 0.a_1a_2a_3\cdots_3$ and $b = 0.b_1b_2b_3\cdots_3$. If some a_i equals 1, then by the previous theorem, we know $a \notin C$. Similarly, if some b_i equals 1, we know $b \notin C$. Otherwise, suppose k is the smallest index such that $a_k \neq b_k$. Our casework forces $a_k = 0$ and $b_k = 2$, so $0.a_1a_2a_3\cdots a_{k-1}1_3 \in [a,b]$, and it follows that $[a,b] \not\subset C$.

□

THEOREM

C is uncountable.

PROOF

Define a function $f: C \to [0,1]$ as follows. If $x = 0.x_1x_2x_3\cdots_3$ is an element in C, set

$$f(x) = 0.(x_1/2)(x_2/2)(x_3/2)\cdots_2.$$

In other words, take the ternary expansion of x, replace every digit 2 with 1, and consider the result as a binary number.

This function is surjective, since any element $y = 0.y_1y_2y_3\cdots_2$ of $[0,1]$ has $f\big(0.(2y_1)(2y_2)(2y_3)\cdots_3\big) = y$.

Now, suppose C is countable. Write $C = \{c_1, c_2, c_3, \cdots\}$. For each $y \in [0,1]$, let $g(y)$ equal the index j of an element c_j with $f(c_j) = y$ (remark: in defining this function g, we are implicitly using the axiom of choice). Then, we may construct a sequence $\{d_i\}_{i\in\mathbb{N}}$ such that $d_{g(y)} = y$ for all $y \in [0,1]$. This constitutes a bijection between $[0,1]$ and $\mathbb{N}$, so it follows that $[0,1]$ is countable. Contradiction!

Hence, we conclude C is uncountable. □

TRY IT YOURSELF

Let $C + C$ denote the sumset of the Cantor set C. That is,

$$C + C = \{x + y : x, y \in C\}.$$

Which of the following statements is true about $C + C$?

- ◯ $[0,2]$ strictly contains $C+C$
- ◯ $C+C$ equals $[0,2]$
- ◯ $C+C$ strictly contains $[0,2]$

References

1. 127 "rect", . *Cantor ternary set, in seven iterations*. Retrieved September 5th, 2016, from https://commons.wikimedia.org/wiki/File:Cantor_set_in_seven_iterations.svg